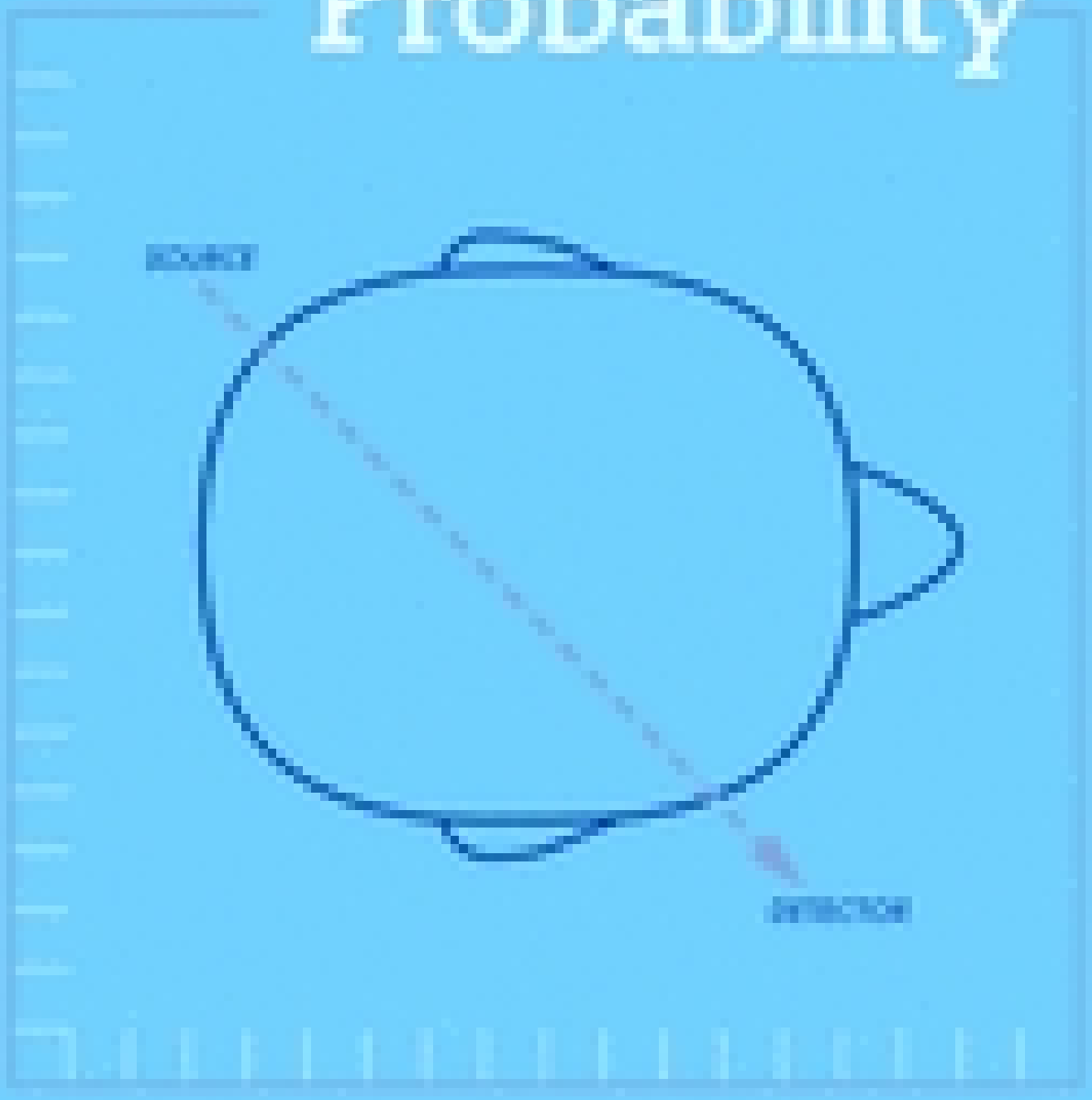
SPRINGER TEXTS IN STATISTICS

Applied Probability



Kenneth Lange

Springer

Kenneth Lange

Applied Probability

Second Edition

Springer

Prof. Kenneth Lange
University of California, Los Angeles
Departments of Biomathematics,
Human Genetics, and Statistics
90095-1766 Los Angeles
California
USA
klange@ucla.edu

ISSN 1431-875X
ISBN 978-1-4419-7164-7 e-ISBN 978-1-4419-7165-4
DOI 10.1007/978-1-4419-7165-4
Springer New York Dordrecht Heidelberg London

Library of Congress Control Number: 2010933109

Preface to the Second Edition

My original intent in writing *Applied Probability* was to strike a balance between theory and applications. Theory divorced from applications runs the risk of alienating many potential practitioners of the art of stochastic modeling. Applications without a clear statement of relevant theory drift in a sea of confusion. To a lesser degree I was also motivated by a desire to promote the nascent field of computational probability. Current students of the mathematical sciences are more computer savvy than ever. Putting the right computational tools in their hands is bound to advance probability and the broader good of science.

The second edition of *Applied Probability* remains true to these aims. I have added two new chapters on asymptotic and numerical methods and an appendix that separates some of the more delicate mathematical theory from the steady flow of examples in the main text. In addition to these major changes, there is now a much more extensive list of exercises. Some of these are trivial, but others will challenge even the best students. Finally, many errors, both large and small, have been corrected.

Chapter 4 on combinatorics includes new sections on bijections, Catalan numbers, and Faà di Bruno's formula. The proof of the inclusion-exclusion formula has been clarified. Chapter 7 on Markov chains contains new material on rates of convergence to equilibrium in reversible finite-state chains. This discussion draws on students' previous exposure to eigenvalues and eigenvectors in linear algebra. Chapter 9 on branching processes features a new section on basic reproduction numbers. Here the idea is to devise easy algebraic tests for deciding when a process is subcritical, critical, or supercritical. Chapter 11 on diffusion processes gives better coverage of

Brownian motion. The last two sections of the chapter have been moved to the new Chapter 13 on numerical methods. The orphan material on convergent sequences of random variables in Chapter 1 has been moved to the new Chapter 12 on asymptotic methods.

Once again I would like to thank the students of my UCLA biomathematics classes for their help. Particularly noteworthy are David Alexander, Kristin Ayers, Forrest Crawford, Kate Crespi, Gabriela Cybis, Lewis Lee, Sarah Nowak, John Ranola, Mary Sehl, Tongtong Wu, and Jin Zhou. I owe an especially heavy debt to Hua Zhou, my former postdoctoral fellow, for suggesting many problems and lecturing in my absence. I also thank my editor, John Kimmel, for his kind support. Finally, I am glad to report that my mother, to whom both editions of this book are dedicated, is alive and well. If I can spread even a fraction of the cheer she has spread, then I will be able to look back over a life well lived.

Preface

Despite the fears of university mathematics departments, mathematics education is growing rather than declining. But the truth of the matter is that the increases are occurring outside departments of mathematics. Engineers, computer scientists, physicists, chemists, economists, statisticians, biologists, and even philosophers teach and learn a great deal of mathematics. The teaching is not always terribly rigorous, but it tends to be better motivated and better adapted to the needs of students. In my own experience teaching students of biostatistics and mathematical biology, I attempt to convey both the beauty and utility of probability. This is a tall order, partially because probability theory has its own vocabulary and habits of thought. The axiomatic presentation of advanced probability typically proceeds via measure theory. This approach has the advantage of rigor, but it inevitably misses most of the interesting applications, and many applied scientists rebel against the onslaught of technicalities. In the current book, I endeavor to achieve a balance between theory and applications in a rather short compass. While the combination of brevity and balance sacrifices many of the proofs of a rigorous course, it is still consistent with supplying students with many of the relevant theoretical tools. In my opinion, it is better to present the mathematical facts without proof rather than omit them altogether.

In the preface to his lovely recent textbook [209], David Williams writes, "Probability and Statistics used to be married; then they separated; then they got divorced; now they hardly see each other." Although this split is doubtless irreversible, at least we ought to be concerned with properly bringing up their children, applied probability and computational statistics.

If we fail, then science as a whole will suffer. You see before you my attempt to give applied probability the attention it deserves. My other recent book [122] covers computational statistics and aspects of computational probability glossed over here.

This graduate-level textbook presupposes knowledge of multivariate calculus, linear algebra, and ordinary differential equations. In probability theory, students should be comfortable with elementary combinatorics, generating functions, probability densities and distributions, expectations, and conditioning arguments. My intended audience includes graduate students in applied mathematics, biostatistics, computational biology, computer science, physics, and statistics. Because of the diversity of needs, instructors are encouraged to exercise their own judgment in deciding what chapters and topics to cover.

Chapter 1 reviews elementary probability while striving to give a brief survey of relevant results from measure theory. Poorly prepared students should supplement this material with outside reading. Well-prepared students can skim Chapter 1 until they reach the less well-known material of the final two sections. Section 1.8 develops properties of the multivariate normal distribution of special interest to students in biostatistics and statistics. This material is applied to optimization theory in Section 3.3 and to diffusion processes in Chapter 11.

We get down to serious business in Chapter 2, which is an extended essay on calculating expectations. Students often complain that probability is nothing more than a bag of tricks. For better or worse, they are confronted here with some of those tricks. Readers may want to skip the final two sections of the chapter on surface area distributions on a first pass through the book.

Chapter 3 touches on advanced topics from convexity, inequalities, and optimization. Besides the obvious applications to computational statistics, part of the motivation for this material is its applicability in calculating bounds on probabilities and moments.

Combinatorics has the odd reputation of being difficult in spite of relying on elementary methods. Chapters 4 and 5 are my stab at making the subject accessible and interesting. There is no doubt in my mind of combinatorics' practical importance. More and more we live in a world dominated by discrete bits of information. The stress on algorithms in Chapter 5 is intended to appeal to computer scientists.

Chapters 6 through 11 cover core material on stochastic processes that I have taught to students in mathematical biology over a span of many years. If supplemented with appropriate sections from Chapters 1 and 2, there is sufficient material here for a traditional semester-long course in stochastic processes. Although my examples are weighted toward biology, particularly genetics, I have tried to achieve variety. The fortunes of this book doubtless will hinge on how compelling readers find these examples.

You can leaf through the table of contents to get a better idea of the topics covered in these chapters.

In the final two chapters, on Poisson approximation and number theory, the applications of probability to other branches of mathematics come to the fore. These chapters are hardly in the mainstream of stochastic processes and are meant for independent reading as much as for classroom presentation.

All chapters come with exercises. (In this second printing, some additional exercises are included at the end of the book.) These are not graded by difficulty, but hints are provided for some of the more difficult ones. My own practice is to require one problem for each hour and a half of lecture. Students are allowed to choose among the problems within each chapter and are graded on the best of the solutions they present. This strategy provides an incentive for the students to attempt more than the minimum number of problems.

I would like to thank my former and current UCLA and University of Michigan students for their help in debugging this text. In retrospect, there were far more contributing students than I can possibly credit. At the risk of offending the many, let me single out Brian Dolan, Ruzong Fan, David Hunter, Wei-hsun Liao, Ben Redelings, Eric Schadt, Marc Suchard, Janet Sinsheimer, and Andy Ming-Ham Yip. I also thank John Kimmel of Springer–Verlag for his editorial assistance.

Finally, I dedicate this book to my mother, Alma Lange, on the occasion of her 80th birthday. Thanks, Mom, for your cheerfulness and generosity in raising me. You were, and always will be, an inspiration to the whole family.

Contents

1
Basic Notions of Probability Theory

1.1 Introduction

This initial chapter covers background material that every serious student of applied probability should master. In no sense is the chapter meant as a substitute for a previous course in applied probability or for a future course in measure-theoretic probability. Our comments are merely meant as reminders and as a bridge. Many mathematical facts will be stated without proof. This is unsatisfactory, but it is even more unsatisfactory to deny students the most powerful tools in the probabilist's toolkit. Quite apart from specific tools, the language and intellectual perspective of modern probability theory also furnish an intuitive setting for solving practical problems. Probability involves modes of thought that are unique within mathematics. As a brief illustration of the material reviewed, we derive properties of the multivariate normal distribution in the final section of this chapter. Later chapters will build on the facts and vocabulary mentioned here and provide more elaborate applications.

1.2 Probability and Expectation

The layman's definition of probability is the long-run frequency of success over a sequence of independent, identically constructed trials. Although this law of large numbers perspective is important, mathematicians have found it helpful to put probability theory on an axiomatic basis [24, 53, 60, 80,

K. Lange, *Applied Probability,*
DOI 10.1007/978-1-4419-7165-4_1, © Springer Science+Business Media, LLC 2010

166, 171, 208]. The modern theory begins with the notion of a sample space Ω and a collection $\mathcal{F}$ of subsets from Ω subject to the following conventions:

(1.2a) $\Omega \in \mathcal{F}$.

(1.2b) If $A \in \mathcal{F}$, then its complement $A^c \in \mathcal{F}$.

(1.2c) If $A_1, A_2, \ldots$ is a finite or countably infinite sequence of subsets from $\mathcal{F}$, then $\bigcup_i A_i \in \mathcal{F}$.

Any collection $\mathcal{F}$ satisfying these postulates is termed a σ-field or σ-algebra. Two immediate consequences of the definitions are that the empty set $\emptyset \in \mathcal{F}$ and that if $A_1, A_2, \ldots$ is a finite or countably infinite sequence of subsets from $\mathcal{F}$, then $\bigcap_i A_i = (\bigcup_i A_i^c)^c \in \mathcal{F}$. In probability theory, we usually substitute everyday language for set theory language. Table 1.1 provides a short dictionary for translating equivalent terms.

TABLE 1.1. A Brief Dictionary of Set Theory and Probability Terms

Set Theory	**Probability**	**Set Theory**	**Probability**
set	event	null set	impossible event
union	or	universal set	certain event
intersection	and	pairwise disjoint	mutually exclusive
complement	not	inclusion	implication

The axiomatic setting of probability theory is completed by introducing a probability measure or distribution Pr on the events in $\mathcal{F}$. This function should satisfy the properties:

(1.2d) $\Pr(\Omega) = 1$.

(1.2e) $\Pr(A) \geq 0$ for any $A \in \mathcal{F}$.

(1.2f) $\Pr(\bigcup_i A_i) = \sum_i \Pr(A_i)$ for any countably infinite sequence of mutually exclusive events $A_1, A_2, \ldots$ from $\mathcal{F}$.

A triple $(\Omega, \mathcal{F}, \Pr)$ constitutes a probability space. An event $A \in \mathcal{F}$ is said to be null when $\Pr(A) = 0$ and almost sure when $\Pr(A) = 1$.

Example 1.2.1 *Discrete Uniform Distribution*

One particularly simple sample space is the set $\Omega = \{1, \ldots, n\}$. Here the natural choice of $\mathcal{F}$ is the collection of all subsets of Ω. The uniform distribution (or normalized counting measure) attributes probability $\Pr(A) = \frac{|A|}{n}$ to a set A, where $|A|$ denotes the number of elements of A. Most of the counting arguments of combinatorics presuppose the discrete uniform distribution. ■

Example 1.2.2 *Continuous Uniform Distribution*

A continuous analog of the discrete uniform distribution is furnished by Lebesgue measure on the unit interval $[0, 1]$. In this case, the best one can do is define $\mathcal{F}$ as the smallest σ-algebra containing all closed subintervals $[a, b]$ of $\Omega = [0, 1]$. The events in $\mathcal{F}$ are then said to be Borel sets. Henri Lebesgue was able to show how to extend the primitive identification $\Pr([a, b]) = b - a$ of the probability of an interval with its length to all Borel sets [171]. Invoking the axiom of choice from set theory, one can prove that it is impossible to attach a probability consistently to all subsets of $[0, 1]$. The existence of nonmeasurable sets makes the whole enterprise of measure-theoretic probability more delicate than mathematicians anticipated. Fortunately, one can ignore such subtleties in most practical problems. ■

The next example is designed to give readers a hint of the complexities involved in defining probability spaces.

Example 1.2.3 *Density in Number Theory*

Consider the natural numbers $\Omega = \{1, 2, \ldots\}$ equipped with the density function

$$\text{den}(A) \quad = \quad \lim_{n \to \infty} \frac{|A \cap \{1, 2, \ldots, n\}|}{n}.$$

Clearly, $0 \leq \text{den}(A) \leq 1$ whenever $\text{den}(A)$ is defined. Some typical densities include $\text{den}(\Omega) = 1$, $\text{den}(\{j\}) = 0$, and $\text{den}(\{j, 2j, 3j, 4j, \ldots\}) = 1/j$. Any σ-algebra $\mathcal{F}$ containing each of the positive integers $\{j\}$ fails the test of countable additivity stated in postulate (1.2f) above. Indeed,

$$\text{den}(\Omega) \quad \neq \quad 0 \quad = \quad \sum_{j=1}^{\infty} \text{den}(\{j\}).$$

Note that $\text{den}(A)$ does satisfy the test of finite additivity. Of course, it is possible to define many legitimate probability distributions on the positive integers. ■

In practice, most questions in probability theory revolve around random variables rather than sample spaces. Readers will doubtless recall that a random variable X is a function from a sample space Ω into the real line R. This is almost correct. To construct a consistent theory of integration, one must insist that a random variable be measurable. This technical condition requires that for every constant c, the set $\{\omega \in \Omega : X(\omega) \leq c\}$ be an event in the σ-algebra $\mathcal{F}$ attached to Ω. Measurability can also be defined in terms of the Borel sets $\mathcal{B}$ of R, which comprise the smallest σ-algebra containing all intervals $[a, b]$ of R. With this definition in mind, X is measurable if and only if the inverse image $X^{-1}(B)$ of every Borel set B is an event in $\mathcal{F}$. This

is analogous to but weaker than defining continuity by requiring that the inverse image of every open set be open. Almost every conceivable function $X : \Omega \mapsto \mathsf{R}$ qualifies as measurable. Formal verification of measurability usually invokes one or more of the many closure properties of measurable functions. For instance, measurability is preserved under the formation of finite sums, products, maxima, minima, and limits of measurable functions. For this reason, we seldom waste time checking measurability.

Measurable functions are candidates for integration. The simplest measurable function is the indicator 1_A of an event A. The integral or expectation $\mathrm{E}(1_A)$ of 1_A is just the corresponding probability $\Pr(A)$. Integration is first extended to simple functions $\sum_{i=1}^n c_i 1_{A_i}$ by the linearity device

$$\begin{aligned} \mathrm{E}\left(\sum_{i=1}^n c_i 1_{A_i}\right) &= \sum_{i=1}^n c_i \,\mathrm{E}(1_{A_i}) \\ &= \sum_{i=1}^n c_i \Pr(A_i) \end{aligned}$$

and from there to the larger class of integrable functions by appropriate limit arguments. Although the rigorous development of integration is one of the intellectual triumphs of modern mathematics, we record here only some of the basic facts. The two most important are linearity and nonnegativity:

(1.2g) $\mathrm{E}(aX + bY) = a\,\mathrm{E}(X) + b\,\mathrm{E}(Y)$.

(1.2h) $\mathrm{E}(X) \geq 0$ for any $X \geq 0$.

From these basic properties, a host of simple results flow. As one example, the inequality $|\,\mathrm{E}(X)| \leq \mathrm{E}(|X|)$ holds whenever $\mathrm{E}(|X|) < \infty$. As another example, taking expectations in the identity $1_{A\cup B} = 1_A + 1_B - 1_{A\cap B}$ produces the identity $\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$. Of course, one can prove this and similar equalities without introducing expectations, but the application of the expectation operator often streamlines proofs.

One of the most impressive achievements of Lebesgue's theory of integration is that it identifies sufficient conditions for the interchange of limits and integrals. Fatou's lemma states that

$$\mathrm{E}\left(\liminf_{n\to\infty} X_n\right) \leq \liminf_{n\to\infty} \mathrm{E}(X_n)$$

for any sequence $X_1, X_2, \ldots$ of nonnegative random variables. Recall that $\liminf_{n\to\infty} a_n = \lim_{n\to\infty} \inf\{a_k\}_{k\geq n}$ for any sequence a_n. In the present case, each sample point ω defines a different sequence $a_n = X_n(\omega)$.

If the sequence of random variables X_n is increasing as well as nonnegative, then the monotone convergence theorem

$$\lim_{n\to\infty} \mathrm{E}(X_n) = \mathrm{E}\left(\lim_{n\to\infty} X_n\right) \tag{1.1}$$

holds, with the possibility $\mathrm{E}(\lim_{n\to\infty} X_n) = \infty$ included. Again we need look no further than indicator functions to apply the monotone convergence theorem. Suppose $A_1 \subset A_2 \subset \cdots$ is an increasing sequence of events with limit $A_\infty = \cup_{n=1}^{\infty} A_n$. Then the continuity property

$$\lim_{n\to\infty} \Pr(A_n) \quad = \quad \Pr(A_\infty)$$

follows trivially from the monotone convergence theorem. Experts might rightfully object that this is circular reasoning because the continuity of probability is one of the ingredients that goes into constructing a rigorous theory of integration in the first place. However, this misses the psychological point that it is easier to remember and apply a general theorem than various special cases of it.

Example 1.2.4 *Borel-Cantelli Lemma*

Suppose a sequence of events $A_1, A_2, \ldots$ satisfies $\sum_{i=1}^{\infty} \Pr(A_i) < \infty$. The Borel-Cantelli lemma says only finitely many of the events occur. To prove this result, let 1_{A_i} be the indicator function of A_i, and let N be the infinite sum $\sum_{i=1}^{\infty} 1_{A_i}$. The monotone convergence theorem implies that

$$\mathrm{E}(N) \quad = \quad \sum_{i=1}^{\infty} \Pr(A_i).$$

If $\mathrm{E}(N) < \infty$ as assumed, then $N < \infty$ with probability 1. In other words, only finitely many of the A_i occur. ■

The dominated convergence theorem relaxes the assumptions that the sequence $X_1, X_2, \ldots$ is monotone and nonnegative but adds the requirement that all X_n satisfy $|X_n| \leq Y$ for some dominating random variable Y with finite expectation. Assuming that $\lim_{n\to\infty} X_n$ exists, the interchange (1.1) is again permissible. If the dominating random variable Y is constant, then most probabilists refer to the dominated convergence theorem as the bounded convergence theorem. Our next example illustrates the power of the dominated convergence theorem.

Example 1.2.5 *Differentiation Under an Expectation Sign*

Let X_t denote a family of random variables indexed by a real parameter t such that (a) $\frac{d}{dt}X_t(\omega)$ exists for all sample points ω and (b) $|\frac{d}{dt}X_t| \leq Y$ for some dominating random variable Y with finite expectation. We claim that $\frac{d}{dt}\mathrm{E}(X_t)$ exists and equals $\mathrm{E}(\frac{d}{dt}X_t)$. To prove this result, consider the difference quotient

$$\frac{\mathrm{E}(X_{t+\Delta t}) - \mathrm{E}(X_t)}{\Delta t} \quad = \quad \mathrm{E}\left(\frac{X_{t+\Delta t} - X_t}{\Delta t}\right).$$

For any sample point ω, the mean value theorem implies that

$$\begin{aligned}\left|\frac{X_{t+\Delta t}(\omega) - X_t(\omega)}{\Delta t}\right| &= \left|\frac{d}{ds}X_s(\omega)\right| \\ &\leq Y(\omega)\end{aligned}$$

for some s between t and $t + \Delta t$. Because the difference quotients converge to the derivative in a dominated fashion as Δt tends to 0, application of the dominated convergence theorem finishes the proof.

As a straightforward illustration, consider the problem of calculating the first moment of a random variable Z from its characteristic function $\mathrm{E}(e^{itZ})$. Assuming that $\mathrm{E}(|Z|)$ is finite, define the family of random variables $X_t = e^{itZ}$. It is then clear that the derivative $\frac{d}{dt}X_t(\omega) = iZ(\omega)e^{itZ(\omega)}$ exists for all sample points ω and that $Y = |iZe^{itZ}| = |Z|$ furnishes an appropriate dominating random variable. Hence, $\mathrm{E}(Z)$ equals the value of $-i\frac{d}{dt}\mathrm{E}(e^{itZ})$ at $t = 0$. ■

1.3 Conditional Probability

Constructing a rigorous theory of conditional probability and conditional expectation is as much a chore as constructing a rigorous theory of integration. Fortunately, most of the theoretic results can be motivated starting with the simple case of conditioning on an event of positive probability. In this case, we define the conditional probability

$$\Pr(B \mid A) = \frac{\Pr(B \cap A)}{\Pr(A)}$$

of any event B relative to A. Because the conditional probability $\Pr(B \mid A)$ is a legitimate probability measure, it is possible to define the conditional expectation $E(Z \mid A)$ of any integrable random variable Z. Fortunately, this boils down to nothing more than

$$\mathrm{E}(Z \mid A) = \frac{\mathrm{E}(Z1_A)}{\Pr(A)}. \tag{1.2}$$

Definition (1.2) has limited scope, and probabilists have generalized it by conditioning on a random variable rather than a single event. If X is a random variable taking only a finite number of values $x_1, \ldots, x_n$, then $\mathrm{E}(Z \mid X)$ is the random variable defined by $\mathrm{E}(Z \mid X = x_i)$ on the event $\{X = x_i\}$. Obviously, the conditional expectation operator inherits the properties of linearity and nonnegativity in Z from the ordinary expectation operator. In addition, there is the further connection

$$\begin{aligned}\mathrm{E}(Z) &= \sum_{i=1}^{n} \mathrm{E}(Z \mid X = x_i) \Pr(X = x_i) \\ &= \mathrm{E}[\mathrm{E}(Z \mid X)] \end{aligned} \tag{1.3}$$

between ordinary and conditional expectations. The final property worth highlighting,

$$\mathrm{E}[f(X)Z] \;=\; \mathrm{E}[f(X)\,\mathrm{E}(Z \mid X)], \tag{1.4}$$

is a consequence of equation (1.3) and the obvious identity

$$\mathrm{E}[f(X)Z \mid X] \;=\; f(X)\,\mathrm{E}(Z \mid X).$$

Example 1.3.1 *The Hypergeometric Distribution*

Consider a finite sequence $X_1, \ldots, X_n$ of independent Bernoulli random variables with common success probability p. Here $\Pr(X_j = 1) = p$ and $\Pr(X_j = 0) = 1 - p$, and the sum $S_n = X_1 + \cdots + X_n$ follows a binomial distribution. The hypergeometric distribution can be recovered in this setting by conditioning. For $m < n$, define the shorter sum $S_m = X_1 + \cdots + X_m$ and calculate

$$\begin{aligned}
\Pr(S_m = j \mid S_n = k) &= \frac{\binom{m}{j} p^j (1-p)^{m-j} \binom{n-m}{k-j} p^{k-j} (1-p)^{n-m+j-k}}{\binom{n}{k} p^k (1-p)^{n-k}} \\
&= \frac{\binom{m}{j}\binom{n-m}{k-j}}{\binom{n}{k}}.
\end{aligned}$$

The mean of this hypergeometric distribution is just the conditional expectation $\mathrm{E}(S_m \mid S_n = k)$. Using symmetry and the additivity of the conditional expectation operator, we find that

$$\begin{aligned}
\mathrm{E}(S_m \mid S_n = k) &= \sum_{i=1}^{m} \mathrm{E}(X_i \mid S_n = k) \\
&= m\,\mathrm{E}(X_1 \mid S_n = k) \\
&= \frac{m}{n}\,\mathrm{E}(S_n \mid S_n = k) \\
&= \frac{mk}{n}.
\end{aligned}$$

It is noteworthy that the identity $\mathrm{E}(S_m \mid S_n) = \frac{m}{n} S_n$ does not require the X_j to be Bernoulli. ■

At the highest level of abstraction, we define conditional expectation $\mathrm{E}(Z \mid \mathcal{G})$ relative to a sub-σ-algebra $\mathcal{G}$ of the underlying σ-algebra $\mathcal{F}$. Here it is important to bear in mind that Z must be integrable and that in most cases $\mathcal{G}$ is the smallest σ-algebra making a random variable X or a random vector $(X_1, \ldots, X_n)$ measurable. The technical requirement that $\mathrm{E}(Z \mid \mathcal{G})$ be measurable with respect to $\mathcal{G}$ then means that $\mathrm{E}(Z \mid \mathcal{G})$ is a function of X or $(X_1, \ldots, X_n)$. Because $\mathcal{G}$ may have an infinity of events, we can no longer rely on defining $\mathrm{E}(Z \mid \mathcal{G})$ by naively conditioning on events

of positive probability. The usual mathematical trick of turning a theorem into a definition, however, comes to the rescue. Thus, $\mathrm{E}(Z \mid \mathcal{G})$ is defined as the essentially unique integrable random variable that is measurable with respect to $\mathcal{G}$ and satisfies the analog

$$\mathrm{E}[1_C Z] \quad = \quad \mathrm{E}[1_C \, \mathrm{E}(Z \mid \mathcal{G})] \tag{1.5}$$

of equation (1.4) for every event C in $\mathcal{G}$. Hidden in this definition is an appeal to the powerful Radon-Nikodym theorem of measure theory. The upshot of these indirect arguments is that the conditional expectation operator is perfectly respectable and continues to enjoy the basic properties mentioned earlier.

In our study of martingales in Chapter 10, we will encounter increasing σ-algebras. We write $\mathcal{F} \subset \mathcal{G}$ if every event of $\mathcal{F}$ is also an event $\mathcal{G}$. In other words, $\mathcal{F}$ is less informative than $\mathcal{G}$. The "tower property"

$$\mathrm{E}[\mathrm{E}(Z \mid \mathcal{G}) \mid \mathcal{F}] \quad = \quad \mathrm{E}(Z \mid \mathcal{F}) \tag{1.6}$$

holds in this case because equation (1.5) implies

$$\begin{aligned} \mathrm{E}[1_C \, \mathrm{E}(Z \mid \mathcal{G})] &= \mathrm{E}[\mathrm{E}(1_C Z \mid \mathcal{G})] \\ &= \mathrm{E}(1_C Z) \\ &= \mathrm{E}[1_C \, \mathrm{E}(Z \mid \mathcal{F})] \end{aligned}$$

for every C in $\mathcal{F}$.

1.4 Independence

Two events A and B are independent if and only if

$$\Pr(A \cap B) \quad = \quad \Pr(A) \Pr(B).$$

This definition is equivalent to $\Pr(B \mid A) = \Pr(B)$ when $\Pr(A) > 0$. A finite or countable sequence $A_1, A_2, \ldots$ of events is independent provided

$$\Pr\left(\bigcap_{j=1}^{n} A_{i_j} \right) \quad = \quad \prod_{j=1}^{n} \Pr(A_{i_j})$$

for all finite subsequences $A_{i_1}, \ldots, A_{i_n}$. Pairwise independence is insufficient to imply independence. A sequence of random variables $X_1, X_2, \ldots$ is independent whenever the sequence of events $A_i = \{X_i \leq c_i\}$ is independent for all possible choices of the constants c_i. In practice, one usually establishes the independence of two random variables U and V by exhibiting them as measurable functions $U = f(X)$ and $V = g(Y)$ of known independent random variables X and Y.

If X and Y are independent random variables with finite expectations, then Fubini's theorem implies $\mathrm{E}(XY) = \mathrm{E}(X)\,\mathrm{E}(Y)$. If X and Y are nonnegative, then equality continues to hold even when one or both random variables have infinite expectations. From equality (1.5), one can deduce that $\mathrm{E}(Y \mid X) = \mathrm{E}(Y)$ whenever Y is independent of X. This result extends to conditioning on a σ-algebra $\mathcal{G}$ when Y is independent of the events in $\mathcal{G}$.

1.5 Distributions, Densities, and Moments

The distribution function $F(x)$ of a random variable X is defined by the formula $F(x) = \Pr(X \le x)$. Readers will recall the familiar properties:

(1.5a) $0 \le F(x) \le 1$,

(1.5b) $F(x) \le F(y)$ for $x \le y$,

(1.5c) $\lim_{x \to y+} F(x) = F(y)$,

(1.5d) $\lim_{x \to -\infty} F(x) = 0$,

(1.5e) $\lim_{x \to \infty} F(x) = 1$,

(1.5f) $\Pr(a < X \le b) = F(b) - F(a)$,

(1.5g) $\Pr(X = x) = F(x) - F(x-)$.

A random variable X is said to be discretely distributed if its possible values are limited to a sequence of points $x_1, x_2, \ldots$. In this case, its discrete density $f(x_i) = \Pr(X = x_i)$ satisfies:

(1.5h) $f(x_i) \ge 0$ for all i,

(1.5i) $\sum_i f(x_i) = 1$,

(1.5j) $F(x) = \sum_{x_i \le x} f(x_i)$.

Example 1.5.1 *The Inverse Method*

The inverse method is one of the simplest and most natural methods of simulating random variables [7]. It depends on the second of the following two properties of a distribution function $F(x)$.

(a) If $F(x)$ is continuous, then $U = F(X)$ is uniformly distributed on $[0, 1]$.

(b) If $F^{[-1]}(y) = \inf\{x : F(x) \ge y\}$ for any $0 < y < 1$, and if U is uniform on $[0, 1]$, then $F^{[-1]}(U)$ has distribution function $F(x)$.

Note that the quantile function $F^{[-1]}(u)$ is the functional inverse of $F(x)$ when $F(x)$ is continuous and strictly increasing. As a preliminary to proving properties (a) and (b), let us demonstrate that

$$\Pr[F(X) \leq F(t)] \quad = \quad F(t). \tag{1.7}$$

To prove this assertion, note that $\{X > t\} \cap \{F(X) < F(t)\} = \emptyset$ and $\{X \leq t\} \cap \{F(X) > F(t)\} = \emptyset$ together entail

$$\{F(X) \leq F(t)\} \quad = \quad \{X \leq t\} \cup \{F(X) = F(t),\ X > t\}.$$

However, the event $\{F(X) = F(t),\ X > t\}$ maps under X to an interval of constancy of $F(x)$ and therefore has probability 0. Equation (1.7) follows immediately.

For claim (a) let $u \in (0, 1)$. Because $F(x)$ is continuous, there exists t with $F(t) = u$. In view of equation (1.7),

$$\Pr[F(X) \leq u] \quad = \quad \Pr[F(X) \leq F(t)] \quad = \quad u.$$

Claim (b) follows if we can show that the events $u \leq F(t)$ and $F^{[-1]}(u) \leq t$ are identical for both u and $F(t)$ in $(0, 1)$. Assume that $F^{[-1]}(u) \leq t$. Because $F(x)$ is increasing and right continuous, the set $\{x : u \leq F(x)\}$ is an interval containing its left endpoint. Hence, $u \leq F(t)$. Conversely, if $u \leq F(t)$, then $F^{[-1]}(u) \leq t$ by definition. This completes the proof.

Because it is easy to generate uniform random numbers on a computer, the inverse method is widely used. For instance, if X is exponentially distributed with mean μ, then $F(x) = 1 - e^{-x/\mu}$ and $F^{[-1]}(u) = -\mu \ln(1-u)$. In view of the symmetry of U and $1 - U$, both of the random variables $-\mu \ln(1-U)$ and $-\mu \ln U$ are distributed as X. The major drawback of the inverse method in other examples is the difficulty of computing the quantile function $F^{[-1]}(u)$. ■

A continuously distributed random variable X has density $f(x)$ defined on the real line and satisfying:

(1.5k) $f(x) \geq 0$ for all x,

(1.5l) $\int_a^b f(x)\,dx = F(b) - F(a)$,

(1.5m) $\frac{d}{dx}F(x) = f(x)$ for almost all x.

One of the primary uses of distribution and density functions is in calculating expectations. If $h(x)$ is Borel measurable and the random variable $h(X)$ has finite expectation, then we can express its expectation as the integral

$$\mathrm{E}[h(X)] \quad = \quad \int h(x)\,dF(x).$$

One proves this result by transferring the probability measure on the underlying sample space to the real line via X. In this scheme, an interval $(a, b]$ is assigned probability $F(b) - F(a)$. In practice, we replace $dF(x)$ by counting measure or Lebesgue measure and evaluate $\mathrm{E}[h(X)]$ by $\sum_i h(x_i) f(x_i)$ in the discrete case and by $\int h(x) f(x) dx$ in the continuous case.

Counting measure and Lebesgue measure on the real line have infinite mass. Such infinite measures share many of the properties of probability measures. For instance, the dominated convergence theorem and Fubini's theorem continue to hold. We will use such properties without comment, relying on the student's training in advanced calculus to lend an air of respectability to our invocation of relevant facts.

The most commonly encountered expectations are moments. The nth moment of X is $\mu_n = \mathrm{E}(X^n)$. If we recenter X around its first moment (or mean) μ_1, then the nth central moment of X is $\mathrm{E}[(X - \mu_1)^n]$. As mentioned earlier, we can recover the mean of X from its characteristic function $\mathrm{E}(e^{itX})$ by differentiation. In general, if $\mathrm{E}(|X|^n)$ is finite, then

$$\mathrm{E}(X^n) \;=\; (-i)^n \frac{d^n}{dt^n} \mathrm{E}(e^{itX})|_{t=0}. \tag{1.8}$$

The characteristic function of a random variable always exists and uniquely determines the distribution function of the random variable [60, 117]. If X has density function $f(x)$, then $\mathrm{E}(e^{itX})$ coincides with the Fourier transform $\hat{f}(t)$ of $f(x)$ [54]. When $\hat{f}(t)$ is integrable, $f(x)$ is recoverable via the inversion formula

$$f(x) \;=\; \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \hat{f}(t)\, dt. \tag{1.9}$$

Appendix A.4 derives formula (1.9) and briefly reviews other properties of the Fourier transform.

For a random variable X possessing moments of all orders, it is usually simpler to deal with the moment generating function $\mathrm{E}(e^{tX})$. For a nonnegative random variable X, we occasionally resort to the Laplace transform $\mathrm{E}(e^{-tX})$. (If X possesses a density function $f(x)$, then $\mathrm{E}(e^{-tX})$ is also the ordinary Laplace transform of $f(x)$ as defined in science and engineering courses.) When X is integer-valued as well as nonnegative, the probability generating function $\mathrm{E}(t^X)$ for $t \in [0, 1]$ also comes in handy. Each of these transforms $M(t)$ possesses the multiplicative property

$$M_{X_1 + \cdots + X_m}(t) \;=\; M_{X_1}(t) \cdots M_{X_m}(t)$$

for a sum of independent random variables $X_1, \ldots, X_m$. The simplest transforms involve constant random variables. For instance, the Laplace transform of the constant c is just e^{-ct}; the probability generating function of the positive integer n is t^n. Chapter 2 introduces more complicated examples.

The second central moment of a random variable is also called the variance of the random variable. Readers will doubtless recall the variance formula

$$\mathrm{Var}\left(\sum_{j=1}^{m} X_j\right) = \sum_{j=1}^{m} \mathrm{Var}(X_j) + \sum_{j=1}^{m} \sum_{k \neq j} \mathrm{Cov}(X_j, X_k) \quad (1.10)$$

for a sum of m random variables. Here

$$\mathrm{Cov}(X_j, X_k) = \mathrm{E}(X_j X_k) - \mathrm{E}(X_j)\,\mathrm{E}(X_k)$$

is the covariance between X_j and X_k. Independent random variables are uncorrelated and exhibit zero covariance, but independence is hardly necessary for two random variables to be uncorrelated. For random variables with zero means, the covariance function serves as an inner product and lends a geometric flavor to many probability arguments. For example, we can think of uncorrelated, zero-mean random variables as being orthogonal. Calculation of variances and covariances is often facilitated by the conditioning formulas

$$\begin{aligned} \mathrm{Var}(X) &= \mathrm{E}[\mathrm{Var}(X \mid Z)] + \mathrm{Var}[\mathrm{E}(X \mid Z)] \\ \mathrm{Cov}(X, Y) &= \mathrm{E}[\mathrm{Cov}(X, Y \mid Z)] + \mathrm{Cov}[\mathrm{E}(X \mid Z), \mathrm{E}(Y \mid Z)] \end{aligned} \quad (1.11)$$

and by the simple formulas $\mathrm{Var}(cX) = c^2 \mathrm{Var}(X)$ and $\mathrm{Var}(X+c) = \mathrm{Var}(X)$ involving a constant c.

Example 1.5.2 *Best Predictor*

Consider a random variable X with finite variance. If Y is a second random variable defined on the same probability space, then it makes sense to inquire what function $f(Y)$ best predicts X. If we use mean square error as our criterion, then we must minimize $\mathrm{Var}[X - f(Y)]$. According to equation (1.11),

$$\begin{aligned} \mathrm{Var}[X - f(Y)] &= \mathrm{E}\{\mathrm{Var}[X - f(Y) \mid Y]\} + \mathrm{Var}\{\mathrm{E}[X - f(Y) \mid Y]\} \\ &= \mathrm{E}[\mathrm{Var}(X \mid Y)] + \mathrm{Var}[\mathrm{E}(X \mid Y) - f(Y)]. \end{aligned}$$

The term $\mathrm{E}[\mathrm{Var}(X \mid Y)]$ does not depend on the function $f(Y)$, and the term $\mathrm{Var}[\mathrm{E}(X \mid Y) - f(Y)]$ is minimized by taking $f(Y) = \mathrm{E}(X \mid Y)$. Thus, $\mathrm{E}(X \mid Y)$ is the best predictor. ■

Table 1.2 lists the densities, means, and characteristic functions of some commonly occurring univariate distributions. Restrictions on the values and parameters of these distributions are not shown. Note that the beta distribution does not possess a simple characteristic function. The version of the geometric distribution given counts the number of Bernoulli trials until a success, not the number of failures.

TABLE 1.2. Common Distributions

Name	Density	Mean	Transform
Binomial	$\binom{n}{x}p^x(1-p)^{n-x}$	np	$(1-p+pe^{it})^n$
Poisson	$\frac{\lambda^x}{x!}e^{-\lambda}$	λ	$e^{\lambda(e^{it}-1)}$
Geometric	$(1-p)^{x-1}p$	$\frac{1}{p}$	$\frac{pe^{it}}{1-(1-p)e^{it}}$
Uniform	$\frac{1}{b-a}$	$\frac{a+b}{2}$	$\frac{e^{itb}-e^{ita}}{it(b-a)}$
Normal	$\frac{1}{\sqrt{2\pi\sigma^2}}e^{-(x-\mu)^2/2\sigma^2}$	μ	$e^{it\mu-\sigma^2t^2/2}$
Exponential	$\lambda e^{-\lambda x}$	$\frac{1}{\lambda}$	$\frac{\lambda}{\lambda-it}$
Beta	$\frac{\Gamma(\alpha+\beta)x^{\alpha-1}(1-x)^{\beta-1}}{\Gamma(\alpha)\Gamma(\beta)}$	$\frac{\alpha}{\alpha+\beta}$	
Gamma	$\frac{\lambda^\alpha x^{\alpha-1}}{\Gamma(\alpha)}e^{-\lambda x}$	$\frac{\alpha}{\lambda}$	$\left(\frac{\lambda}{\lambda-it}\right)^\alpha$

In statistical applications, densities often depend on parameters. The parametric families displayed in Table 1.2 are typical. Viewed as a function of its parameters, a density $f(x)$, either discrete or continuous, is called a likelihood. For purposes of estimation, one can ignore any factor of $f(x)$ that depends only on the data x and not on the parameters. In maximum likelihood estimation, $f(x)$ is maximized with respect to its parameters. The parameters giving the maximum likelihood are the maximum likelihood estimates. These distinctions carry over to multidimensional densities.

1.6 Convolution

If X and Y are independent random variables with distribution functions $F(x)$ and $G(y)$, then $F * G(z)$ denotes the distribution function of the sum $Z = X + Y$. Fubini's theorem permits us to write this convolution of distribution functions as

$$\begin{aligned} F * G(z) &= \int\int 1_{\{x+y\le z\}}\, dF(x)\, dG(y) \\ &= \int F(z-y)\, dG(y). \end{aligned}$$

If the random variable X possesses density $f(x)$, then executing the change of variables $w = x + y$ and interchanging the order of integration yield

$$F * G(z) = \int\int_{-\infty}^{z-y} f(x)\, dx\, dG(y)$$

$$\begin{aligned} &= \int \int_{-\infty}^{z} f(w-y)\,dw\,dG(y) \\ &= \int_{-\infty}^{z} \int f(w-y)\,dG(y)\,dw. \end{aligned}$$

Thus, $Z = X + Y$ has density $\int f(z-y)\,dG(y)$. When Y has density $g(y)$, this simplifies to the convolution $f * g(z) = \int f(z-y)g(y)\,dy$ of the two density functions.

Other functions of X and Y produce similar results. For example, if we suppose that $Y > 0$, then the product $U = XY$ and ratio $V = X/Y$ have distribution functions $\int_0^\infty F(uy^{-1})\,dG(y)$ and $\int_0^\infty F(vy)\,dG(y)$, respectively. Differentiation of these by u and v leads to the corresponding densities $\int_0^\infty f(uy^{-1})y^{-1}\,dG(y)$ and $\int_0^\infty f(vy)y\,dG(y)$. Problem 17 asks the reader to verify these claims rigorously. Example 1.7.2 treats a ratio when the denominator possesses a density.

1.7 Random Vectors

Random vectors with dependent components arise in many problems. Tools for manipulating random vectors are therefore crucially important. For instance, we define the expectation of a random vector $X = (X_1, \ldots, X_n)^t$ componentwise by $\mathrm{E}(X) = [\mathrm{E}(X_1), \ldots, \mathrm{E}(X_n)]^t$. Linearity carries over from the scalar case in the sense that

$$\begin{aligned} \mathrm{E}(X+Y) &= \mathrm{E}(X) + \mathrm{E}(Y) \\ \mathrm{E}(AX) &= A\,\mathrm{E}(X) \end{aligned}$$

for a compatible random vector Y and a compatible constant matrix A. Similar definitions and results come into play for random matrices when we calculate the covariance matrix

$$\begin{aligned} \mathrm{Cov}(X,Y) &= \mathrm{E}\{[X - \mathrm{E}(X)][Y - \mathrm{E}(Y)]^t\} \\ &= \mathrm{E}(XY^t) - \mathrm{E}(X)\,\mathrm{E}(Y)^t \end{aligned}$$

of two random vectors X and Y. The covariance operator is linear in each of its arguments and vanishes when these arguments are independent. Furthermore, one can readily check that

$$\mathrm{Cov}(AX, BY) = A\,\mathrm{Cov}(X,Y)B^t$$

for compatible constant matrices A and B. The variance matrix of X is expressible as $\mathrm{Var}(X) = \mathrm{Cov}(X,X)$ and is nonnegative definite.

We define the distribution function $F(x)$ of X via

$$F(x) = \Pr\left(\cap_{i=1}^n \{X_i \le x_i\}\right).$$

This function is increasing in each component x_i of x holding the other components fixed. The marginal distribution function of a subvector of X is recoverable by taking the limit of $F(x)$ as the corresponding components of x tend to ∞. The components of X are independent if and only if $F(x)$ factors as the product $\prod_{i=1}^n F_i(x_i)$ of the marginal distribution functions. In many practical problems, X possesses a density $f(x)$. We then have

$$\Pr(X \in C) = \int_C f(x)\,dx$$

for every Borel set C. Here the indicated integral is multidimensional. The distribution and density functions are related by

$$F(x) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f(y)\,dy_1 \cdots dy_n.$$

The marginal density of a subvector of X is recoverable by integrating $f(x)$ over the components of x corresponding to the complementary subvector. Furthermore, the components of X are independent if and only if $f(x)$ factors as the product $\prod_{i=1}^n f_i(x_i)$ of the marginal densities. For discrete random vectors, similar results hold provided we interpret $f(x)$ as a discrete density and replace multiple integrals by multiple sums.

Conditional expectations are often conveniently calculated using conditional densities. Consider a bivariate random vector $X = (X_1, X_2)$ with density $f(x_1, x_2)$. The formula

$$f_{2|1}(x_2 \mid x_1) = \frac{f(x_1, x_2)}{f_1(x_1)}$$

determines the conditional density of X_2 given X_1. To compute the conditional expectation of a function $h(X)$ of X, we form

$$\mathrm{E}[\,h(X) \mid X_1 = x_1] = \int h(x_1, x_2) f_{2|1}(x_2 \mid x_1)\,dx_2.$$

This works because

$$\begin{aligned} \mathrm{E}[1_C(X_1)h(X)] &= \int_C \int h(x_1, x_2) f_{2|1}(x_2 \mid x_1)\,dx_2 f_1(x_1)\,dx_1 \\ &= \mathrm{E}\{1_C(X_1)\,\mathrm{E}[\,h(X) \mid X_1]\} \end{aligned}$$

mirrors equation (1.5).

Example 1.7.1 *Bayes' Rule*

In many statistical applications, it is common to know one conditional density $f_{2|1}(x_2 \mid x_1)$ but not the other $f_{1|2}(x_1 \mid x_2)$. Since

$$\begin{aligned} f_2(x_2) &= \int f(x_1, x_2)\,dx_1 \\ &= \int f_1(x_1) f_{2|1}(x_2 \mid x_1)\,dx_1, \end{aligned}$$

it follows that

$$\begin{aligned} f_{1|2}(x_1 \mid x_2) &= \frac{f(x_1, x_2)}{f_2(x_2)} \\ &= \frac{f_1(x_1) f_{2|1}(x_2 \mid x_1)}{f_2(x_2)} \\ &= \frac{f_1(x_1) f_{2|1}(x_2 \mid x_1)}{\int f_1(x_1) f_{2|1}(x_2 \mid x_1)\, dx_1}. \end{aligned}$$

Variants of this simple formula underlie all of Bayesian statistics. ■

Often probability models involve transformations of one random vector into another. Let $T(x)$ be a continuously differentiable transformation of an open set U containing the range of X onto an open set V. The density $g(y)$ of the random vector $Y = T(X)$ is determined by the standard change of variables formula

$$\begin{aligned} \Pr(Y \in C) &= \int_{T^{-1}(C)} f(x)\, dx \\ &= \int_C f \circ T^{-1}(y) |\det dT^{-1}(y)|\, dy \end{aligned} \tag{1.12}$$

from advanced calculus [97, 173]. Here we assume that $T(x)$ is invertible and that its differential (or Jacobian matrix)

$$dT(x) = \left[\frac{\partial}{\partial x_j} T_i(x)\right]$$

of partial derivatives is invertible at each point $x \in U$. Under these circumstances, the chain rule applied to $T^{-1}[T(x)] = x$ produces

$$dT^{-1}(y) dT(x) = I$$

for $y = T(x)$. This permits us to substitute $|\det T(x)|^{-1}$ for $|\det dT^{-1}(y)|$ in the change of variables formula.

Example 1.7.2 *Density of a Ratio*

Let X_1 and X_2 be independent random variables with densities $f_1(x_1)$ and $f_2(x_2)$. The transformation

$$T\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1/x_2 \\ x_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

has inverse

$$T^{-1}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} y_1 y_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

and differential and Jacobian

$$\begin{aligned} dT(x) &= \begin{pmatrix} 1/x_2 & -x_1/x_2^2 \\ 0 & 1 \end{pmatrix} \\ \det dT(x) &= 1/x_2. \end{aligned}$$

The fact that $T(x)$ is undefined on $\{x : x_2 = 0\}$ is harmless since this closed set has probability 0. The change of variables formula (1.12) implies that $Y_1 = X_1/X_2$ and $Y_2 = X_2$ have joint density

$$f_1(y_1y_2)f_2(y_2)|y_2|.$$

Integrating over y_2 gives the marginal density

$$\int_{-\infty}^{\infty} f_1(y_1y_2)f_2(y_2)|y_2|\,dy_2$$

of Y_1.

As a concrete example, suppose that X_1 and X_2 have standard normal densities. Then the ratio $Y_1 = X_1/X_2$ has density

$$\begin{aligned} \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-(y_1y_2)^2/2}e^{-y_2^2/2}|y_2|\,dy_2 &= \frac{1}{\pi}\int_0^{\infty} e^{-y_2^2(y_1^2+1)/2}y_2\,dy_2 \\ &= -\frac{1}{\pi(y_1^2+1)}e^{-y_2^2(y_1^2+1)/2}\Big|_0^{\infty} \\ &= \frac{1}{\pi(y_1^2+1)}. \end{aligned}$$

This is the density of a Cauchy random variable. Because X_1/X_2 and X_2/X_1 are identically distributed, the reciprocal of a Cauchy is Cauchy. ■

To recover the moments of a random vector X, we can differentiate its characteristic function $s \mapsto \mathrm{E}(e^{is^tX})$. In particular,

$$\begin{aligned} \mathrm{E}(X_j) &= -i\frac{\partial}{\partial s_j}\mathrm{E}(e^{is^tX})|_{s=\mathbf{0}} \\ \mathrm{E}(X_j^2) &= -\frac{\partial^2}{\partial s_j^2}\mathrm{E}(e^{is^tX})|_{s=\mathbf{0}} \\ \mathrm{E}(X_jX_k) &= -\frac{\partial^2}{\partial s_j\partial s_k}\mathrm{E}(e^{is^tX})|_{s=\mathbf{0}}. \end{aligned}$$

The characteristic function of X uniquely determines its distribution.

1.8 Multivariate Normal Random Vectors

As an illustration of the material reviewed, we now consider the multivariate normal distribution. Among the many possible definitions, we adopt

the one most widely used in stochastic simulation. Our point of departure will be random vectors with independent standard normal components. If such a random vector X has n components, then its density is

$$\prod_{j=1}^{n} \frac{1}{\sqrt{2\pi}} e^{-x_j^2/2} = \left(\frac{1}{2\pi}\right)^{n/2} e^{-x^t x/2}.$$

As demonstrated in Chapter 2, the standard normal distribution has mean 0, variance 1, and characteristic function $e^{-s^2/2}$. It follows that X has mean vector $\mathbf{0}$, variance matrix I, and characteristic function

$$\mathrm{E}(e^{is^t X}) = \prod_{j=1}^{n} e^{-s_j^2/2} = e^{-s^t s/2}.$$

We now define any affine transformation $Y = AX + \mu$ of X to be multivariate normal [164]. This definition has several practical consequences. First, it is clear that $\mathrm{E}(Y) = \mu$ and $\mathrm{Var}(Y) = A\,\mathrm{Var}(X)A^t = AA^t = \Omega$. Second, any affine transformation $BY+\nu = BAX+B\mu+\nu$ of Y is also multivariate normal. Third, any subvector of Y is multivariate normal. Fourth, the characteristic function of Y is

$$\mathrm{E}(e^{is^t Y}) = e^{is^t \mu}\,\mathrm{E}(e^{is^t AX}) = e^{is^t \mu - s^t AA^t s/2} = e^{is^t \mu - s^t \Omega s/2}.$$

This enumeration omits two more subtle issues. One is whether Y possesses a density. Observe that Y lives in an affine subspace of dimension equal to or less than the rank of A. Thus, if Y has m components, then $n \geq m$ must hold in order for Y to possess a density. A second issue is the existence and nature of the conditional density of a set of components of Y given the remaining components. We can clarify both of these issues by making canonical choices of X and A based on the classical QR decomposition of a matrix, which follows directly from the Gram-Schmidt orthogonalization procedure. See Problem 22 or reference [39].

Assuming that $n \geq m$, we can write

$$A^t = Q \begin{pmatrix} R \\ \mathbf{0} \end{pmatrix}, \tag{1.13}$$

where Q is an $n \times n$ orthogonal matrix and $R = L^t$ is an $m \times m$ upper-triangular matrix with nonnegative diagonal entries. (If $n = m$, we omit the zero matrix in the QR decomposition.) It follows that

$$AX = (\,L \quad \mathbf{0}^t\,)\,Q^t X = (\,L \quad \mathbf{0}^t\,)\,Z.$$

In view of the change of variables formula (1.12) and the facts that the orthogonal matrix Q^t preserves inner products and has determinant ± 1, the random vector Z has n independent standard normal components and

serves as a substitute for X. Not only is this true, but we can dispense with the last $n - m$ components of Z because they are multiplied by the matrix $\mathbf{0}^t$. Thus, we can safely assume $n = m$ and calculate the density of $Y = LZ + \mu$ when L is invertible. In this situation, $\Omega = LL^t$ is termed the Cholesky decomposition, and the change of variables formula (1.12) shows that Y has density

$$\begin{aligned} f(y) &= \left(\frac{1}{2\pi}\right)^{n/2} |\det L^{-1}| e^{-(y-\mu)^t (L^{-1})^t L^{-1}(y-\mu)/2} \\ &= \left(\frac{1}{2\pi}\right)^{n/2} |\det \Omega|^{-1/2} e^{-(y-\mu)^t \Omega^{-1}(y-\mu)/2}, \end{aligned}$$

where $\Omega = LL^t$ is the variance matrix of Y. In this formula for the density, the absolute value signs on $\det \Omega$ and $\det L^{-1}$ are redundant because these determinants are positive.

To address the issue of conditional densities, consider the compatibly partitioned vectors $Y^t = (Y_1^t, Y_2^t)$, $X^t = (X_1^t, X_2^t)$, and $\mu^t = (\mu_1^t, \mu_2^t)$ and matrices

$$L = \begin{pmatrix} L_{11} & \mathbf{0} \\ L_{21} & L_{22} \end{pmatrix} \qquad \Omega = \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix}.$$

Now suppose that X is standard normal, that $Y = LX + \mu$, and that L_{11} has full rank. For $Y_1 = y_1$ fixed, the equation $y_1 = L_{11}X_1 + \mu_1$ shows that X_1 is fixed at the value $x_1 = L_{11}^{-1}(y_1 - \mu_1)$. Because no restrictions apply to X_2, we have

$$Y_2 = L_{22}X_2 + L_{21}L_{11}^{-1}(y_1 - \mu_1) + \mu_2.$$

Thus, Y_2 given Y_1 is normal with mean $L_{21}L_{11}^{-1}(y_1 - \mu_1) + \mu_2$ and variance $L_{22}L_{22}^t$. To express these in terms of the blocks of $\Omega = LL^t$, observe that

$$\begin{aligned} \Omega_{11} &= L_{11}L_{11}^t \\ \Omega_{21} &= L_{21}L_{11}^t \\ \Omega_{22} &= L_{21}L_{21}^t + L_{22}L_{22}^t. \end{aligned}$$

The first two of these equations imply that $L_{21}L_{11}^{-1} = \Omega_{21}\Omega_{11}^{-1}$. The last equation then gives

$$\begin{aligned} L_{22}L_{22}^t &= \Omega_{22} - L_{21}L_{21}^t \\ &= \Omega_{22} - \Omega_{21}(L_{11}^t)^{-1}L_{11}^{-1}\Omega_{12} \\ &= \Omega_{22} - \Omega_{21}\Omega_{11}^{-1}\Omega_{12}. \end{aligned}$$

These calculations do not require that Y_2 possess a density. In summary, the conditional distribution of Y_2 given Y_1 is normal with mean and variance

$$\begin{aligned} \mathrm{E}(Y_2 \mid Y_1) &= \Omega_{21}\Omega_{11}^{-1}(Y_1 - \mu_1) + \mu_2 \\ \mathrm{Var}(Y_2 \mid Y_1) &= \Omega_{22} - \Omega_{21}\Omega_{11}^{-1}\Omega_{12}. \end{aligned}$$

1.9 Problems

1. Let Ω be an infinite set. A subset $S \subset \Omega$ is said to be cofinite when its complement S^c is finite. Demonstrate that the family of subsets

$$\mathcal{F} \quad = \quad \{S \subset \Omega : S \text{ is finite or cofinite}\}$$

is not a σ-algebra. What property fails? If we define $\mathrm{P}(S) = 0$ for S finite and $\mathrm{P}(S) = 1$ for S cofinite, then prove that $\mathrm{P}(S)$ is finitely additive but not countably additive.

2. The symmetric difference $A \triangle B$ of two events A and B is defined as $(A \cap B^c) \cup (A^c \cap B)$. Show that $A \triangle B$ has indicator $|1_A - 1_B|$. Use this fact to prove the triangle inequality

$$\Pr(A \triangle C) \quad \leq \quad \Pr(A \triangle B) + \Pr(B \triangle C).$$

It follows that if we ignore events of probability 0, then the collection of events forms a metric space.

3. Suppose A, B, and C are three events with $\Pr(A \cap B) > 0$. Show that A and C are conditionally independent given B if and only if the Markov property $\Pr(C \mid A \cap B) = \Pr(C \mid B)$ holds.

4. Suppose X_n is a sequence of nonnegative random variables that converges pointwise to the random variable X. If X is integrable, then Scheffe's lemma declares that $\lim_{n\to\infty} \mathrm{E}(|X_n - X|) = 0$ if and only if $\lim_{n\to\infty} \mathrm{E}(X_n) = \mathrm{E}(X)$. Prove this equivalence. (Hints: Let A_n and B_n be the events $X_n - X > 0$ and $X_n - X \leq 0$. Write

$$\begin{aligned} \mathrm{E}(X_n - X) &= \mathrm{E}[1_{A_n}(X_n - X)] + \mathrm{E}[1_{B_n}(X_n - X)] \\ \mathrm{E}(|X_n - X|) &= \mathrm{E}[1_{A_n}(X_n - X)] - \mathrm{E}[1_{B_n}(X_n - X)] \end{aligned}$$

and apply the dominated convergence theorem to the rightmost expectations.)

5. Consider a sequence of independent events $A_1, A_2, \ldots$ satisfying

$$\sum_{i=1}^{\infty} \Pr(A_i) \quad = \quad \infty.$$

As a partial converse to the Borel-Cantelli lemma, prove that infinitely many of the A_i occur. (Hints: Express the event that infinitely many of the events occur as $\cap_{n=1}^{\infty} \cup_{i=n}^{\infty} A_i$. Use the inequality $1 - x \leq e^{-x}$ to bound an infinite product by the exponential of an infinite sum.)

6. Use Problem 5 to prove that the pattern SFS of a success, failure, and success occurs infinitely many times in a sequence of Bernoulli trials. This result obviously generalizes to more complex patterns.

7. Consider a sequence $X_1, X_2, \ldots$ of independent random variables that are exponentially distributed with mean 1. Show that

$$\begin{aligned} 1 &= \limsup_{n\to\infty} \frac{X_n}{\ln n} \\ 1 &= \limsup_{n\to\infty} \frac{X_n - \ln n}{\ln \ln n} \\ 1 &= \limsup_{n\to\infty} \frac{X_n - \ln n - \ln \ln n}{\ln \ln \ln n}. \end{aligned}$$

(Hints: Use the sums

$$\begin{aligned} \infty &= \sum_{n=1}^{\infty} \frac{1}{n} \\ \infty &= \sum_{n=1}^{\infty} \frac{1}{n \ln n} \\ \infty &= \sum_{n=1}^{\infty} \frac{1}{n(\ln n)(\ln \ln n)} \end{aligned}$$

from pages 54 and 55 of [173], and apply Problem 5.)

8. Discuss how you would use the inverse method of Example 1.5.1 to generate a random variable with (a) the continuous logistic density

$$f(x|\mu,\sigma) = \frac{e^{-\frac{x-\mu}{\sigma}}}{\sigma[1+e^{-\frac{x-\mu}{\sigma}}]^2},$$

(b) the Pareto density

$$f(x|\alpha,\beta) = \frac{\beta\alpha^\beta}{x^{\beta+1}} 1_{(\alpha,\infty)}(x),$$

and (c) the Weibull density

$$f(x|\delta,\gamma) = \frac{\gamma}{\delta} x^{\gamma-1} e^{-\frac{x^\gamma}{\delta}} 1_{(0,\infty)}(x),$$

where α, β, γ, δ, and σ are taken positive.

9. Let the random variable X have distribution function $F(x)$. Demonstrate that

$$\mathrm{E}\{h[F(X)]\} = \int_0^1 h(u)\, du$$

for any integrable function $h(u)$ on $[0,1]$.

10. Let the random variable X have symmetric density $f(x) = f(-x)$. Prove that the corresponding distribution function $F(x)$ satisfies the identity $\int_{-a}^{a} F(x)\,dx = a$ for all $a \geq 0$ [183].

11. Suppose X has a continuous, strictly increasing distribution function $F(x)$ and $Y = -X$ has distribution function $G(y)$. Show that X is symmetrically distributed around some point μ if and only if the function $x \mapsto x - G^{-1}[F(x)]$ is constant, where $G^{-1}[G(y)] = y$ for all y.

12. Prove the two conditioning formulas in equation (1.11) for calculating variances and covariances.

13. If X and Y are independent random variables with finite variances, then show that

$$\mathrm{Var}(XY) \quad = \quad \mathrm{Var}(X)\,\mathrm{Var}(Y) + \mathrm{E}(X)^2\,\mathrm{Var}(Y) + \mathrm{E}(Y)^2\,\mathrm{Var}(X).$$

14. Suppose X and Y are independent random variables with finite variances. Define Z to be either X or Y depending on the outcome of a coin toss. In other words, set $Z = X$ with probability p and $Z = Y$ with probability $q = 1-p$. Find the mean, variance, and characteristic function of Z.

15. Let $S_n = X_1 + \cdots + X_n$ be the sum of n independent random variables, each distributed uniformly over the set $\{1, 2, \ldots, m\}$. For example, imagine tossing an m-sided die n times and recording the total score. Calculate $\mathrm{E}(S_n)$ and $\mathrm{Var}(S_n)$.

16. Suppose Y has exponential density e^{-y} with unit mean. Given Y, let a point X be chosen uniformly from the interval $[0, Y]$. Show that X has density $E_1(x) = \int_x^\infty e^{-y} y^{-1}\,dy$ and distribution function $1 - e^{-x} + xE_1(x)$. Calculate $\mathrm{E}(X)$ and $\mathrm{Var}(X)$.

17. Validate the formulas for the distribution and density functions of the product XY and the ratio X/Y of independent random variables X and $Y > 0$ given in Section 1.6. (Hint: Mimic the arguments used in establishing the convolution formulas.)

18. Suppose X and Y are independent random variables concentrated on the interval $(0, \infty)$. If $\mathrm{E}(X) < \infty$ and Y has density $g(y)$, then show that the ratio X/Y has finite expectation if and only if

$$\int_0^1 y^{-1} g(y) dy \quad < \quad \infty.$$

19. Let X_1 and X_2 be independent random variables with common exponential density $\lambda e^{-\lambda x}$ on $(0, \infty)$. Show that the random variables $Y_1 = X_1 + X_2$ and $Y_2 = X_1/X_2$ are independent, and find their densities.

20. Let $X_1, \ldots, X_n$ be a sequence of independent standard normal random variables. Prove that $\chi_n^2 = X_1^2 + \cdots + X_n^2$ has a gamma distribution. Calculate the mean and variance of χ_n^2.

21. Continuing Problem 20, let X be a multivariate normal random vector with mean vector μ and invertible variance matrix Ω. If X has n components, then show that the quadratic form $(X-\mu)^t\Omega^{-1}(X-\mu)$ has a χ_n^2 distribution.

22. For $n \geq m$, verify the QR decomposition (1.13). (Hints: Write

$$\begin{aligned} A^t &= (a_1, \ldots, a_m) \\ Q &= (q_1, \ldots, q_n) \\ R &= (r_1, \ldots, r_m). \end{aligned}$$

The Gram-Schmidt orthogonalization process applied to the columns of A^t yields orthonormal column vectors $q_1, \ldots, q_m$ satisfying

$$a_i = \sum_{j=1}^{i} q_j r_{ji}.$$

Complete this orthonormal basis by adding vectors $q_{m+1}, \ldots, q_n$.)

23. The Hadamard product $C = A \circ B$ of two matrices $A = (a_{ij})$ and $B = (b_{ij})$ of the same dimensions has entries $c_{ij} = a_{ij}b_{ij}$. If A and B are nonnegative definite matrices, then show that $A \circ B$ is nonnegative definite. If in addition A is positive definite, and B has positive diagonal entries, then show that $A \circ B$ is positive definite. (Hints: Let X and Y be multivariate normal random vectors with mean $\mathbf{0}$ and variance matrices A and B. Show that the vector Z with entries $Z_i = X_i Y_i$ has variance matrix $A \circ B$. To prove that $A \circ B$ is positive definite, demonstrate that $v^t Z$ has positive variance for $v \neq \mathbf{0}$. This can be done via the equality $\mathrm{Var}(v^t Z) = \mathrm{E}[(v \circ Y)^t A (v \circ Y)]$ based on formula (1.11).)

2
Calculation of Expectations

2.1 Introduction

Many of the hardest problems in applied probability revolve around the calculation of expectations of one sort or another. On one level, these are merely humble exercises in integration or summation. However, we should not be so quick to dismiss the intellectual challenges. Readers are doubtless already aware of the clever applications of characteristic and moment generating functions. This chapter is intended to review and extend some of the tools that probabilists routinely call on. Readers can consult the books [34, 59, 60, 78, 80, 166] for many additional examples of these tools in action.

2.2 Indicator Random Variables and Symmetry

Many counting random variables can be expressed as the sum of indicator random variables. If $S = \sum_{i=1}^{n} 1_{A_i}$ for events $A_1, \ldots, A_n$, then straightforward calculations and equation (1.10) give

$$\begin{aligned} \mathrm{E}(S) &= \sum_{i=1}^{n} \Pr(A_i) && (2.1) \\ \mathrm{Var}(S) &= \sum_{i=1}^{n} \Pr(A_i) + \sum_{i=1}^{n} \sum_{j \neq i} \Pr(A_i \cap A_j) - \mathrm{E}(S)^2. && (2.2) \end{aligned}$$

K. Lange, *Applied Probability,*
DOI 10.1007/978-1-4419-7165-4_2, © Springer Science+Business Media, LLC 2010

Example 2.2.1 *Fixed Points of a Random Permutation*

There are $n!$ permutations π of the set $\{1, \ldots, n\}$. Under the uniform distribution, each of these permutations is equally likely. If A_i is the event that $\pi(i) = i$, then $S = \sum_{i=1}^{n} 1_{A_i}$ is the number of fixed points of π. By symmetry, $\Pr(A_i) = \frac{1}{n}$ and

$$\begin{aligned} \Pr(A_i \cap A_j) &= \Pr(A_j \mid A_i) \Pr(A_i) \\ &= \frac{1}{(n-1)n}. \end{aligned}$$

Hence, the formulas in (2.2) yield $\mathrm{E}(S) = \frac{n}{n} = 1$ and

$$\begin{aligned} \mathrm{Var}(S) &= \frac{n}{n} + \sum_{i=1}^{n} \sum_{j \neq i} \frac{1}{(n-1)n} - 1^2 \\ &= 1. \end{aligned}$$

The equality $\mathrm{E}(S) = \mathrm{Var}(S)$ suggests that S is approximately Poisson distributed. We will verify this conjecture in Example 4.3.1. ■

Example 2.2.2 *Pattern Matching*

Consider a random string of n letters drawn uniformly and independently from the alphabet $\{1, \ldots, m\}$. Let S equal the number of occurrences of a given word of length $k \leq n$ in the string. For example, with $m = 2$ and $n = 10$, all strings have probability 2^{-10}. The word 101 is present in the string 1101011101 three times. Represent S as

$$S = \sum_{j=1}^{n-k+1} 1_{A_j},$$

where A_j is the event that the given word occurs beginning at position j in the string. In view of equation (2.1), it is obvious that

$$\mathrm{E}(S) = \sum_{j=1}^{n-k+1} \Pr(A_j) = (n-k+1)p^k$$

for the choice $p = m^{-1}$. Calculation of $\mathrm{Var}(S)$ is more subtle. Equation (2.2) and symmetry imply

$$\mathrm{Var}(S) = (n-k+1)\mathrm{Var}(1_{A_1}) + 2\sum_{j=2}^{l}(n-k-j+2)\mathrm{Cov}(1_{A_1}, 1_{A_j}),$$

where $l = \min\{k, n-k+1\}$, the multiplier $2(n-k-j+2)$ of $\mathrm{Cov}(1_{A_1}, 1_{A_j})$ equals the number of pairs (r, x) with $|r - s| = j - 1$, and the events A_r and A_s are independent whenever $|r - s| \geq k$. Although it is clear that

$$\mathrm{Var}(1_{A_1}) = p^k - p^{2k},$$

the covariance terms present more of a challenge because of the possibility of overlapping occurrences of the word. Let ϵ_l equal 1 or 0, depending on whether the last l letters of the word taken as a block coincide with the first l letters of the word taken as a block. For the particular word 101, $\epsilon = 1$ and $\epsilon_2 = 0$. With this convention, we calculate

$$\text{Cov}(1_{A_1}, 1_{A_j}) \quad = \quad \epsilon_{k-j+1} p^{k+j-1} - p^{2k}$$

for $2 \leq j \leq k$. ■

Both of the previous examples exploit symmetry as well as indicator random variables. Here is another example from sampling theory that depends crucially on symmetry [40].

Example 2.2.3 *Sampling without Replacement*

Assume that m numbers $Y_1, \ldots, Y_m$ are drawn randomly without replacement from n numbers $x_1, \ldots, x_n$. It is of interest to calculate the mean and variance of the sample average $S = \frac{1}{m} \sum_{i=1}^{m} Y_i$. Clearly,

$$\text{E}(S) \quad = \quad \frac{1}{m} \sum_{i=1}^{m} \text{E}(Y_i) \quad = \quad \bar{x},$$

where $\bar{x}$ is the sample average of the x_i. To calculate the variance of S, let $s^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2$ denote the sample variance of the x_i. Now imagine filling out the sample to $Y_1, \ldots, Y_n$ so that all n values $x_1, \ldots, x_n$ are exhausted. Because the sum $Y_1 + \cdots + Y_n = n\bar{x}$ is constant, symmetry and equation (1.10) imply that

$$\begin{aligned} 0 &= \text{Var}(Y_1 + \cdots + Y_n) \\ &= ns^2 + n(n-1) \, \text{Cov}(Y_1, Y_2). \end{aligned}$$

In verifying that $\text{Cov}(Y_i, Y_j) = \text{Cov}(Y_1, Y_2)$, it is helpful to think of the sampling being done simultaneously rather than sequentially. In any case, $\text{Cov}(Y_1, Y_2) = -\frac{s^2}{n-1}$, and the formula

$$\begin{aligned} \text{Var}(S) &= \frac{1}{m^2} \Big[ms^2 + m(m-1) \, \text{Cov}(Y_1, Y_2) \Big] \\ &= \frac{1}{m^2} \left[ms^2 - \frac{m(m-1)s^2}{n-1} \right] \\ &= \frac{(n-m)s^2}{m(n-1)} \end{aligned}$$

follows directly. ■

The next problem, the first of a long line of problems in geometric probability, also yields to symmetry arguments [116].

Example 2.2.4 *Buffon Needle Problem*

Suppose we draw an infinite number of equally distant parallel lines on the plane R^2. If we drop a needle (or line segment) of fixed length randomly onto the plane, then the needle may or may not intersect one of the parallel lines. Figure 2.1 shows the needle intersecting a line. Buffon's problem is to calculate the probability of an intersection. Without loss of generality, we assume that the spacing between lines is 1 and the length of the needle is d. Let X_d be the random number of lines that the needle intersects. If $d < 1$, then X_d equals 0 or 1, and $\Pr(X_d = 1) = \mathrm{E}(X_d)$. Thus, Buffon's problem reduces to calculating an expectation for a short needle. Our task is to construct the function $f(d) = \mathrm{E}(X_d)$. This function is clearly nonnegative, increasing, and continuous in d.

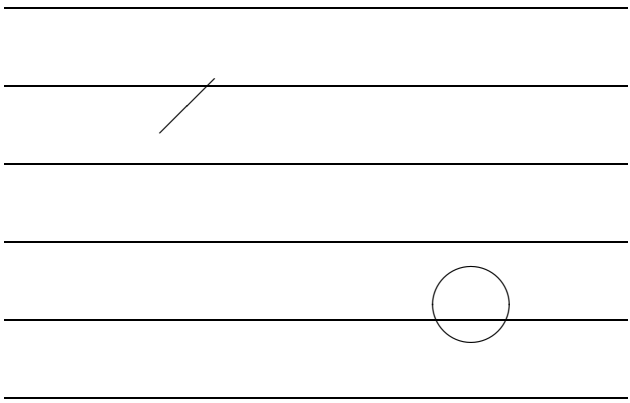


FIGURE 2.1. Diagram of the Buffon Needle Problem

Now imagine randomly dropping two needles simultaneously of lengths d_1 and d_2. The expected number of intersections of both needles obviously amounts to $\mathrm{E}(X_{d_1}) + \mathrm{E}(X_{d_2})$. This result holds whether we drop the two needles independently or dependently, as long as we drop them randomly. We can achieve total dependence by welding the end of one needle to the start of the other needle. If the weld is just right, then the two needles will form a single needle of length $d_1 + d_2$. This shows that

$$f(d_1 + d_2) \quad = \quad f(d_1) + f(d_2). \tag{2.3}$$

The only functions $f(d)$ that are nonnegative, increasing, and additive in d are the linear functions $f(d) = cd$ with $c \geq 0$. To find the proportionality constant c, we take the experiment of welding needles together to its logical extreme. Thus, a rigid wire of welded needles with perimeter p determines

on average cp intersections. In the limit, we can replace the wire by any reasonable curve. The key to finding c is to take a circle of diameter 1. This particular curve has perimeter π and either is tangent to two lines or intersects the same line twice. Figure 2.1 depicts the latter case. The equation $2 = c\pi$ now determines $c = 2/\pi$ and $f(d) = 2d/\pi$. ■

2.3 Conditioning

A third way to calculate expectations is to condition. Two of the next three examples use conditioning to derive a recurrence relation. In the family planning model, the recurrence is difficult to solve exactly, but as with most recurrences, it is easy to implement by hand or computer.

Example 2.3.1 *Beta-Binomial Distribution*

Consider a random variable P with beta density

$$f_{\alpha\beta}(p) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1}(1-p)^{\beta-1}$$

on the unit interval. In Section 2.9, we generalize the beta distribution to the Dirichlet distribution. In the meantime, the reader may recall the moment calculation

$$\begin{aligned} \mathrm{E}[P^i(1-P)^j] &= \int_0^1 p^i(1-p)^j f_{\alpha\beta}(p)\,dp \\ &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(\alpha+i)\Gamma(\beta+j)}{\Gamma(\alpha+\beta+i+j)} \int_0^1 f_{\alpha+i,\beta+j}(p)\,dp \\ &= \frac{(\alpha+i-1)\cdots\alpha(\beta+j-1)\cdots\beta}{(\alpha+\beta+i+j-1)\cdots(\alpha+\beta)} \end{aligned}$$

invoking the factorial property $\Gamma(x+1) = x\Gamma(x)$ of the gamma function. This gives, for example,

$$\begin{aligned} \mathrm{E}(P) &= \frac{\alpha}{\alpha+\beta} \\ \mathrm{E}[P(1-P)] &= \frac{\alpha\beta}{(\alpha+\beta)(\alpha+\beta+1)} \\ \mathrm{Var}(P) &= \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}. \end{aligned} \tag{2.4}$$

Now suppose we carry out n Bernoulli trials with the same random success probability P pertaining to all n trials. The number of successes S_n follows a beta-binomial distribution. Application of equations (1.3) and

(1.11) yields

$$\begin{aligned} \mathrm{E}(S_n) &= n\,\mathrm{E}(P) \\ \mathrm{Var}(S_n) &= \mathrm{E}[\mathrm{Var}(S_n \mid P)] + \mathrm{Var}[\mathrm{E}(S_n \mid P)] \\ &= \mathrm{E}[nP(1-P)] + \mathrm{Var}(nP) \\ &= n\,\mathrm{E}[P(1-P)] + n^2\,\mathrm{Var}(P), \end{aligned}$$

which can be explicitly evaluated using the moments in equation (2.4). Problem 4 provides the density of S_n. ■

Example 2.3.2 *Repeated Uniform Sampling*

Suppose we construct a sequence of dependent random variables X_n by taking $X_0 = 1$ and sampling X_n uniformly from the interval $[0, X_{n-1}]$. To calculate the moments of X_n, we use the facts $\mathrm{E}(X_n^k) = \mathrm{E}[\mathrm{E}(X_n^k \mid X_{n-1})]$ and

$$\begin{aligned} \mathrm{E}(X_n^k \mid X_{n-1}) &= \frac{1}{X_{n-1}} \int_0^{X_{n-1}} x^k dx \\ &= \frac{x^{k+1}}{X_{n-1}(k+1)}\Big|_0^{X_{n-1}} \\ &= \frac{1}{k+1} X_{n-1}^k. \end{aligned}$$

Hence,

$$\mathrm{E}(X_n^k) = \frac{1}{k+1}\,\mathrm{E}(X_{n-1}^k) = \left(\frac{1}{k+1}\right)^n.$$

For example, $\mathrm{E}(X_n) = 2^{-n}$ and $\mathrm{Var}(X_n) = 3^{-n} - 2^{-2n}$. It is interesting that if we standardize by defining $Y_n = 2^n X_n$, then the mean $\mathrm{E}(Y_n) = 1$ is stable, but the variance $\mathrm{Var}(Y_n) = (\frac{4}{3})^n - 1$ tends to ∞.

The clouds of mystery lift a little when we rewrite X_n as the product

$$X_n = U_n X_{n-1} = \prod_{i=1}^n U_i$$

of independent uniform random variables $U_1, \ldots, U_n$ on $[0, 1]$. The product rule for expectations now gives $\mathrm{E}(X_n^k) = \mathrm{E}(U_1^k)^n = (k+1)^{-n}$. Although we cannot stabilize X_n, it is possible to stabilize $\ln X_n$. Indeed, Problem 5 notes that $\ln X_n = \sum_{i=1}^n \ln U_i$ follows a $-\frac{1}{2}\chi_{2n}^2$ distribution with mean $-n$ and variance n. Thus for large n, the central limit theorem implies that $(\ln X_n + n)/\sqrt{n}$ has an approximate standard normal distribution. ■

Example 2.3.3 *Expected Family Size*

A married couple desires a family consisting of at least s sons and d daughters. At each birth, the mother independently bears a son with probability p and a daughter with probability $q = 1 - p$. They will quit having children when their objective is reached. Let N_{sd} be the random number of children born to them. Suppose we wish to calculate the expected value $\mathrm{E}(N_{sd})$. Two cases are trivial. If either $s = 0$ or $d = 0$, then N_{sd} follows a negative binomial distribution. Therefore, $\mathrm{E}(N_{0d}) = d/q$ and $\mathrm{E}(N_{s0}) = s/p$. When both s and d are positive, the distribution of N_{sd} is not so obvious. Conditional on the sex of the first child, the random variable $N_{sd} - 1$ is either a probabilistic copy $N^*_{s-1,d}$ of $N_{s-1,d}$ or a probabilistic copy $N^*_{s,d-1}$ of $N_{s,d-1}$. Because in both cases the copy is independent of the sex of the first child, the recurrence relation

$$\begin{aligned} \mathrm{E}(N_{sd}) &= p[1 + \mathrm{E}(N_{s-1,d})] + q[1 + \mathrm{E}(N_{s,d-1})] \\ &= 1 + p\,\mathrm{E}(N_{s-1,d}) + q\,\mathrm{E}(N_{s,d-1}) \end{aligned}$$

follows from conditioning on this outcome.

There are many variations on this idea. For instance, suppose we wish to compute the probability R_{sd} that they reach their quota of s sons before their quota of d daughters. Then the R_{sd} satisfy the boundary conditions $R_{0d} = 1$ for $d > 0$ and $R_{s0} = 0$ for $s > 0$. When s and d are both positive, we have the recurrence relation

$$R_{sd} = pR_{s-1,d} + qR_{s,d-1}.$$

2.4 Moment Transforms

Each of the moment transforms reviewed in Section 1.5 can be differentiated to capture the moments of a random variable. Equally important, these transforms often solve other theoretical problems with surprising ease. The next seven examples illustrate these two roles.

Example 2.4.1 *Characteristic Function of a Normal Density*

To find the characteristic function $\hat{\psi}(t) = \mathrm{E}(e^{itX})$ of a standard normal random variable X with density $\psi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$, we derive and solve a differential equation. Differentiation under the integral sign and integration by parts together imply that

$$\frac{d}{dt}\hat{\psi}(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{itx} ix e^{-\frac{x^2}{2}} dx$$

$$\begin{aligned} &= -\frac{i}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{itx} \frac{d}{dx} e^{-\frac{x^2}{2}} dx \\ &= \frac{-i}{\sqrt{2\pi}} e^{itx} e^{-\frac{x^2}{2}} \Big|_{-\infty}^{\infty} - \frac{t}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{itx} e^{-\frac{x^2}{2}} dx \\ &= -t\hat{\psi}(t). \end{aligned}$$

The unique solution to this differential equation with initial value $\hat{\psi}(0) = 1$ is $\hat{\psi}(t) = e^{-t^2/2}$.

If X is a standard normal random variable, then $\mu + \sigma X$ is a normal random variable with mean μ and variance σ^2. The general identity $\mathrm{E}[e^{it(\mu+\sigma X)}] = e^{it\mu}\,\mathrm{E}[e^{i(\sigma t)X}]$ permits us to express the characteristic function of the normal distribution with mean μ and variance σ^2 as

$$\hat{\psi}_{\mu,\sigma^2}(t) = e^{it\mu}\hat{\psi}(\sigma t) = e^{it\mu - \frac{\sigma^2 t^2}{2}}.$$

The first two derivatives

$$\begin{aligned} \frac{d}{dt}\hat{\psi}_{\mu,\sigma^2}(t) &= (i\mu - \sigma^2 t) e^{it\mu - \frac{\sigma^2 t^2}{2}} \\ \frac{d^2}{dt^2}\hat{\psi}_{\mu,\sigma^2}(t) &= -\sigma^2 e^{it\mu - \frac{\sigma^2 t^2}{2}} + (i\mu - \sigma^2 t)^2 e^{it\mu - \frac{\sigma^2 t^2}{2}} \end{aligned}$$

evaluated at $t = 0$ determine the mean μ and second moment $\sigma^2 + \mu^2$ as indicated in equation (1.8). ■

Example 2.4.2 *Characteristic Function of a Gamma Density*

A random variable X with exponential density $\lambda e^{-\lambda x} 1_{\{x>0\}}$ has characteristic function

$$\begin{aligned} \int_0^{\infty} e^{itx} \lambda e^{-\lambda x} dx &= \frac{\lambda}{it - \lambda} e^{(it-\lambda)x} \Big|_0^{\infty} \\ &= \frac{\lambda}{\lambda - it}. \end{aligned}$$

An analogous calculation yields the Laplace transform $\lambda/(\lambda + t)$. Differentiation of either of these transforms produces

$$\begin{aligned} \mathrm{E}(X) &= \frac{1}{\lambda} \\ \mathrm{Var}(X) &= \frac{1}{\lambda^2}. \end{aligned}$$

The gamma density $\lambda^n x^{n-1} e^{-\lambda x} 1_{\{x>0\}}/\Gamma(n)$ is the convolution of n exponential densities with common intensity λ. The corresponding random variable X_n therefore has

$$\mathrm{E}(X_n) = \frac{n}{\lambda}$$

$$\begin{aligned}
\mathrm{Var}(X_n) &= \frac{n}{\lambda^2} \\
\mathrm{E}(e^{itX_n}) &= \left(\frac{\lambda}{\lambda - it}\right)^n \\
\mathrm{E}(e^{-tX_n}) &= \left(\frac{\lambda}{\lambda + t}\right)^n.
\end{aligned}$$

Problem 16 indicates that these results carry over to non-integer $n > 0$. ■

Example 2.4.3 *Factorial Moments*

Let X be a nonnegative, integer-valued random variable. Repeated differentiation of its probability generating function $G(u) = \mathrm{E}(u^X)$ yields its factorial moments $\mathrm{E}[X(X-1)\cdots(X-j+1)] = \frac{d^j}{du^j}G(1)$. The first two central moments

$$\begin{aligned}
\mathrm{E}(X) &= G'(1) \\
\mathrm{Var}(X) &= \mathrm{E}[X(X-1)] + \mathrm{E}(X) - \mathrm{E}(X)^2 \\
&= G''(1) + G'(1) - G'(1)^2
\end{aligned}$$

are worth committing to memory. As an example, suppose X is Poisson distributed with mean λ. Then

$$G(u) = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} u^k = e^{-\lambda(1-u)}.$$

Repeated differentiation yields $\frac{d^j}{du^j}G(1) = \lambda^j$. In particular, $\mathrm{E}(X) = \lambda$ and $\mathrm{Var}(X) = \lambda$. For another example, let X follow a binomial distribution with n trials and success probability p per trial. In this case $G(u) = (1-p+pu)^n$, $\mathrm{E}(X) = np$, and $\mathrm{Var}(X) = np(1-p)$. ■

Example 2.4.4 *Random Sums*

Suppose $X_1, X_2, \ldots$ is a sequence of independent identically distributed (i.i.d.) random variables. Consider the random sum $S_N = \sum_{i=1}^{N} X_i$, where the random number of terms N is independent of the X_i, and where we adopt the convention $S_0 = 0$. For example in an ecological study, the number of animal litters N in a plot of land might have a Poisson distribution to a good approximation. The random variable X_i then represents the number of offspring in litter i, and the compound Poisson random variable S_N counts the number of offspring over the whole plot.

If N has probability generating function $G(u) = \mathrm{E}(u^N)$, then the characteristic function of S_N is

$$\mathrm{E}(e^{itS_N}) = \sum_{n=0}^{\infty} \mathrm{E}(e^{itS_N} \mid N = n) \Pr(N = n)$$

$$\begin{aligned} &= \sum_{n=0}^{\infty} \mathrm{E}(e^{itX_1})^n \Pr(N=n) \\ &= G[\mathrm{E}(e^{itX_1})]. \end{aligned}$$

This composition rule carries over to other moment transforms. For instance, if the X_i are nonnegative and integer-valued with probability generating function $H(u)$, then a similar argument gives $\mathrm{E}(u^{S_N}) = G[H(u)]$.

We can extract the moments of S_N by differentiation. Alternatively, conditioning on N produces

$$\mathrm{E}(S_N) = \mathrm{E}[N\,\mathrm{E}(X_1)] = \mathrm{E}(N)\,\mathrm{E}(X_1)$$

and

$$\begin{aligned} \mathrm{Var}(S_N) &= \mathrm{E}[\mathrm{Var}(S_N \mid N)] + \mathrm{Var}[\mathrm{E}(S_N \mid N)] \\ &= \mathrm{E}[N\,\mathrm{Var}(X_1)] + \mathrm{Var}[N\,\mathrm{E}(X_1)] \\ &= \mathrm{E}(N)\,\mathrm{Var}(X_1) + \mathrm{Var}(N)\,\mathrm{E}(X_1)^2. \end{aligned}$$

For instance, if N has a Poisson distribution with mean λ and the X_i have a binomial distribution with parameters n and p, then S_N has

$$\begin{aligned} \mathrm{E}(u^{S_N}) &= e^{-\lambda[1-(1-p+pu)^n]} \\ \mathrm{E}(S_N) &= \lambda np \\ \mathrm{Var}(S_N) &= \lambda np(1-p) + \lambda n^2 p^2 \end{aligned}$$

as its probability generating function, mean, and variance, respectively. ■

Example 2.4.5 *Sum of Uniforms*

In Example 2.3.2, we considered the product of n independent random variables $U_1, \ldots, U_n$ uniformly distributed on $[0,1]$. We now turn to the problem of finding the density of the sum $S_n = U_1 + \cdots + U_n$. Our strategy will be to calculate and invert the Laplace transform of the density of S_n, keeping in mind that the Laplace transform of a random variable coincides with the Laplace transform of its density. Because the Laplace transform of a single U_i is $\int_0^1 e^{-tx}dx = (1-e^{-t})/t$, the Laplace transform of S_n is

$$\frac{(1-e^{-t})^n}{t^n} = \frac{1}{t^n}\sum_{k=0}^{n}\binom{n}{k}(-1)^k e^{-kt}.$$

In view of the linearity of the Laplace transform, it therefore suffices to invert the term e^{-kt}/t^n. Since multiplication by e^{-kt} in the transform domain corresponds to an argument shift of k in the original domain, all we need to do is find the function with transform t^{-n} and shift it by k. We now make an inspired guess that the function x^{n-1} is relevant. Because

the Laplace transform deals with functions defined on $[0, \infty)$, we exchange x^{n-1} for the function $(x)_+^{n-1}$, which equals 0 for $x \leq 0$ and x^{n-1} for $x > 0$. The change of variables $u = tx$ and the definition of the gamma function show that $(x)_+^{n-1}$ has transform

$$\begin{aligned} \int_0^\infty x^{n-1} e^{-tx} dx &= \frac{1}{t^n} \int_0^\infty u^{n-1} e^{-u} du \\ &= \frac{(n-1)!}{t^n}. \end{aligned}$$

Up to a constant, this is just what we need. Hence, we conclude that S_n has density

$$f(x) = \frac{1}{(n-1)!} \sum_{k=0}^{n} \binom{n}{k} (-1)^k (x-k)_+^{n-1}.$$

The corresponding distribution function

$$F(x) = \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} (-1)^k (x-k)_+^n$$

emerges after integration with respect to x. ■

Example 2.4.6 *A Nonexistence Problem*

Is it always possible to represent a random variable X as the difference $Y - Z$ of two independent, identically distributed random variables Y and Z? The answer is clearly no unless X is symmetrically distributed around 0. For a symmetrically distributed X, the question is more subtle. Suppose that Y and Z exist for such an X. Then the characteristic function of X reduces to

$$\begin{aligned} \mathrm{E}[e^{it(Y-Z)}] &= \mathrm{E}(e^{itY}) \, \mathrm{E}(e^{-itZ}) \\ &= \mathrm{E}(e^{itY}) \, \mathrm{E}(e^{itY})^* \\ &= |\, \mathrm{E}(e^{itY})|^2, \end{aligned}$$

where the superscript $*$ denotes complex conjugation. (It is trivial to check that conjugation commutes with expectation for complex random variables possessing only a finite number of values, and this property persists in the limit for all complex random variables.) In any case, if the representation $X = Y - Z$ holds, then the characteristic function of X is nonnegative. Thus, to construct a counterexample, all we need to do is find a symmetrically distributed random variable whose characteristic function fails the test of nonnegativity. For instance, if we take X to be uniformly distributed on $[-\frac{1}{2}, \frac{1}{2}]$, then its characteristic function

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} e^{itx} dx = \frac{e^{itx}}{it} \Big|_{-\frac{1}{2}}^{\frac{1}{2}} = \frac{\sin(\frac{t}{2})}{\frac{t}{2}}$$

oscillates in sign. ■

Example 2.4.7 *Characterization of the Standard Normal Distribution*

Consider a random variable X with mean 0, variance 1, and characteristic function $\hat{\psi}(t)$. If X is standard normal and Y is an independent copy of X, then for all a and b the sum $aX + bY$ has the same distribution as $\sqrt{a^2 + b^2}X$. This distributional identity implies the characteristic function identity

$$\hat{\psi}(at)\hat{\psi}(bt) \quad = \quad \hat{\psi}\left(\sqrt{a^2 + b^2}t\right) \tag{2.5}$$

for all t.

Conversely, suppose the functional equation (2.5) holds for a random variable X with mean 0 and variance 1. Let us show that X possesses a standard normal distribution. The special case $a = -1$ and $b = 0$ of equation (2.5) amounts to $\hat{\psi}(-t) = \hat{\psi}(t)$, from which it immediately follows that $\hat{\psi}(t)^* = \hat{\psi}(-t) = \hat{\psi}(t)$. Thus, $\hat{\psi}(t)$ is real and even. It is also differentiable because $\mathrm{E}(|X|) < \infty$. Now define $g(t^2) = \hat{\psi}(t)$ for $t > 0$. Setting $t = 1$ and replacing a^2 by a and b^2 by b in the functional equation (2.5) produce the revised functional equation

$$g(a)g(b) \quad = \quad g(a + b). \tag{2.6}$$

Taking $f(t) = \ln g(t)$ lands us right back at equation (2.3), except that it is no longer clear that $f(t)$ is monotone. Rather than rely on our previous hand-waving solution, we can differentiate equation (2.6), first with respect to $a \geq 0$ and then with respect to $b \geq 0$. This yields

$$g'(a)g(b) \quad = \quad g'(a + b) \quad = \quad g(a)g'(b). \tag{2.7}$$

If we take $b > 0$ sufficiently small, then $g(b) > 0$, owing to the continuity of $g(t)$ and the initial condition $g(0) = 1$. Dividing equation (2.7) by $g(b)$ and defining $\lambda = -g'(b)/g(b)$ leads to the differential equation $g'(a) = -\lambda g(a)$ with solution $g(a) = e^{-\lambda a}$. To determine λ, note that the equality

$$g''(t^2)4t^2 + g'(t^2)2 \quad = \quad \hat{\psi}''(t)$$

yields $-2\lambda = -1$ in the limit as t approaches 0. Thus, $\hat{\psi}(t) = e^{-t^2/2}$ as required. ■

2.5 Tail Probability Methods

Consider a nonnegative random variable X with distribution function $F(x)$. The right-tail probability $\Pr(X > t) = 1 - F(t)$ turns out to be helpful

in calculating certain expectations relative to X. Let $h(t)$ be an integrable function on each finite interval $[0, x]$. If we define $H(x) = H(0) + \int_0^x h(t)\,dt$ and suppose that $\int_0^\infty |h(t)|[1-F(t)]\,dt < \infty$, then Fubini's theorem justifies the calculation

$$\begin{aligned} \mathrm{E}[H(X)] &= H(0) + \mathrm{E}\left[\int_0^X h(t)\,dt\right] \\ &= H(0) + \int_0^\infty \int_0^x h(t)\,dt\,dF(x) \\ &= H(0) + \int_0^\infty \int_t^\infty dF(x)h(t)\,dt \qquad (2.8) \\ &= H(0) + \int_0^\infty h(t)[1-F(t)]\,dt. \end{aligned}$$

If X is concentrated on the integers $\{0, 1, 2, \ldots\}$, the right-tail probability $1 - F(t)$ is constant except for jumps at these integers. Equation (2.8) therefore reduces to

$$\mathrm{E}[H(X)] = H(0) + \sum_{k=0}^{\infty} [H(k+1) - H(k)][1 - F(k)]. \qquad (2.9)$$

Example 2.5.1 *Moments from Right-Tail Probabilities*

The choices $h(t) = nt^{n-1}$ and $H(0) = 0$ yield $H(x) = x^n$. Hence, equations (2.8) and (2.9) become

$$\mathrm{E}[X^n] = n\int_0^\infty t^{n-1}[1 - F(t)]\,dt$$

and

$$\mathrm{E}[X^n] = \sum_{k=0}^{\infty} [(k+1)^n - k^n][1 - F(k)],$$

respectively. For instance, if X is exponentially distributed, then the right-tail probability $1 - F(t) = e^{-\lambda t}$ and $E(X) = \int_0^\infty e^{-\lambda t}dt = \lambda^{-1}$. If X is geometrically distributed with failure probability q, then $1 - F(k) = q^k$ and $\mathrm{E}(X) = \sum_{k=0}^\infty q^k = (1-q)^{-1}$. ■

Example 2.5.2 *Laplace Transforms*

Equation (2.8) also determines the relationship between the Laplace transform $\mathrm{E}(e^{-sX})$ of a nonnegative random variable X and the ordinary Laplace transform $\tilde{F}(s)$ of its distribution function $F(x)$. For this purpose, we choose $h(t) = -se^{-st}$ and $H(0) = 1$. The resulting integral $H(x) = e^{-sx}$

and equation (2.8) together yield the formula

$$\begin{aligned} \mathrm{E}(e^{-sX}) &= 1 - s\int_0^\infty e^{-st}[1-F(t)]\,dt \\ &= s\int_0^\infty e^{-st}F(t)\,dt \\ &= s\tilde{F}(s). \end{aligned}$$

For example, if X is exponentially distributed with intensity λ, then the Laplace transform $\mathrm{E}(e^{-sX}) = \lambda/(s+\lambda)$ mentioned in Example 2.4.2 leads to $\tilde{F}(s) = \lambda/[s(s+\lambda)]$. ■

2.6 Moments of Reciprocals and Ratios

Ordinarily we differentiate Laplace transforms to recover moments. However, to recover an inverse moment, we need to integrate [43]. Suppose X is a positive random variable with Laplace transform $L(t)$. If $n > 0$, then Fubini's theorem and the change of variables $s = tX$ shows that

$$\begin{aligned} \int_0^\infty t^{n-1}L(t)\,dt &= \mathrm{E}\left(\int_0^\infty t^{n-1}e^{-tX}dt\right) \\ &= \mathrm{E}\left(X^{-n}\int_0^\infty s^{n-1}e^{-s}ds\right) \\ &= \mathrm{E}\left(X^{-n}\right)\Gamma(n). \end{aligned}$$

The formula

$$\mathrm{E}\left(X^{-n}\right) = \frac{1}{\Gamma(n)}\int_0^\infty t^{n-1}L(t)\,dt \tag{2.10}$$

can be evaluated exactly in some cases. In other cases, for instance when n fails to be an integer, the formula can be evaluated numerically.

Example 2.6.1 *Mean and Variance of an Inverse Gamma*

Because a gamma random variable X with intensity λ and shape parameter β has Laplace transform $L(t) = [\lambda/(t+\lambda)]^\beta$, formula (2.10) gives

$$\begin{aligned} \mathrm{E}\left(X^{-1}\right) &= \int_0^\infty \left(\frac{\lambda}{t+\lambda}\right)^\beta dt \\ &= \frac{\lambda}{\beta-1} \end{aligned}$$

for $\beta > 1$ and

$$\mathrm{E}\left(X^{-2}\right) = \int_0^\infty t\left(\frac{\lambda}{t+\lambda}\right)^\beta dt$$

$$\begin{aligned} &= \int_0^\infty \lambda\Big(\frac{\lambda}{t+\lambda}\Big)^{\beta-1} dt - \int_0^\infty \lambda\Big(\frac{\lambda}{t+\lambda}\Big)^{\beta} dt \\ &= \frac{\lambda^2}{\beta-2} - \frac{\lambda^2}{\beta-1} \end{aligned}$$

for $\beta > 2$. It follows that $\mathrm{Var}(X^{-1}) = \lambda^2/[(\beta-1)^2(\beta-2)]$ for $\beta > 2$. ■

To calculate the expectation of a ratio X^m/Y^n for a positive random variable Y and an arbitrary random variable X, we consider the mixed characteristic function and Laplace transform $M(s,t) = \mathrm{E}(e^{isX-tY})$. Assuming that $\mathrm{E}(|X|^m) < \infty$ for some positive integer m, we can write

$$\frac{\partial^m}{\partial s^m} M(s,t) = \mathrm{E}\left[(iX)^m e^{isX-tY}\right]$$

by virtue of Example 1.2.5 with dominating random variable $|X|^k e^{-tY}$ for the kth partial derivative. For $n > 0$ and $\mathrm{E}(|X|^m Y^{-n}) < \infty$, we now invoke Fubini's theorem and calculate

$$\begin{aligned} \int_0^\infty t^{n-1} \frac{\partial^m}{\partial s^m} M(0,t)\, dt &= \int_0^\infty t^{n-1}\, \mathrm{E}\left[(iX)^m e^{-tY}\right] dt \\ &= \mathrm{E}\left[\int_0^\infty t^{n-1} (iX)^m e^{-tY} dt\right] \\ &= \mathrm{E}\left[\frac{(iX)^m}{Y^n} \int_0^\infty r^{n-1} e^{-r} dr\right] \\ &= \mathrm{E}\left[\frac{(iX)^m}{Y^n}\right] \Gamma(n). \end{aligned}$$

Rearranging this yields

$$\mathrm{E}\left(\frac{X^m}{Y^n}\right) = \frac{1}{i^m \Gamma(n)} \int_0^\infty t^{n-1} \frac{\partial^m}{\partial s^m} M(0,t)\, dt. \tag{2.11}$$

Example 2.6.2 *Mean of a Beta Random Variable*

If U and V are independent gamma random variables with common intensity λ and shape parameters α and β, then the ratio $U/(U+V)$ has a beta distribution with parameters α and β. The reader is asked to prove this fact in Problem 32. It follows that the mixed characteristic function and Laplace transform

$$\begin{aligned} M_{U,U+V}(s,t) &= \mathrm{E}\left[e^{isU-t(U+V)}\right] \\ &= \mathrm{E}\left[e^{(is-t)U}\right] \mathrm{E}\left[e^{-tV}\right] \\ &= \Big(\frac{\lambda}{\lambda - is + t}\Big)^{\alpha} \Big(\frac{\lambda}{\lambda + t}\Big)^{\beta}. \end{aligned}$$

Equation (2.11) consequently gives the mean of the beta distribution as

$$\begin{aligned}
\mathrm{E}\left(\frac{U}{U+V}\right) &= \frac{1}{i}\int_0^\infty \alpha i \frac{\lambda^\alpha}{(\lambda+t)^{\alpha+1}}\left(\frac{\lambda}{\lambda+t}\right)^\beta dt \\
&= -\frac{\alpha\lambda^{\alpha+\beta}}{(\alpha+\beta)(\lambda+t)^{\alpha+\beta}}\Big|_0^\infty \\
&= \frac{\alpha}{\alpha+\beta}.
\end{aligned}$$

2.7 Reduction of Degree

The method of reduction of degree also involves recurrence relations. However, instead of creating these by conditioning, we now employ integration by parts and simple algebraic transformations. The Stein and Chen lemmas given below find their most important applications in the approximation theories featured in the books [18, 187].

Example 2.7.1 *Stein's Lemma*

Suppose X is normally distributed with mean μ and variance σ^2 and $g(x)$ is a differentiable function such that $|g(X)(X-\mu)|$ and $|g'(X)|$ have finite expectations. Stein's lemma [187] asserts that

$$\mathrm{E}[g(X)(X-\mu)] = \sigma^2\,\mathrm{E}[g'(X)].$$

To prove this formula, we note that integration by parts produces

$$\begin{aligned}
&\frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^\infty g(x)(x-\mu)e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx \\
&= \lim_{a_n\to-\infty}\lim_{b_n\to\infty}\left[\frac{-\sigma^2 g(x)e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}}\Big|_{a_n}^{b_n} + \frac{\sigma^2}{\sqrt{2\pi\sigma^2}}\int_{a_n}^{b_n} g'(x)e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx\right] \\
&= \frac{\sigma^2}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^\infty g'(x)e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx.
\end{aligned}$$

The boundary terms vanish for carefully chosen sequences a_n and b_n tending to $\pm\infty$ because the integrable function $|g(x)(x-\mu)|\exp[-\frac{(x-\mu)^2}{2\sigma^2}]$ cannot be bounded away from 0 as $|x|$ tends to ∞. To illustrate the repeated application of Stein's lemma, take $g(x) = (x-\mu)^{2n-1}$. Then the important moment identity

$$\begin{aligned}
\mathrm{E}[(X-\mu)^{2n}] &= \sigma^2(2n-1)\,\mathrm{E}[(X-\mu)^{2n-2}] \\
&= \sigma^{2n}(2n-1)(2n-3)\cdots 1
\end{aligned}$$

follows immediately. ■

Example 2.7.2 *Reduction of Degree for the Gamma*

A random variable X with gamma density $\lambda^\alpha x^{\alpha-1}e^{-\lambda x}/\Gamma(\alpha)$ on $(0,\infty)$ satisfies the analogous reduction of degree formula

$$\mathrm{E}[g(X)X] = \frac{1}{\lambda}\mathrm{E}[g'(X)X] + \frac{\alpha}{\lambda}\mathrm{E}[g(X)].$$

Provided the required moments exist and $\lim_{x\to 0} g(x)x^\alpha = 0$, the integration by parts calculation

$$\begin{aligned}
&\frac{\lambda^\alpha}{\Gamma(\alpha)}\int_0^\infty g(x)xx^{\alpha-1}e^{-\lambda x}dx \\
= &\frac{\lambda^\alpha}{\Gamma(\alpha)\lambda}\left[-g(x)xx^{\alpha-1}e^{-\lambda x}\Big|_0^\infty + \int_0^\infty g'(x)xx^{\alpha-1}e^{-\lambda x}dx\right. \\
&\left. + \int_0^\infty g(x)\alpha x^{\alpha-1}e^{-\lambda x}dx\right]
\end{aligned}$$

is valid. The special case $g(x) = x^{n-1}$ yields the recurrence relation

$$\mathrm{E}(X^n) = \frac{(n-1+\alpha)}{\lambda}\mathrm{E}(X^{n-1})$$

for the moments of X. ■

Example 2.7.3 *Chen's Lemma*

Chen [36] pursues the formula $\mathrm{E}[Zg(Z)] = \lambda\mathrm{E}[g(Z+1)]$ for a Poisson random variable Z with mean λ as a kind of discrete analog to Stein's lemma. The proof of Chen's result

$$\begin{aligned}
\sum_{j=0}^\infty jg(j)\frac{\lambda^j e^{-\lambda}}{j!} &= \lambda\sum_{j=1}^\infty g(j)\frac{\lambda^{j-1}e^{-\lambda}}{(j-1)!} \\
&= \lambda\sum_{k=0}^\infty g(k+1)\frac{\lambda^k e^{-\lambda}}{k!}
\end{aligned}$$

is almost trivial. The choice $g(z) = z^{n-1}$ gives the recurrence relation

$$\begin{aligned}
\mathrm{E}(Z^n) &= \lambda\mathrm{E}[(Z+1)^{n-1}] \\
&= \lambda\sum_{k=0}^{n-1}\binom{n-1}{k}\mathrm{E}(Z^k)
\end{aligned}$$

for the moments of Z. ■

2.8 Spherical Surface Measure

In this section and the next, we explore probability measures on surfaces. Surface measures are usually treated using differential forms and manifolds [185]. With enough symmetry, one can dispense with these complicated mathematical objects and fall back on integration on R^n. This concrete approach has the added benefit of facilitating the calculation of certain expectations.

Let $g(\|x\|)$ be any probability density such as $e^{-\pi\|x\|^2}$ on R^n that depends only on the Euclidean distance $\|x\|$ of a point x from the origin. Given a choice of $g(\|x\|)$, one can define the integral of a continuous, real-valued function $f(s)$ on the unit sphere $S_{n-1} = \{x \in \mathsf{R}^n : \|x\| = 1\}$ by

$$\int_{S_{n-1}} f(s)\, d\omega_{n-1}(s) \quad = \quad a_{n-1} \int f\left(\frac{x}{\|x\|}\right) g(\|x\|)\, dx \tag{2.12}$$

for a positive constant a_{n-1} to be specified [16]. It is trivial to show that this yields an invariant integral in the sense that

$$\int_{S_{n-1}} f(Ts)\, d\omega_{n-1}(s) \quad = \quad \int_{S_{n-1}} f(s)\, d\omega_{n-1}(s)$$

for any orthogonal transformation T. In this regard note that $|\det(T)| = 1$ and $\|Tx\| = \|x\|$. Taking $f(s) = 1$ produces a total mass of a_{n-1} for the surface measure ω_{n-1}.

Of course, the constant a_{n-1} is hardly arbitrary. We can pin it down by proving the product measure formula

$$\int h(x)\, dx \quad = \quad \int_0^\infty \int_{S_{n-1}} h(rs)\, d\omega_{n-1}(s)\, r^{n-1}\, dr \tag{2.13}$$

for any integrable function $h(x)$ on R^n. Formula (2.13) says that we can integrate over R^n by cumulating the surface integrals over successive spherical shells. To prove (2.13), we interchange orders of integration as needed and execute the successive changes of variables $t = r\|x\|^{-1}$, $z = tx$, and $t = \|z\|r^{-1}$. These maneuvers turn the right-hand side of formula (2.13) into

$$\begin{aligned}
& \int_0^\infty \int_{S_{n-1}} h(rs)\, d\omega_{n-1}(s)\, r^{n-1}\, dr \\
= \quad & a_{n-1} \int_0^\infty \int h(rx/\|x\|) g(\|x\|)\, dx\, r^{n-1}\, dr \\
= \quad & a_{n-1} \int \int_0^\infty h(rx/\|x\|)\, r^{n-1}\, dr\, g(\|x\|)\, dx \\
= \quad & a_{n-1} \int \int_0^\infty h(tx)(t\|x\|)^{n-1}\|x\|\, dt\, g(\|x\|)\, dx
\end{aligned}$$

$$\begin{aligned}
&= a_{n-1}\int_0^\infty \int h(tx)\|x\|^n g(\|x\|)\,dx\,t^{n-1}\,dt \\
&= a_{n-1}\int_0^\infty \int h(z)(\|z\|/t)^n g(\|z\|/t)\,t^{-n}\,dz\,t^{n-1}\,dt \\
&= a_{n-1}\int\int_0^\infty (\|z\|/t)^{n-1} g(\|z\|/t)\|z\|t^{-2}\,dt\,h(z)\,dz \\
&= a_{n-1}\int_0^\infty r^{n-1}g(r)\,dr\int h(z)\,dz.
\end{aligned}$$

This establishes equality (2.13) provided we take $a_{n-1}\int_0^\infty r^{n-1}g(r)\,dr = 1$. For the choice $g(r) = e^{-\pi r^2}$, we calculate

$$\int_0^\infty r^{n-1}e^{-\pi r^2}dr = \int_0^\infty \left(\frac{t}{\pi}\right)^{(n-2)/2} e^{-t}\frac{1}{2\pi}\,dt = \frac{\Gamma(\frac{n}{2})}{2\pi^{n/2}}. \tag{2.14}$$

Thus, the surface area a_{n-1} of S_{n-1} reduces to $2\pi^{n/2}/\Gamma(\frac{n}{2})$. Omitting the constant a_{n-1} in the definition (2.12) yields the uniform probability distribution on S_{n-1}.

Besides offering a method of evaluating integrals, formula (2.13) demonstrates that the definition of surface measure does not depend on the choice of the function $g(\|x\|)$. In fact, consider the extension

$$h(x) = f\left(\frac{x}{\|x\|}\right)1_{\{1\le\|x\|\le c\}}$$

of a function $f(x)$ on S_{n-1}. If we take $c = \sqrt[n]{n+1}$, then $\int_1^c r^{n-1}\,dr = 1$, and formula (2.13) amounts to

$$\int_{S_{n-1}} f(s)\,d\omega_{n-1}(s) = \frac{\int h(x)\,dx}{\int_1^c r^{n-1}dr} = \int h(x)\,dx,$$

which, as Baker notes [16], affords a definition of the surface integral that does not depend on the choice of the probability density $g(\|x\|)$. As a by-product of this result, it follows that the surface area a_{n-1} of S_{n-1} also does not depend on $g(\|x\|)$.

Example 2.8.1 *Moments of $\|x\|$ Relative to $e^{-\pi\|x\|^2}$*

Formula (2.13) gives

$$\begin{aligned}
\int \|x\|^k e^{-\pi\|x\|^2}dx &= \int_0^\infty \int_{S_{n-1}} r^k e^{-\pi r^2}\,d\omega_{n-1}(s)r^{n-1}dr \\
&= a_{n-1}\int_0^\infty r^{n+k-1}e^{-\pi r^2}dr \\
&= \frac{a_{n-1}}{a_{n+k-1}}.
\end{aligned}$$

Negative as well as positive values of $k > -n$ are permitted. ■

Example 2.8.2 *Integral of a Polynomial*

The function $f(x) = x_1^{k_1} \cdots x_n^{k_n}$ is a monomial when $k_1, \ldots, k_n$ are nonnegative integers. A linear combination of monomials is a polynomial. To find the integral of $f(x)$ on S_{n-1}, it is convenient to put $k = \sum_{j=1}^n k_j$ and use the probability density $g(\|x\|) = a_{n+k-1}\|x\|^k e^{-\pi\|x\|^2}/a_{n-1}$. With these choices,

$$\begin{aligned}
\int_{S_{n-1}} f(s)\, d\omega_{n-1}(s) &= a_{n-1}\frac{a_{n+k-1}}{a_{n-1}} \int \frac{x_1^{k_1} \cdots x_n^{k_n}}{\|x\|^k} \|x\|^k e^{-\pi\|x\|^2} dx \\
&= a_{n+k-1} \int x_1^{k_1} \cdots x_n^{k_n} e^{-\pi\|x\|^2} dx \\
&= a_{n+k-1} \prod_{j=1}^n \int_{-\infty}^{\infty} x_j^{k_j} e^{-\pi x_j^2} dx_j.
\end{aligned}$$

If any k_j is odd, then the corresponding one-dimensional integral in the last product vanishes. Hence, the surface integral of the monomial vanishes as well. If all k_j are even, then the same reasoning that produced equation (2.14) leads to

$$\begin{aligned}
\int_{-\infty}^{\infty} x_j^{k_j} e^{-\pi x_j^2} dx_j &= 2\int_0^{\infty} x_j^{k_j} e^{-\pi x_j^2} dx_j \\
&= \frac{\Gamma(\frac{k_j+1}{2})}{\pi^{(k_j+1)/2}}.
\end{aligned}$$

It follows that

$$\begin{aligned}
\int_{S_{n-1}} x_1^{k_1} \cdots x_n^{k_n}\, d\omega_{n-1}(s) &= \frac{2\pi^{(n+k)/2}}{\Gamma(\frac{n+k}{2})} \prod_{j=1}^n \frac{\Gamma(\frac{k_j+1}{2})}{\pi^{(k_j+1)/2}} \\
&= \frac{2\prod_{j=1}^n \Gamma(\frac{k_j+1}{2})}{\Gamma(\frac{n+k}{2})}
\end{aligned}$$

when all k_j are even. ■

2.9 Dirichlet Distribution

The Dirichlet distribution generalizes the beta distribution. As such, it lives on the unit simplex $T_{n-1} = \{x \in \mathsf{R}_+^n : \|x\|_1 = 1\}$, where $\|x\|_1 = \sum_{j=1}^n |x_j|$ and

$$\mathsf{R}_+^n = \{x \in \mathsf{R}^n : x_j > 0,\ j = 1, \ldots, n\}.$$

By analogy with our definition (2.12) of spherical surface measure, one can define the simplex surface measure μ_{n-1} on T_{n-1} through the equation

$$\int_{T_{n-1}} f(s)\,d\mu_{n-1}(s) \;=\; b_{n-1}\int_{\mathsf{R}^n_+} f\left(\frac{x}{\|x\|_1}\right)g(\|x\|_1)\,dx \qquad (2.15)$$

for any continuous function $f(s)$ on T_{n-1}. In this setting, $g(\|x\|_1)$ is a probability density on R^n_+ that depends only on the distance $\|x\|_1$ of x from the origin.

One can easily show that this definition of surface measure is invariant under permutation of the coordinates. One can also prove the product measure formula

$$\int_{\mathsf{R}^n_+} h(x)\,dx \;=\; \frac{1}{\sqrt{n}}\int_0^\infty \int_{T_{n-1}} h(rs)\,d\mu_{n-1}(s)\,r^{n-1}\,dr. \qquad (2.16)$$

The appearance of the factor $1/\sqrt{n}$ here can be explained by appealing to geometric intuition. In formula (2.16) we integrate $h(x)$ by summing its integrals over successive slabs multiplied by the thicknesses of the slabs. Now the thickness of a slab amounts to nothing more than the distance between two slices $(r+dr)T_{n-1}$ and rT_{n-1}. Given that the corresponding centers of mass are $(r+dr)n^{-1}\mathbf{1}$ and $rn^{-1}\mathbf{1}$, the slab thickness is $dr/\sqrt{n}$.

The proof of formula (2.16) is virtually identical to the proof of formula (2.13). In the final stage of the proof, we must set

$$\frac{b_{n-1}}{\sqrt{n}}\int_0^\infty r^{n-1}g(r)\,dr \;=\; 1.$$

The choice $g(r) = e^{-r}$ immediately gives $\int_0^\infty r^{n-1}e^{-r}dr = \Gamma(n)$. It follows that the surface area b_{n-1} of T_{n-1} is $\sqrt{n}/\Gamma(n)$. Omitting the constant b_{n-1} in the definition (2.15) yields the uniform probability distribution on T_{n-1}.

As before we evaluate the moment

$$\begin{aligned}\int_{\mathsf{R}^n_+} \|x\|_1^k e^{-\|x\|_1}dx &= \frac{1}{\sqrt{n}}\int_0^\infty\int_{T_{n-1}} r^k e^{-r}d\mu_{n-1}(s)r^{n-1}dr\\ &= \frac{b_{n-1}}{\sqrt{n}}\int_0^\infty r^{n+k-1}e^{-r}dr\\ &= \frac{\Gamma(n+k)}{\Gamma(n)}.\end{aligned}$$

For the multinomial $f(x) = x_1^{k_1}\cdots x_n^{k_n}$ with $k = \sum_{j=1}^n k_j$, we then evaluate

$$\begin{aligned}\int_{T_{n-1}} f(s)\,d\mu_{n-1}(s) &= b_{n-1}\frac{\Gamma(n)}{\Gamma(n+k)}\int_{\mathsf{R}^n_+}\frac{x_1^{k_1}\cdots x_n^{k_n}}{\|x\|_1^k}\|x\|_1^k e^{-\|x\|_1}dx\\ &= b_{n-1}\frac{\Gamma(n)}{\Gamma(n+k)}\int_{\mathsf{R}^n_+} x_1^{k_1}\cdots x_n^{k_n}e^{-\|x\|_1}dx\end{aligned}$$

$$= \quad b_{n-1}\frac{\Gamma(n)}{\Gamma(n+k)}\prod_{j=1}^{n}\Gamma(k_j+1)$$

using the probability density

$$g(\|x\|_1) \quad = \quad \frac{\Gamma(n)}{\Gamma(n+k)}\|x\|_1^k e^{-\|x\|_1}$$

on R_+^n. This calculation identifies the Dirichlet distribution

$$\frac{\Gamma(k)}{b_{n-1}\Gamma(n)\prod_{j=1}^{n}\Gamma(k_j)}\prod_{j=1}^{n}s_j^{k_j-1}$$

as a probability density on T_{n-1} relative to μ_{n-1} with moments

$$\mathrm{E}\left(s_1^{l_1}\cdots s_n^{l_n}\right) \quad = \quad \frac{\Gamma(k)\prod_{j=1}^{n}\Gamma(k_j+l_j)}{\Gamma(k+l)\prod_{j=1}^{n}\Gamma(k_j)},$$

where $l=\sum_{j=1}^{n}l_j$. Note that $k_j>0$ need not be an integer.

2.10 Problems

1. Let X represent the number of fixed points of a random permutation of the set $\{1,\ldots,n\}$. Demonstrate that X has the falling factorial moment

$$\mathrm{E}[X(X-1)\cdots(X-k+1)] \quad = \quad k!\,\mathrm{E}\left[\binom{X}{k}\right] \quad = \quad 1$$

for $1\le k\le n$. (Hints: Note that $\binom{X}{k}$ is the number of ways of choosing k points among the available fixed points. Choose the points first, and then calculate the probability that they are fixed.)

2. In a certain building, p people enter an elevator stationed on the ground floor. There are n floors above the ground floor, and each is an equally likely destination. If the people exit the elevator independently, then show that the elevator makes on average

$$n\left[1-\left(1-\frac{1}{n}\right)^p\right]$$

stops in discharging all p people.

3. Numbers are drawn randomly from the set $\{1,2,\ldots,n\}$ until their sum exceeds k for $0\le k\le n$. Show that the expected number of draws equals

$$e_k \quad = \quad \left(1+\frac{1}{n}\right)^k.$$

In particular, $e_n\approx e$. (Hint: Show that $e_k=1+\frac{1}{n}[e_0+\cdots+e_{k-1}]$.)

4. Show that the beta-binomial distribution of Example 2.3.1 has discrete density

$$\Pr(S_n = k) = \binom{n}{k} \frac{\Gamma(\alpha+\beta)\Gamma(\alpha+k)\Gamma(\beta+n-k)}{\Gamma(\alpha)\Gamma(\beta)\Gamma(\alpha+\beta+n)}.$$

5. Prove that the $\ln X_n = \sum_{i=1}^n \ln U_i$ random variable of Example 2.3.2 follows a $-\frac{1}{2}\chi^2_{2n}$ distribution.

6. A noncentral chi-square random variable X has a χ^2_{n+2Y} distribution conditional on a Poisson random variable Y with mean λ. Show that $\mathrm{E}(X) = n + 2\lambda$ and $\mathrm{Var}(X) = 2n + 8\lambda$.

7. Consider an urn with $b \geq 1$ black balls and $w \geq 0$ white balls. Balls are extracted from the urn without replacement until a black ball is encountered. Show that the number of balls N_{bw} extracted has mean $\mathrm{E}(N_{bw}) = (b+w+1)/(b+1)$. (Hint: Derive a recurrence relation and boundary conditions for $\mathrm{E}(N_{bw})$ and solve.)

8. Give a recursive method for computing the second moments $\mathrm{E}(N_{sd}^2)$ in the family planning model.

9. In the family planning model, suppose the couple has an upper limit m on the number of children they can afford. Hence, they stop whenever they reach their goal of s sons and d daughters or m total children, whichever comes first. Let N_{sdm} now be their random number of children. Give a recursive method for computing $\mathrm{E}(N_{sdm})$.

10. In table tennis suppose that player B wins a set with probability p and player A wins a set with probability $q = 1 - p$. Each set counts 1 point. The winner of a match is the first to accumulate 21 points and at least 2 points more than the opposing player. How can one calculate the probability that player B wins? Assume that A has already accumulated i points and B has already accumulated j points. Let w_{ij} denote the probability that B wins the match. Let t_{ij} denote the corresponding expected number of further points scored before the match ends. Show that these quantities satisfy the recurrences

$$\begin{aligned} w_{ij} &= pw_{i,j+1} + qw_{i+1,j} \\ t_{ij} &= 1 + pt_{i,j+1} + qt_{i+1,j} \end{aligned}$$

for i and j between 0 and 20. The first recurrence allows one to compute w_{00} from the boundary values $w_{i,21}$ and $w_{21,j}$, where $i \leq 21$ and $j \leq 21$. In these situations, the quota of 21 total points is irrelevant, and only the excess points criterion is operative. The second recurrence has similar implications for the t_{ij}. On the boundary, the

winning probability reduces to either 0 or 1 or to the hitting probability considered in Problem 38 of Chapter 7. Using the results stated there, tabulate or graph w_{00} and t_{00} as a function of p.

11. Consider the integral

$$I(a,p,y) = \int_{-\infty}^{y} \frac{1}{(a+x^2)^p}\,dx$$

for $p > \frac{1}{2}$ and $a > 0$. As an example of the method of parametric differentiation [29], prove that

$$I(a,p+n,y) = \frac{(-1)^n}{p(p+1)\cdots(p+n-1)} \frac{d^n}{da^n} I(a,p,y).$$

In the particular case $p = \frac{3}{2}$, show that

$$I\Big(a,\frac{3}{2},y\Big) = \frac{y}{a\sqrt{a+y^2}} + \frac{1}{a}.$$

Use these facts to show that the t-distribution with $2m$ degrees of freedom has finite expansion

$$\begin{aligned} &\frac{\Gamma(m+1/2)}{\sqrt{2\pi m}\Gamma(m)} \int_{-\infty}^{y} \left(1+\frac{x^2}{2m}\right)^{-m-1/2} dx \\ &= \frac{1}{2\sqrt{2m}} \left[\frac{y}{\sqrt{\pi}} \sum_{j=0}^{m-1} \frac{\Gamma(j+1/2)}{j!} \left(1+\frac{y^2}{2m}\right)^{-j-1/2} + \sqrt{2m} \right]. \end{aligned}$$

(Hints: For the case $p = \frac{3}{2}$ apply the fundamental theorem of calculus. To expand the t-distribution, use Leibniz's rule for differentiating a product.)

12. Let X be a nonnegative integer-valued random variable with probability generating function $Q(s)$. Find the probability generating functions of $X + k$ and kX in terms of $Q(s)$ for any nonnegative integer k.

13. Let $S_n = X_1 + \cdots + X_n$ be the sum of n independent random variables, each distributed uniformly over the set $\{1, 2, \ldots, m\}$. Find the probability generating function of S_n, and use it to calculate $\mathrm{E}(S_n)$ and $\mathrm{Var}(S_n)$.

14. Let $X_1, X_2, \ldots$ be an i.i.d. sequence of Bernoulli random variables with success probability p. Thus, $X_i = 1$ with probability p, and $X_i = 0$ with probability $1 - p$. Demonstrate that the infinite series $S = \sum_{i=1}^{\infty} 2^{-i} X_i$ has mean p and variance $\frac{1}{3}p(1-p)$. When $p = \frac{1}{2}$,

the random variable S is uniformly distributed on $[0,1]$, and X_i can be interpreted as the ith binary digit of S [192]. In this special case also prove the well-known identity

$$\frac{\sin\theta}{\theta} = \prod_{j=1}^{\infty} \cos\left(\frac{\theta}{2^j}\right)$$

by calculating the characteristic function of $S - \frac{1}{2}$ in two different ways.

15. Consider a sequence $X_1, X_2, \ldots$ of independent, integer-valued random variables with common logarithmic distribution

$$\Pr(X_i = k) = -\frac{q^k}{k\ln(1-q)}$$

for $k \geq 1$. Let N be a Poisson random variable with mean λ that is independent of the X_i. Show that the random sum $S_N = \sum_{i=1}^{N} X_i$ has a negative binomial distribution that counts only failures. Note that the required "number of successes" in the negative binomial need not be an integer [59].

16. Suppose X has gamma density $\lambda^\beta x^{\beta-1}e^{-\lambda x}/\Gamma(\beta)$ on $(0,\infty)$, where β is not necessarily an integer. Show that X has characteristic function $(\frac{\lambda}{\lambda - it})^\beta$ and Laplace transform $(\frac{\lambda}{\lambda+t})^\beta$. Use either of these to calculate the mean and variance of X. (Hint: For the characteristic function, derive and solve a differential equation. Alternatively, calculate the Laplace transform directly by integration and show that it can be extended to an analytic function in a certain region of the complex plane.)

17. Let X have the gamma density defined in Problem 16. Conditional on X, let Y have a Poisson distribution with mean X. Prove that Y has probability generating function

$$\mathrm{E}(s^Y) = \left(\frac{\lambda}{\lambda + 1 - s}\right)^\beta.$$

18. Show that the bilateral exponential density $\frac{1}{2}e^{-|x|}$ has characteristic function $1/(1+t^2)$. Use this fact to calculate its mean and variance.

19. Example 2.4.6 shows that it is impossible to write a random variable U uniformly distributed on $[-1,1]$ as the difference of two i.i.d. random variables X and Y. It is also true that it is impossible to write U as the sum of two i.i.d. random variables X and Y. First of all it is clear that X and Y have support on $[-1/2, 1/2]$. Hence, they possess

moments of all orders, and it is possible to represent the characteristic function of X by the series

$$\mathrm{E}\left(e^{itX}\right) = \sum_{n=0}^{\infty} \mathrm{E}(X^n)\frac{(it)^n}{n!}.$$

If one can demonstrate that the odd moments of X vanish, then it follows that its characteristic function is real and that

$$\mathrm{E}\left(e^{itU}\right) = \left[\mathrm{E}\left(e^{itX}\right)\right]^2$$

can never be negative. This contradicts the fact that $t^{-1}\sin t$ oscillates in sign. Supply the missing steps in this argument. (Hints: Why does $\mathrm{E}(X) = 0$? Assuming this is true, take n odd, expand $\mathrm{E}[(X+Y)^n]$ by the binomial theorem, and apply induction.)

20. Calculate the Laplace transform of the probability density

$$\frac{1+a^2}{a^2}e^{-x}[1-\cos(ax)]1_{\{x\geq 0\}}.$$

21. Card matching is one way of testing extrasensory perception (ESP). The tester shuffles a deck of cards labeled 1 through n and turns cards up one by one. The subject is asked to guess the value of each card and is told whenever he or she gets a match. No information is revealed for a nonmatch. According to Persi Diaconis, the optimal strategy the subject can adopt is to guess the value 1 until it turns up, then guess the value 2 until it turns up, then guess the value 3 until it turns up, and so forth. Note that this strategy gives a single match if card 2 is turned up before card 1. Show that the first two moments of the number of matches X are

$$\begin{aligned}\mathrm{E}(X) &= \sum_{j=1}^{n}\frac{1}{j!} \approx e-1\\ \mathrm{E}(X^2) &= 2\sum_{j=1}^{n}\frac{1}{(j-1)!} - \mathrm{E}(X) \approx e+1.\end{aligned}$$

(Hint: Why does the tail probability $\Pr(X \geq j)$ equal $\frac{1}{j!}$?)

22. Suppose the right-tail probability of a nonnegative random variable X satisfies $|1-F(x)| \leq cx^{-n-\epsilon}$ for all sufficiently large x, where n is a positive integer, and ϵ and c are positive real numbers. Show that $\mathrm{E}(X^n)$ is finite.

23. Let the positive random variable X have Laplace transform $L(t)$. Prove that $\mathrm{E}[(aX+b)^{-1}] = \int_0^\infty e^{-bt}L(at)\,dt$ for $a \geq 0$ and $b > 0$.

24. Let X be a nonnegative integer-valued random variable with probability generating function $G(u)$. Prove that

$$\begin{aligned} & \mathrm{E}\left[\frac{1}{(X+k+j)(X+k+j-1)\cdots(X+k)}\right] \\ = \; & \frac{1}{j!}\int_0^1 u^{k-1}(1-u)^j G(u)\,du \end{aligned}$$

by taking the expectation of $\int_0^1 u^{X+k-1}(1-u)^j\,du$.

25. Suppose X has a binomial distribution with success probability p over n trials. Show that

$$\mathrm{E}\left(\frac{1}{X+1}\right) = \frac{1-(1-p)^{n+1}}{(n+1)p}.$$

26. Let χ_n^2 and χ_{n+2}^2 be chi-square random variables with n and $n+2$ degrees of freedom, respectively. Demonstrate that

$$\mathrm{E}[f(\chi_n^2)] = n\,\mathrm{E}\left[\frac{f(\chi_{n+2}^2)}{\chi_{n+2}^2}\right]$$

for any well-behaved function $f(x)$ for which the two expectations exist. Use this identity to calculate the mean and variance of χ_n^2 [34].

27. Suppose X has a binomial distribution with success probability p over n trials. Show that

$$\mathrm{E}[Xf(X)] = \frac{p}{1-p}\,\mathrm{E}\Big[(n-X)f(X+1)\Big]$$

for any function $f(x)$. Use this identity to calculate the mean and variance of X.

28. Consider a negative binomial random variable X with density

$$\Pr(X=k) = \binom{k-1}{n-1}p^n q^{k-n}$$

for $q = 1-p$ and $k \geq n$. Prove that for any function $f(x)$

$$\mathrm{E}[qf(X)] = \mathrm{E}\left[\frac{(X-n)f(X-1)}{X-1}\right].$$

Use this identity to calculate the mean of X [100].

29. Demonstrate that the unit ball $\{x \in \mathsf{R}^n : \|x\| \leq 1\}$ has volume $\pi^{n/2}/\Gamma(n/2+1)$ and the standard simplex $\{x \in \mathsf{R}^n_+ : \|x\|_1 \leq 1\}$ has volume $1/n!$.

30. An Epanechnikov random vector X has density

$$f(x) = \frac{n+2}{2v_n}\left(1 - \|x\|^2\right)$$

supported on the unit ball $\{x \in \mathsf{R}^n : \|x\| \leq 1\}$. Here v_n is the volume of the ball as given in Problem 29. Demonstrate that X has mean vector $\mathbf{0}$ and variance matrix $\frac{1}{n+4}I$, where I is the identity matrix. (Hint: Use equation (2.13) and Example 2.8.2.)

31. Suppose the random vector X is uniformly distributed on the unit simplex T_{n-1}. Let m be a positive integer with $m < n$ and v be a vector in R^n with positive components. Show that the expected value of $(v^t X)^{-m}$ is

$$\mathrm{E}[(v^t X)^{-m}] = m\binom{n-1}{m}\int_0^\infty t^{m-1}\prod_{i=1}^n \frac{1}{tv_i+1}dt.$$

See the article [158] for explicit evaluation of the last one-dimensional integral. (Hints: Show that

$$\mathrm{E}[(v^t X)^{-m}] = \frac{\Gamma(n)}{\Gamma(n-m)}\int_{\mathsf{R}^n_+}(v^t x)^{-m}\|x\|_1^m \frac{1}{\|x\|_1^m}e^{-\|x\|_1}dx.$$

Then let $s_{n-1} = \sum_{i=2}^n v_i x_i$, and demonstrate that

$$\int_0^\infty \frac{1}{(v_1 x_1 + s_{n-1})^m}e^{-x_1}dx_1 = \frac{1}{\Gamma(m)}\int_0^\infty \frac{t^{m-1}}{tv_1+1}e^{-ts_{n-1}}dt$$

by invoking equation (2.10).)

32. One can generate the Dirichlet distribution by a different mechanism than the one developed in the text [114]. Take n independent gamma random variables $X_1, \ldots, X_n$ of unit scale and form the ratios

$$Y_i = \frac{X_i}{\sum_{j=1}^n X_j}.$$

Here X_i has density $x_i^{k_i-1}e^{-x_i}/\Gamma(k_i)$ on $(0,\infty)$ for some $k_i > 0$. Clearly, each $Y_i \geq 0$ and $\sum_{i=1}^n Y_i = 1$. Show that $(Y_1, \ldots, Y_n)^t$ follows a Dirichlet distribution on T_{n-1}.

33. Continuing Problem 32, calculate $\mathrm{E}(Y_i)$, $\mathrm{Var}(Y_i)$, and $\mathrm{Cov}(Y_i, Y_j)$ for $i \neq j$. Also show that $(Y_1 + Y_2, Y_3, \ldots, Y_n)^t$ has a Dirichlet distribution.

34. Continuing Problem 32, prove that the random variables

$$\frac{X_1}{X_1+X_2},\ \frac{X_1+X_2}{X_1+X_2+X_3},\ \ldots,\ \frac{X_1+\cdots+X_{n-1}}{X_1+\cdots+X_n},\ X_1+\cdots+X_n$$

are independent. What distributions to these random variables follow? (Hints: Denote the random variables $Z_1,\ldots,Z_n$. Make a multidimensional change of variables to find their joint density using the identities $X_1=\prod_{i=1}^n Z_i$ and $X_j=(1-Z_{j-1})\prod_{i=j}^n Z_i$ for $j>1$.)

3

Convexity, Optimization, and Inequalities

3.1 Introduction

Convexity is one of the key concepts of mathematical analysis and has interesting consequences for optimization theory, statistical estimation, inequalities, and applied probability. Despite this fact, students seldom see convexity presented in a coherent fashion. It always seems to take a backseat to more pressing topics. The current chapter is intended as a partial remedy to this pedagogical gap.

Our emphasis will be on convex functions rather than convex sets. It is helpful to have a variety of tests to recognize such functions. We present such tests and discuss the important class of log-convex functions. A strictly convex function has at most one minimum point. This property tremendously simplifies optimization. For a few functions, we are fortunate enough to be able to find their optima explicitly. For other functions, we must iterate. Section 3.4 introduces a class of optimization algorithms that exploit convexity. These algorithms are ideal for high-dimensional problems in statistics.

The concluding section of this chapter rigorously treats several inequalities. Our inclusion of Bernstein's proof of Weierstrass's approximation theorem provides a surprising application of Chebyshev's inequality and illustrates the role of probability theory in illuminating problems outside its usual sphere of influence. The less familiar inequalities of Jensen, Schlömilch, and Hölder find numerous applications in optimization theory and functional analysis.

K. Lange, *Applied Probability,*
DOI 10.1007/978-1-4419-7165-4_3, © Springer Science+Business Media, LLC 2010

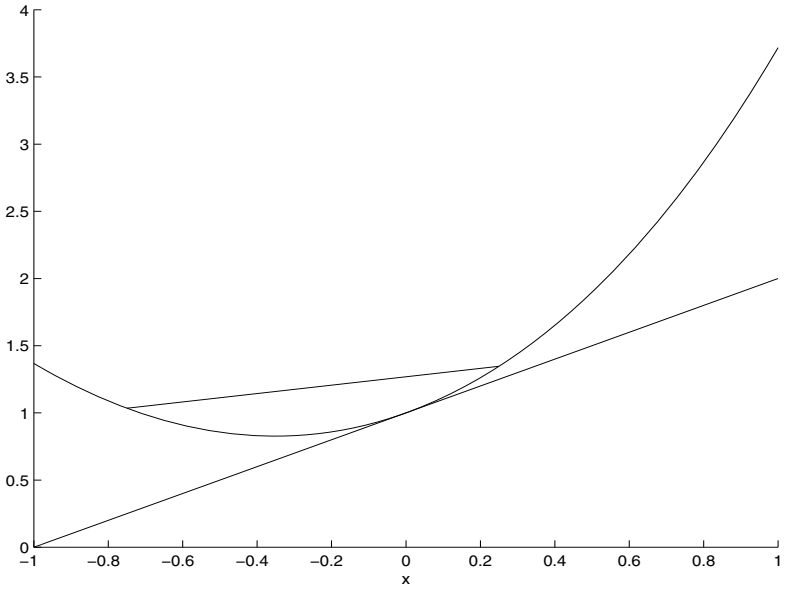


FIGURE 3.1. Plot of the Convex Function $e^x + x^2$

3.2 Convex Functions

A set $S \subset \mathsf{R}^m$ is said to be convex if the line segment between any two points x and y of S lies entirely within S. Formally, this means that whenever $x, y \in S$ and $\alpha \in [0, 1]$, the point $z = \alpha x + (1 - \alpha)y \in S$. In general, any convex combination $\sum_{i=1}^n \alpha_i x_i$ of points $x_1, \ldots, x_n$ in S must also reside in S. Here, the coefficients α_i must be nonnegative and sum to 1.

Convex functions are defined on convex sets. A real-valued function $f(x)$ defined on a convex set S is convex provided

$$f[\alpha x + (1 - \alpha)y] \quad \leq \quad \alpha f(x) + (1 - \alpha)f(y) \tag{3.1}$$

for all $x, y \in S$ and $\alpha \in [0, 1]$. Figure 3.1 depicts how in one dimension definition (3.1) requires the chord connecting any two points on the curve $x \mapsto f(x)$ to lie above the curve. If strict inequality holds in (3.1) for every $x \neq y$ and $\alpha \in (0, 1)$, then $f(x)$ is said to strictly convex. One can prove by induction that inequality (3.1) extends to

$$f\left(\sum_{i=1}^n \alpha_i x_i\right) \quad \leq \quad \sum_{i=1}^n \alpha_i f(x_i)$$

for any convex combination of points from S.

Figure 3.1 also illustrates how a tangent line to the curve lies below the curve. This property characterizes convex functions. In stating and proving the property, we will deal with differentiable functions. Recall that $f(x)$ is

differentiable at x if there exists a row vector $df(x)$ satisfying

$$f(y) \quad = \quad f(x) + df(x)(y-x) + r(y-x)$$

for all y near x and a remainder $r(y-x)$ that is $o(\|y-x\|)$ [39, 97]. Here $o(t)$ denotes a quantity such that $\lim_{t\to 0_+} o(t)t^{-1} = 0$. If the differential $df(x)$ exists, then the first partial derivatives of $f(x)$ exist at x as well, and we can identify $df(x)$ with the row vector of these partials.

Proposition 3.2.1 *Let $f(x)$ be a differentiable function on the open, convex set $S \subset \mathsf{R}^m$. Then $f(x)$ is convex if and only if*

$$f(y) \quad \geq \quad f(x) + df(x)(y-x) \tag{3.2}$$

for all $x, y \in S$. Furthermore, $f(x)$ is strictly convex if and only if strict inequality holds in inequality (3.2) when $y \neq x$.

Proof: If $f(x)$ is convex, then we can rearrange inequality (3.1) to give

$$\frac{f[\alpha x + (1-\alpha)y] - f(x)}{1-\alpha} \quad \leq \quad f(y) - f(x).$$

Letting α tend to 1 proves inequality (3.2). To demonstrate the converse, let $z = \alpha x + (1-\alpha)y$. Then with obvious notational changes, inequality (3.2) implies

$$\begin{aligned} f(x) &\geq f(z) + df(z)(x-z) \\ f(y) &\geq f(z) + df(z)(y-z). \end{aligned}$$

Multiplying the first of these inequalities by α and the second by $1-\alpha$ and adding the results produce

$$\alpha f(x) + (1-\alpha)f(y) \quad \geq \quad f(z) + df(z)(z-z) \quad = \quad f(z),$$

which is just inequality (3.1). The claims about strict convexity are left to the reader. ■

Example 3.2.1 *Linear Functions*

For a linear function $f(x) = a^t x + b$, both of the inequalities (3.1) and (3.2) are equalities. Thus, a linear function is convex. ■

Example 3.2.2 *Euclidean Norm*

The Euclidean norm $f(x) = \|x\| = \sqrt{\sum_{i=1}^m x_i^2}$ satisfies the standard triangle inequality and the homogeneity condition $\|cx\| = |c|\,\|x\|$. Thus,

$$\|\alpha x + (1-\alpha)y\| \quad \leq \quad \|\alpha x\| + \|(1-\alpha)y\| \quad \leq \quad \alpha\|x\| + (1-\alpha)\|y\|$$

for any $\alpha \in [0,1]$. It is noteworthy that $\|x\| = |x|$ is not differentiable at $x = 0$ when $m = 1$. However, closer examination of the proof of Proposition

3.2.1 makes it clear that $f(x)$ is convex throughout S if and only if the "supporting hyperplane" condition

$$f(y) \quad \geq \quad f(x) + c(x)^t(y - x)$$

holds for all y and x, where $c(x)$ is a vector depending on x. For example when $f(x) = |x|$, any scalar $c(0)$ satisfying $|c(0)| \leq 1$ works at $x = 0$. ■

It is useful to have simpler tests for convexity than inequality (3.1) or (3.2). One such test involves the second differential $d^2 f(x)$ of a function $f(x)$. We can view $d^2 f(x)$ as the Hessian matrix of second partial derivatives of $f(x)$.

Proposition 3.2.2 *Consider a twice continuously differentiable function $f(x)$ on the open, convex set $S \subset \mathsf{R}^m$. If its second differential $d^2 f(x)$ is positive semidefinite, then $f(x)$ is convex. If $d^2 f(x)$ is positive definite, then $f(x)$ is strictly convex.*

Proof: Note that $f(z)$ is a twice continuously differentiable function of the real variable α along the line $z = \alpha x + (1 - \alpha)y$. Executing a second-order Taylor expansion around $\alpha = 1$ therefore gives

$$f(y) \quad = \quad f(x) + df(x)(y - x) + \frac{1}{2}(y - x)^t d^2 f(z)(y - x)$$

for some z on the line between x and y. The claim follows directly from this equality and Proposition 3.2.1. ■

Example 3.2.3 *Arithmetic-Geometric Mean Inequality*

The second derivative test shows that the function e^x is strictly convex. Taking $y_i = e^{x_i}$, $\sum_{i=1}^n \alpha_i = 1$, and all $\alpha_i \geq 0$ produces the generalized arithmetic-geometric mean inequality

$$\prod_{i=1}^n y_i^{\alpha_i} \quad \leq \quad \sum_{i=1}^n \alpha_i y_i. \tag{3.3}$$

Equality holds if and only if all y_i coincide. ■

Example 3.2.4 *Positive Definite Quadratic Functions*

If the matrix A is positive definite, then Proposition 3.2.2 implies that the quadratic function $f(x) = \frac{1}{2}x^t Ax + b^t x + c$ is strictly convex. ■

Even Proposition 3.2.2 can be difficult to apply. The next proposition helps us to recognize convex functions by their closure properties.

Proposition 3.2.3 *Convex functions satisfy the following:*

(a) If $f(x)$ is convex and $g(x)$ is convex and increasing, then the functional composition $g \circ f(x)$ is convex.

(b) If $f(x)$ is convex, then the functional composition $f(Ax+b)$ of $f(x)$ with an affine function $Ax+b$ is convex.

(c) If $f(x)$ and $g(x)$ are convex and α and β are nonnegative constants, then $\alpha f(x)+\beta g(x)$ is convex.

(d) If $f(x)$ and $g(x)$ are convex, then $\max\{f(x), g(x)\}$ is convex.

(e) If $f_n(x)$ is a sequence of convex functions, then $\lim_{n\to\infty} f_n(x)$ is convex whenever it exists.

Proof: To prove assertion (a), we calculate

$$\begin{aligned} g \circ f[\alpha x + (1-\alpha)y] &\leq g[\alpha f(x) + (1-\alpha)f(y)] \\ &\leq \alpha g \circ f(x) + (1-\alpha) g \circ f(y). \end{aligned}$$

The remaining assertions are left to the reader. ■

Part (a) of Proposition 3.2.3 implies that $e^{f(x)}$ is convex when $f(x)$ is convex and that $f(x)^\alpha$ is convex when $f(x)$ is nonnegative and convex and $\alpha > 1$. One case not covered by the proposition is products. The counterexample $x^3 = x^2 x$ shows that the product of two convex functions is not necessarily convex.

Example 3.2.5 *Differences of Convex Functions*

Although the class of convex functions is rather narrow, most well-behaved functions can be expressed as the difference of two convex functions. For example, consider a polynomial $p(x) = \sum_{m=0}^{n} p_m x^m$. The second derivative test shows that x^m is convex whenever m is even. If m is odd, then x^m is convex on $[0, \infty)$, and $-x^m$ is convex on $(-\infty, 0)$. Therefore,

$$x^m = \max\{x^m, 0\} - \max\{-x^m, 0\}$$

is the difference of two convex functions. Because the class of differences of convex functions is closed under the formation of linear combinations, it follows that $p(x)$ belongs to this larger class. ■

A positive function $f(x)$ is said to be log-convex if and only if $\ln f(x)$ is convex. Log-convex functions have excellent closure properties as documented by the next proposition.

Proposition 3.2.4 *Log-convex functions satisfy the following:*

(a) If $f(x)$ is log-convex, then $f(x)$ is convex.

(b) If $f(x)$ is convex and $g(x)$ is log-convex and increasing, then the functional composition $g \circ f(x)$ is log-convex.

(c) If $f(x)$ is log-convex, then the functional composition $f(Ax+b)$ of $f(x)$ with an affine function $Ax+b$ is log-convex.

(d) If $f(x)$ is log-convex, then $f(x)^\alpha$ and $\alpha f(x)$ are log-convex for any $\alpha > 0$.

(e) If $f(x)$ and $g(x)$ are log-convex, then $f(x) + g(x)$, $f(x)g(x)$, and $\max\{f(x), g(x)\}$ are log-convex.

(f) If $f_n(x)$ is a sequence of log-convex functions, then $\lim_{n\to\infty} f_n(x)$ is log-convex whenever it exists and is positive.

Proof: Assertion (a) follows from part (a) of Proposition 3.2.3 after composing the functions e^x and $\ln f(x)$. To prove that the sum of log-convex functions is log-convex, let $h(x) = f(x) + g(x)$ and apply Hölder's inequality (Example 3.5.3) to random variables U and V defined on the sample space $\{0, 1\}$ with the uniform distribution. At the point 0, U equals $f(x)^\alpha$ and V equals $f(y)^{1-\alpha}$, and at the point 1, U equals $g(x)^\alpha$ and V equals $g(y)^{1-\alpha}$. These considerations imply

$$\begin{aligned} h[\alpha x + (1-\alpha)y] &= f[\alpha x + (1-\alpha)y] + g[\alpha x + (1-\alpha)y] \\ &\le f(x)^\alpha f(y)^{1-\alpha} + g(x)^\alpha g(y)^{1-\alpha} \\ &\le [f(x) + g(x)]^\alpha [f(y) + g(y)]^{1-\alpha} \\ &= h(x)^\alpha h(y)^{1-\alpha}. \end{aligned}$$

The remaining assertions are left to the reader. ■

Example 3.2.6 *Gamma Function*

Gauss's representation of the gamma function

$$\Gamma(z) = \lim_{n\to\infty} \frac{n! n^z}{z(z+1)\cdots(z+n)}$$

shows that it is log-convex on $(0, \infty)$ [95]. Indeed, one can easily check that n^z and $(z+k)^{-1}$ are log-convex and then apply the closure of the set of log-convex functions under the formation of products and limits. Note that invoking convexity in this argument is insufficient because the set of convex functions is not closed under the formation of products. Alternatively, one can deduce log-convexity from Euler's definition

$$\Gamma(z) = \int_0^\infty x^{z-1} e^{-x} dx$$

by viewing the integral as the limit of Riemann sums, each of which is log-convex. ■

Example 3.2.7 *Log-concavity of* $\det \Sigma$ *for* Σ *Positive Definite*

Let Ω be an $m \times m$ positive definite matrix. According to Section 1.7, the function

$$f(y) = \left(\frac{1}{2\pi}\right)^{m/2} |\det \Omega|^{-1/2} e^{-y^* \Omega^{-1} y/2}$$

is a probability density. Integrating over all $y \in \mathsf{R}^m$ produces the identity

$$|\det \Omega|^{1/2} = \frac{1}{(2\pi)^{m/2}} \int e^{-y^* \Omega^{-1} y/2} dy.$$

We can restate this identity in terms of the inverse matrix $\Sigma = \Omega^{-1}$ as

$$\ln \det \Sigma = m \ln(2\pi) - 2 \ln \int e^{-y^* \Sigma y/2} dy.$$

By the reasoning of the last example, the integral on the right is log-convex. Because Σ is positive definite if and only if Ω is positive definite, it follows that $\ln \det \Sigma$ is concave in the positive definite matrix Σ. ■

3.3 Minimization of Convex Functions

Optimization theory is much simpler for convex functions than for ordinary functions [89, 140, 155]. For instance, we have the following:

Proposition 3.3.1 *Suppose that $f(x)$ is a convex function on the convex set $S \subset \mathsf{R}^m$. If z is a local minimum of $f(x)$, then it is also a global minimum, and the set $\{x : f(x) = f(z)\}$ is convex.*

Proof: If $f(x) \leq f(z)$ and $f(y) \leq f(z)$, then

$$\begin{aligned} f[\alpha x + (1-\alpha)y] &\leq \alpha f(x) + (1-\alpha) f(y) \\ &\leq f(z) \end{aligned} \tag{3.4}$$

for any $\alpha \in [0, 1]$. This shows that the set $\{x : f(x) \leq f(z)\}$ is convex. Now suppose that $f(x) < f(z)$. Strict inequality then prevails between the extreme members of inequality (3.4) provided $\alpha > 0$. Taking $y = z$ and α close to 0 shows that z cannot serve as a local minimum. Thus, z must be a global minimum. ■

Example 3.3.1 *Piecewise Linear Functions*

The function $f(x) = |x|$ on the real line is piecewise linear. It attains its minimum of 0 at the point $x = 0$. The convex function $f(x) = \max\{1, |x|\}$ is also piecewise linear, but it attains its minimum throughout the interval $[-1, 1]$. In both cases the set $\{y : f(y) = \min_x f(x)\}$ is convex. In higher dimensions, the convex function $f(x) = \max\{1, \|x\|\}$ attains its minimum of 1 throughout the closed ball $\|x\| \leq 1$. ■

Proposition 3.3.2 *Let $f(x)$ be a convex, differentiable function on the convex set $S \subset \mathsf{R}^m$. If the point $z \in S$ satisfies*

$$df(z)(x - z) \geq 0$$

for every point $x \in S$, then z is a global minimum of $f(x)$. In particular, any stationary point of $f(x)$ is a global minimum.

Proof: This assertion follows immediately from inequality (3.2) characterizing convex functions. ■

Example 3.3.2 *Minimum of x on $[0,\infty)$.*

The convex function $f(x) = x$ has derivative $df(x) = 1$. On the convex set $[0,\infty)$, we have $df(0)(x-0) = x \geq 0$ for any $x \in [0,\infty)$. Hence, 0 provides the minimum of x. ■

Example 3.3.3 *Minimum of a Positive Definite Quadratic Function*

The quadratic function $f(x) = \frac{1}{2}x^t Ax + b^t x + c$ has differential

$$df(x) \quad = \quad x^t A + b^t$$

for A symmetric. Assuming that A is also invertible, the sole stationary point of $f(x)$ is $-A^{-1}b$. This point furnishes the minimum of $f(x)$ when A is positive definite. ■

Example 3.3.4 *Maximum Likelihood for the Multivariate Normal*

The sample mean and sample variance

$$\begin{aligned} \bar{y} &= \frac{1}{k}\sum_{j=1}^{k} y_j \\ S &= \frac{1}{k}\sum_{j=1}^{k}(y_j - \bar{y})(y_j - \bar{y})^t \end{aligned}$$

are also the maximum likelihood estimates of the theoretical mean μ and theoretical variance Ω of a random sample $y_1, \ldots, y_k$ from a multivariate normal. To prove this fact, we first note that maximizing the loglikelihood function

$$\begin{aligned} & -\frac{k}{2}\ln\det\Omega - \frac{1}{2}\sum_{j=1}^{k}(y_j-\mu)^t\Omega^{-1}(y_j-\mu) \\ = \; & -\frac{k}{2}\ln\det\Omega - \frac{k}{2}\mu^t\Omega^{-1}\mu + \left(\sum_{j=1}^{k} y_j\right)^t \Omega^{-1}\mu - \frac{1}{2}\sum_{j=1}^{k} y_j^t\Omega^{-1}y_j \\ = \; & -\frac{k}{2}\ln\det\Omega - \frac{1}{2}\mathrm{tr}\left[\Omega^{-1}\sum_{j=1}^{k}(y_j-\mu)(y_j-\mu)^t\right] \end{aligned}$$

constitutes a special case of the previous example with $A = k\Omega^{-1}$ and $b = -\Omega^{-1}\sum_{j=1}^{k} y_j$. This leads to the same estimate, $\hat{\mu} = \bar{y}$, regardless of the value of Ω.

To estimate Ω, we invoke the Cholesky decompositions $\Omega = LL^t$ and $S = MM^t$ under the assumption that both Ω and S are invertible. Given that $\Omega^{-1} = (L^{-1})^t L^{-1}$ and $\det \Omega = (\det L)^2$, the loglikelihood becomes

$$\begin{aligned} & k \ln \det L^{-1} - \frac{k}{2} \mathrm{tr}\left[(L^{-1})^t L^{-1} M M^t\right] \\ = \; & k \ln \det \left(L^{-1}M\right) - \frac{k}{2} \mathrm{tr}\left[(L^{-1}M)(L^{-1}M)^t\right] - k \ln \det M \end{aligned}$$

using the cyclic permutation property of the matrix trace function. Because products and inverses of lower triangular matrices are lower triangular, the matrix $R = L^{-1}M$ ranges over the set of lower triangular matrices with positive diagonal entries as L ranges over the same set. This permits us to reparameterize and estimate $R = (r_{ij})$ instead of L. Up to an irrelevant constant, the loglikelihood reduces to

$$k \ln \det R - \frac{k}{2} \mathrm{tr}(RR^t) \;=\; k \sum_i \ln r_{ii} - \frac{k}{2} \sum_i \sum_{j=1}^{i} r_{ij}^2.$$

Clearly, this is maximized by taking $r_{ij} = 0$ for $j \neq i$. Differentiation of the concave function $k \ln r_{ii} - \frac{k}{2} r_{ii}^2$ shows that it is maximized by taking $r_{ii} = 1$. In other words, the maximum likelihood estimator $\hat{R}$ is the identity matrix I. This implies that $\hat{L} = M$ and consequently that $\hat{\Omega} = S$. ■

3.4 The MM Algorithm

Most practical optimization problems defy exact solution. In this section we discuss a minimization method that relies heavily on convexity arguments and is particularly useful in high-dimensional problems such as image reconstruction [128]. We call this method the MM algorithm; the first M of this two-stage algorithm stands for majorize and the second M for minimize. When it is successful, the MM algorithm substitutes a simple optimization problem for a difficult optimization problem. Simplicity can be attained by (a) avoiding large matrix inversions, (b) linearizing an optimization problem, (c) separating the variables of an optimization problem, (d) dealing with equality and inequality constraints gracefully, and (e) turning a nondifferentiable problem into a smooth problem. The price we pay for simplifying the original problem is that we must iterate.

A function $g(x \mid x_n)$ is said to majorize a function $f(x)$ at x_n provided

$$\begin{aligned} f(x_n) &= g(x_n \mid x_n) \\ f(x) &\leq g(x \mid x_n) \qquad x \neq x_n. \end{aligned} \tag{3.5}$$

In other words, the surface $x \mapsto g(x \mid x_n)$ lies above the surface $x \mapsto f(x)$ and is tangent to it at the point $x = x_n$. Here x_n represents the current

iterate in a search of the surface $x \mapsto f(x)$. In the MM algorithm, we minimize the surrogate function $g(x \mid x_n)$ rather than the actual function $f(x)$. If x_{n+1} denotes the minimum of $g(x \mid x_n)$, then we can show that the MM procedure forces $f(x)$ downhill. Indeed, the inequality

$$\begin{aligned} f(x_{n+1}) &= g(x_{n+1} \mid x_n) + f(x_{n+1}) - g(x_{n+1} \mid x_n) \\ &\leq g(x_n \mid x_n) + f(x_n) - g(x_n \mid x_n) \\ &= f(x_n) \end{aligned} \tag{3.6}$$

follows directly from the fact $g(x_{n+1} \mid x_n) \leq g(x_n \mid x_n)$ and definition (3.5). The descent property (3.6) lends the MM algorithm remarkable numerical stability. When $f(x)$ is strictly convex, one can show with a few additional mild hypotheses that the iterates x_n converge to the global minimum of $f(x)$ regardless of the initial point x_0.

With obvious changes, the MM algorithm applies to maximization rather than minimization. To maximize a function $f(x)$, we minorize it by a surrogate function $g(x \mid x_n)$ and maximize $g(x \mid x_n)$ to produce the next iterate x_{n+1}. In this case, the letters MM stand for minorize/maximize rather than majorize/minimize. Chapter 6 discusses an MM algorithm for maximum likelihood estimation in transmission tomography. Here is a simpler example relevant to sports.

Example 3.4.1 *Bradley-Terry Model of Ranking*

In the sports version of the Bradley and Terry model [28, 109], each team i in a league of teams is assigned a rank parameter $r_i > 0$. Assuming ties are impossible, team i beats team j with probability $r_i/(r_i + r_j)$. If this outcome occurs y_{ij} times during a season of play, then the probability of the whole season is

$$L(r) = \prod_{i,j} \Big(\frac{r_i}{r_i + r_j}\Big)^{y_{ij}},$$

assuming the games are independent. To rank the teams, we find the values $\hat{r}_i$ that maximize $f(r) = \ln L(r)$. The team with the largest $\hat{r}_i$ is considered best, the team with the smallest $\hat{r}_i$ is considered worst, and so forth. In view of the fact that $-\ln u$ is convex, inequality (3.2) implies

$$\begin{aligned} f(r) &= \sum_{i,j} y_{ij} \Big[\ln r_i - \ln(r_i + r_j)\Big] \\ &\geq \sum_{i,j} y_{ij} \left[\ln r_i - \ln(r_i^n + r_j^n) - \frac{r_i + r_j - r_i^n - r_j^n}{r_i^n + r_j^n}\right] \\ &= g(r \mid r^n), \end{aligned}$$

where the superscript n indicates iteration number. Equality occurs in this minorizing inequality when $r = r^n$. Differentiating $g(r \mid r^n)$ with respect

to r_i and setting the result equal to 0 produces the next iterate

$$r_i^{n+1} = \frac{\sum_{j \neq i} y_{ij}}{\sum_{j \neq i}(y_{ij} + y_{ji})/(r_i^n + r_j^n)}.$$

Because $L(r) = L(cr)$ for any $c > 0$, we constrain $r_1 = 1$ and omit the update r_1^{n+1}. In this example, the MM algorithm separates parameters and allows us to maximize $g(r \mid r^n)$ parameter by parameter. The values $\hat{r}_i$ are referred to as maximum likelihood estimates. ■

Example 3.4.2 *Least Absolute Deviation Regression*

Statisticians often estimate parameters by the method of least squares. This classical method suffers from the fact that it is strongly influenced by observations far removed from their predicted values. To review the situation, consider p independent experiments with outcomes $y_1, \ldots, y_p$. We wish to predict y_i from q covariates $x_{i1}, \ldots, x_{iq}$ known in advance. For instance, y_i might be the height of the ith child in a classroom of p children. Relevant covariates might be the heights x_{i1} and x_{i2} of i's mother and father and the sex of i coded as $x_{i3} = 1$ for a girl and $x_{i4} = 1$ for a boy. Here we take $q = 4$ and force $x_{i3}x_{i4} = 0$ so that only one sex is possible. If we use a linear predictor $\sum_{j=1}^q x_{ij}\theta_j$ of y_i, it is natural to estimate the regression coefficients θ_j by minimizing the sum of squares

$$f(\theta) = \sum_{i=1}^p \left(y_i - \sum_{j=1}^q x_{ij}\theta_j \right)^2.$$

Differentiating $f(\theta)$ with respect to θ_j and setting the result equal to 0 produce

$$\sum_{i=1}^p x_{ij} y_i = \sum_{i=1}^p \sum_{k=1}^q x_{ij} x_{ik} \theta_k.$$

If we let y denote the column vector with entries y_i and X denote the matrix with entry x_{ij} in row i and column j, these q normal equations can be written in vector form as

$$X^t y = X^t X \theta$$

and solved as

$$\hat{\theta} = (X^t X)^{-1} X^t y.$$

In the method of least absolute deviation regression, we replace $f(\theta)$ by

$$h(\theta) = \sum_{i=1}^p \left| y_i - \sum_{j=1}^q x_{ij}\theta_j \right|.$$

Traditionally, one simplifies this expression by defining the residual

$$r_i(\theta) = y_i - \sum_{j=1}^{q} x_{ij}\theta_j.$$

We are now faced with minimizing a nondifferentiable function. Fortunately, the MM algorithm can be implemented by exploiting the convexity of the function $-\sqrt{u}$ in inequality (3.2). Because

$$-\sqrt{u} \geq -\sqrt{u^n} - \frac{u - u^n}{2\sqrt{u^n}},$$

we find that

$$\begin{aligned} h(\theta) &= \sum_{i=1}^{p} \sqrt{r_i(\theta)^2} \\ &\leq h(\theta^n) + \frac{1}{2}\sum_{i=1}^{p} \frac{r_i^2(\theta) - r_i^2(\theta^n)}{\sqrt{r_i^2(\theta^n)}} \\ &= g(\theta \mid \theta^n). \end{aligned}$$

Minimizing $g(\theta \mid \theta^n)$ is accomplished by minimizing the weighted sum of squares

$$\sum_{i=1}^{p} w_i(\theta^n) r_i(\theta)^2$$

with ith weight $w_i(\theta^n) = |r_i(\theta^n)|^{-1}$. A slight variation of the above argument for minimizing a sum of squares leads to

$$\theta^{n+1} = [X^t W(\theta^n) X]^{-1} X^t W(\theta^n) y,$$

where $W(\theta^n)$ is the diagonal matrix with ith diagonal entry $w_i(\theta^n)$. Unfortunately, the possibility that some $w_i(\theta^n) = \infty$ cannot be ruled out. Problem 13 suggests a simple remedy. ■

3.5 Moment Inequalities

Inequalities give important information about the magnitude of probabilities and expectations without requiring their exact calculation. The Cauchy-Schwarz inequality $|\mathrm{E}(XY)| \leq \mathrm{E}(X^2)^{1/2}\,\mathrm{E}(Y^2)^{1/2}$ is one of the most useful of the classical inequalities. It is also one of the easiest to remember because it is equivalent to the fact that a correlation coefficient must lie on the interval $[-1, 1]$. Equality occurs in the Cauchy-Schwarz inequality if and only if X is proportional to Y or vice versa.

Markov's inequality is another widely applied bound. Let $g(x)$ be a nonnegative, increasing function, and let X be a random variable such that $g(X)$ has finite expectation. Then Markov's inequality

$$\Pr(X \geq c) \quad \leq \quad \frac{\mathrm{E}[g(X)]}{g(c)}$$

holds for any constant c for which $g(c) > 0$. This result follows upon taking expectations in the inequality $g(c)1_{\{X \geq c\}} \leq g(X)$. Chebyshev's inequality is the special case of Markov's inequality with $g(x) = x^2$ applied to the random variable $X - \mathrm{E}(X)$. In large deviation theory, we take $g(x) = e^{tx}$ and $c > 0$ and choose $t > 0$ to minimize the right-hand side of the inequality $\Pr(X \geq c) \leq e^{-ct}\,\mathrm{E}(e^{tX})$ involving the moment generating function of X. As an example, suppose X follows a standard normal distribution. The moment generating function $e^{t^2/2}$ of X is derived by a minor variation of the argument given in Example 2.4.1 for the characteristic function of X. The large deviation inequality

$$\Pr[X \geq c] \quad \leq \quad \inf_t e^{-ct} e^{t^2/2} \quad = \quad e^{-c^2/2}$$

is called a Chernoff bound. Problem 19 discusses another typical Chernoff bound.

Our next example involves a nontrivial application of Chebyshev's inequality.

Example 3.5.1 *Weierstrass's Approximation Theorem*

Weierstrass showed that a continuous function $f(x)$ on $[0, 1]$ can be uniformly approximated to any desired degree of accuracy by a polynomial. Bernstein's lovely proof of this fact relies on applying Chebyshev's inequality to a binomial random variable S_n with n trials and success probability x per trial [60]. The corresponding candidate polynomial is defined by the expectation

$$\mathrm{E}\left[f\Big(\frac{S_n}{n}\Big)\right] \quad = \quad \sum_{k=0}^{n} f\Big(\frac{k}{n}\Big)\binom{n}{k} x^k (1-x)^{n-k}.$$

Note that $\mathrm{E}(S_n/n) = x$ and $\mathrm{Var}(S_n/n) = x(1-x)/n \leq 1/(4n)$. Now given an arbitrary $\epsilon > 0$, one can find by the uniform continuity of $f(x)$ a $\delta > 0$ such that $|f(u) - f(v)| < \epsilon$ whenever $|u - v| < \delta$. If $||f||_\infty = \sup |f(x)|$ on $[0, 1]$, then Chebyshev's inequality implies

$$\begin{aligned}
& \left|\mathrm{E}\left[f\Big(\frac{S_n}{n}\Big)\right] - f(x)\right| \\
\leq \quad & \mathrm{E}\left[\left|f\Big(\frac{S_n}{n}\Big) - f(x)\right|\right] \\
\leq \quad & \epsilon \Pr\left(\left|\frac{S_n}{n} - x\right| < \delta\right) + 2||f||_\infty \Pr\left(\left|\frac{S_n}{n} - x\right| \geq \delta\right)
\end{aligned}$$

$$\begin{aligned} &\leq \epsilon + \frac{2||f||_\infty x(1-x)}{n\delta^2} \\ &\leq \epsilon + \frac{||f||_\infty}{2n\delta^2}. \end{aligned}$$

Taking $n \geq ||f||_\infty/(2\epsilon\delta^2)$ then gives $\left|\mathrm{E}\left[f\left(\frac{S_n}{n}\right)\right] - f(x)\right| \leq 2\epsilon$ regardless of the chosen $x \in [0,1]$. ■

Proposition 3.5.1 (Jensen's Inequality) *Let the values of the random variable W be confined to the possibly infinite interval (a,b). If $h(w)$ is convex on (a,b), then $\mathrm{E}[h(W)] \geq h[\mathrm{E}(W)]$, provided both expectations exist. For a strictly convex function $h(w)$, equality holds in Jensen's inequality if and only if $W = \mathrm{E}(W)$ almost surely.*

Proof: For the sake of simplicity, suppose that $h(w)$ is differentiable. If we let $v = \mathrm{E}(W)$, then Jensen's inequality follows from Proposition 3.2.1 after taking expectations in the inequality

$$h(W) \geq h(v) + dh(v)(W - v). \tag{3.7}$$

If $h(w)$ is strictly convex, and W is not constant, then inequality (3.7) is strict with positive probability. Hence, strict inequality prevails in Jensen's inequality. ■

Jensen's inequality is the key to a host of other inequalities. Here are two nontrivial examples.

Example 3.5.2 *Schlömilch's Inequality for Weighted Means*

If X is a positive random variable, then we define the weighted mean function $\mathrm{M}(p) = \mathrm{E}(X^p)^{\frac{1}{p}}$. For the sake of argument, we assume that $\mathrm{M}(p)$ exists and is finite for all real p. Typical values of $\mathrm{M}(p)$ are $\mathrm{M}(1) = \mathrm{E}(X)$ and $\mathrm{M}(-1) = 1/\mathrm{E}(X^{-1})$. To make $\mathrm{M}(p)$ continuous at $p = 0$, it turns out that we should set $\mathrm{M}(0) = e^{\mathrm{E}(\ln X)}$. The reader is asked to check this fact in Problem 24. Here we are more concerned with proving Schlömilch's assertion that $\mathrm{M}(p)$ is an increasing function of p. If $0 < p < q$, then the function $x \mapsto x^{q/p}$ is convex, and Jensen's inequality says

$$\mathrm{E}(X^p)^{q/p} \leq \mathrm{E}(X^q).$$

Taking the qth root of both sides of this inequality yields $\mathrm{M}(p) \leq \mathrm{M}(q)$. On the other hand if $p < q < 0$, then the function $x \mapsto x^{q/p}$ is concave, and Jensen's inequality says

$$\mathrm{E}(X^p)^{q/p} \geq \mathrm{E}(X^q).$$

Taking the qth root reverses the inequality and again yields $\mathrm{M}(p) \leq \mathrm{M}(q)$. When either p or q is 0, we have to change tactics. One approach is to

invoke the continuity of $\mathrm{M}(p)$ at $p = 0$. Another approach is to exploit the concavity of $\ln x$. Jensen's inequality now gives

$$\mathrm{E}(\ln X^p) \quad \leq \quad \ln \mathrm{E}(X^p),$$

which on exponentiation becomes

$$e^{p\,\mathrm{E}(\ln X)} \quad \leq \quad \mathrm{E}(X^p).$$

If $p > 0$, then taking the pth root produces

$$M(0) \quad = \quad e^{\mathrm{E}(\ln X)} \quad \leq \quad \mathrm{E}(X^p)^{\frac{1}{p}},$$

and if $p < 0$, then taking the pth root produces the opposite inequality

$$M(0) \quad = \quad e^{\mathrm{E}(\ln X)} \quad \geq \quad \mathrm{E}(X^p)^{\frac{1}{p}}.$$

When the random variable X is defined on the space $\{1, \ldots, n\}$ equipped with the uniform distribution, Schlömilch's inequalities for $p = -1$, 0, and 1 reduce to the classical inequalities

$$\frac{1}{\frac{1}{n}\left(\frac{1}{x_1} + \cdots \frac{1}{x_n}\right)} \quad \leq \quad \Big(x_1 \cdots x_n\Big)^{\frac{1}{n}} \quad \leq \quad \frac{1}{n}\Big(x_1 + \cdots + x_n\Big)$$

relating the harmonic, geometric, and arithmetic means. ■

Example 3.5.3 *Hölder's Inequality*

Consider two random variables X and Y and two numbers $p > 1$ and $q > 1$ such that $p^{-1} + q^{-1} = 1$. Then Hölder's inequality

$$|\,\mathrm{E}(XY)| \quad \leq \quad \mathrm{E}(|X|^p)^{\frac{1}{p}}\,\mathrm{E}(|Y|^q)^{\frac{1}{q}} \tag{3.8}$$

generalizes the Cauchy-Schwarz inequality whenever the indicated expectations on its right exist. To prove (3.8), it clearly suffices to assume that X and Y are nonnegative. It also suffices to take $\mathrm{E}(X^p) = \mathrm{E}(Y^q) = 1$ once we divide the left-hand side of (3.8) by its right-hand side. Now set $r = p^{-1}$, and let Z be a random variable equal to $u \geq 0$ with probability r and equal to $v \geq 0$ with probability $1 - r$. Schlömilch's inequality $\mathrm{M}(0) \leq \mathrm{M}(1)$ for Z says

$$u^r v^{1-r} \quad \leq \quad ru + (1 - r)v.$$

If we substitute X^p for u and Y^q for v in this inequality and take expectations, then we find that $\mathrm{E}(XY) \leq r + 1 - r = 1$ as required. ■

3.6 Problems

1. On which intervals are the following functions convex: e^x, e^{-x}, x^n, $|x|^p$ for $p \geq 1$, $\sqrt{1+x^2}$, $x \ln x$, and $\cosh x$? On these intervals, which functions are log-convex?

2. Show that Riemann's zeta function
$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$$
is log-convex for $s > 1$.

3. Demonstrate that the function $f(x) = x^n - na \ln x$ is convex on $(0, \infty)$ for $a > 0$. Where does its minimum occur?

4. Prove the strict convexity assertions of Proposition 3.2.1.

5. Prove parts (b), (c), and (d) of Proposition 3.2.3.

6. Prove the unproved assertions of Proposition 3.2.4.

7. Suppose that $f(x)$ is a convex function on the real line. If a and y are vectors in R^m, then show that $f(a^t y)$ is a convex function of y. For $m > 1$ show that $f(a^t y)$ is not strictly convex.

8. Let $f(x)$ be a continuous function on the real line satisfying
$$f\left[\frac{1}{2}(x+y)\right] \leq \frac{1}{2}f(x) + \frac{1}{2}f(y).$$
Prove that $f(x)$ is convex.

9. If $f(x)$ is a nondecreasing function on the interval $[a, b]$, then show that $g(x) = \int_a^x f(y)\, dy$ is a convex function on $[a, b]$.

10. Heron's classical formula for the area of a triangle with sides of length a, b, and c is $\sqrt{s(s-a)(s-b)(s-c)}$, where $s = (a+b+c)/2$ is the semiperimeter. Using inequality (3.8), show that the triangle of fixed perimeter with greatest area is equilateral.

11. Let $H_n = 1 + \frac{1}{2} + \cdots + \frac{1}{n}$. Using inequality (3.8), verify the inequality $n\sqrt[n]{n+1} \leq n + H_n$ for any positive integer n (Putnam Competition, 1975).

12. Show that the loglikelihood $L(r)$ in Example 3.4.1 is concave under the reparameterization $r_i = e^{\theta_i}$.

13. Suppose that in Example 3.4.2 we minimize the function

$$h_\epsilon(\theta) = \sum_{i=1}^{p} \left\{ \left[y_i - \sum_{j=1}^{q} x_{ij}\theta_j \right]^2 + \epsilon \right\}^{1/2}$$

instead of $h(\theta)$ for a small, positive number ϵ. Show that the same MM algorithm applies with revised weights $w_i(\theta^n) = 1/\sqrt{r_i(\theta^n)^2 + \epsilon}$.

14. Let $X_1, \ldots, X_n$ be n independent random variables from a common distributional family. Suppose the variance $\sigma^2(\mu)$ of a generic member of this family is a function of the mean μ. Now consider the sum $S = X_1 + \cdots + X_n$. If the mean $\omega = \mathrm{E}(S)$ is fixed, it is of some interest to determine whether taking $\mathrm{E}(X_i) = \mu_i = \omega/n$ minimizes or maximizes $\mathrm{Var}(S)$. Show that the minimum occurs when $\sigma^2(\mu)$ is convex in μ and the maximum occurs when $\sigma^2(\mu)$ is concave in μ [148]. What do you deduce in the special cases where the family is binomial, Poisson, and exponential?

15. Suppose the random variables X and Y have densities $f(u)$ and $g(u)$ such that $f(u) \geq g(u)$ for $u \leq a$ and $f(u) \leq g(u)$ for $u > a$. Prove that $\mathrm{E}(X) \leq \mathrm{E}(Y)$. If in addition $f(u) = g(u) = 0$ for $u < 0$, show that $\mathrm{E}(X^n) \leq \mathrm{E}(Y^n)$ for all positive integers n [60].

16. If the random variable X has values in the interval $[a, b]$, then show that $\mathrm{Var}(X) \leq (b-a)^2/4$ and that this bound is sharp. (Hints: Reduce to the case $[a, b] = [0, 1]$. If $\mathrm{E}(X) = p$, then demonstrate that $\mathrm{Var}(X) \leq p(1-p)$.)

17. Let X be a random variable with $\mathrm{E}(X) = 0$ and $\mathrm{E}(X^2) = \sigma^2$. Show that

$$\Pr(X \geq c) \leq \frac{a^2 + \sigma^2}{(a+c)^2} \tag{3.9}$$

for all nonnegative a and c. Prove that the choice $a = \sigma^2/c$ minimizes the right-hand side of (3.9) and that for this choice

$$\Pr(X \geq c) \leq \frac{\sigma^2}{\sigma^2 + c^2}.$$

This is Cantelli's inequality [60].

18. Suppose $g(x)$ is a function such that $g(x) \leq 1$ for all x and $g(x) \leq 0$ for $x \leq c$. Demonstrate the inequality

$$\Pr(X \geq c) \geq \mathrm{E}[g(X)] \tag{3.10}$$

for any random variable X [60]. Verify that the polynomial

$$g(x) = \frac{(x-c)(c+2d-x)}{d^2}$$

with $d > 0$ satisfies the stated conditions leading to inequality (3.10). If X is nonnegative with $\mathrm{E}(X) = 1$ and $\mathrm{E}(X^2) = \beta$ and $c \in (0,1)$, then prove that the choice $d = \beta/(1-c)$ yields

$$\Pr(X \geq c) \geq \frac{(1-c)^2}{\beta}.$$

Finally, if $\mathrm{E}(X^2) = 1$ and $\mathrm{E}(X^4) = \beta$, show that

$$\Pr(|X| \geq c) \geq \frac{(1-c^2)^2}{\beta}.$$

19. Let X be a Poisson random variable with mean λ. Demonstrate that the Chernoff bound

$$\Pr(X \geq c) \leq \inf_{t>0} e^{-ct} \mathrm{E}(e^{tX})$$

amounts to

$$\Pr(X \geq c) \leq \frac{(\lambda e)^c}{c^c} e^{-\lambda}$$

for any integer $c > \lambda$.

20. Let $B_n f(x) = \mathrm{E}[f(S_n/n)]$ denote the Bernstein polynomial of degree n approximating $f(x)$ as discussed in Example 3.5.1. Prove that

    (a) $B_n f(x)$ is linear in $f(x)$,
    (b) $B_n f(x) \geq 0$ if $f(x) \geq 0$,
    (c) $B_n f(x) = f(x)$ if $f(x)$ is linear,
    (d) $B_n x(1-x) = \frac{n-1}{n} x(1-x)$.

21. Suppose the function $f(x)$ has continuous derivative $f'(x)$. For $\delta > 0$ show that Bernstein's polynomial satisfies the bound

$$\left| \mathrm{E}\left[f\left(\frac{S_n}{n}\right)\right] - f(x) \right| \leq \delta ||f'||_\infty + \frac{||f||_\infty}{2n\delta^2}.$$

Conclude from this estimate that

$$\left|\left| \mathrm{E}\left[f\left(\frac{S_n}{n}\right)\right] - f \right|\right|_\infty = O(n^{-\frac{1}{3}}).$$

22. Let $f(x)$ be a convex function on $[0,1]$. Prove that the Bernstein polynomial of degree n approximating $f(x)$ is also convex. (Hint: Show that

$$\begin{aligned}\frac{d^2}{dx^2}\,\mathrm{E}\left[f\Big(\frac{S_n}{n}\Big)\right] &= n(n-1)\Big\{\,\mathrm{E}\left[f\Big(\frac{S_{n-2}+2}{n}\Big)\right] \\ &\quad -2\;\mathrm{E}\left[f\Big(\frac{S_{n-2}+1}{n}\Big)\right]+\mathrm{E}\left[f\Big(\frac{S_{n-2}}{n}\Big)\right]\Big\}\end{aligned}$$

in the notation of Example 3.5.1.)

23. Suppose $1 \le p < \infty$. For a random variable X with $\mathrm{E}(|X|^p) < \infty$, define the norm $||X||_p = \mathrm{E}(X^p)^{\frac{1}{p}}$. Now prove Minkowski's triangle inequality $||X+Y||_p \le ||X||_p + ||Y||_p$. (Hint: Apply Hölder's inequality to the right-hand side of

$$\mathrm{E}(|X+Y|^p) \le \mathrm{E}(|X|\cdot|X+Y|^{p-1}) + \mathrm{E}(|Y|\cdot|X+Y|^{p-1})$$

and rearrange the result.

24. Suppose X is a random variable satisfying $0 < a \le X \le b < \infty$. Use l'Hôpital's rule to prove that the weighted mean $\mathrm{M}(p) = \mathrm{E}(X^p)^{\frac{1}{p}}$ is continuous at $p = 0$ if we define $\mathrm{M}(0) = e^{\mathrm{E}(\ln X)}$.

4
Combinatorics

4.1 Introduction

Combinatorics is the bane of many a student of probability theory. Even elementary combinatorial problems can be frustratingly subtle. The cure for this ill is more exposure, not less. Because combinatorics has so many important applications, serious students of the mathematical sciences neglect it at their peril. Here we explore a few topics in combinatorics that have maximum intersection with probability. Our policy is to assume that readers have a nodding familiarity with combinations and permutations. Based on this background, we discuss bijections, inclusion-exclusion (sieve) methods, Catalan numbers, Stirling numbers of the first and second kind, and the pigeonhole principle. Along the way we meet some applications that we hope will whet readers' appetites for further study. The books [21, 22, 26, 59, 78, 139, 207] are especially recommended.

4.2 Bijections

Many combinatorial identities can be derived by posing and answering the same counting problem in two different ways. The bijection method consists in constructing a one-to-one correspondence between two sets, both of which we can count. The correspondence shows that the sets have the same cardinality. This idea is more fertile than it sounds. For instance, it

K. Lange, *Applied Probability,*
DOI 10.1007/978-1-4419-7165-4_4, © Springer Science+Business Media, LLC 2010

forms the basis of many recurrence relations for computing combinatorial quantities.

Example 4.2.1 *Pascal's Triangle*

Let $\binom{n}{k}$ be the number of subsets of size k from a set of size n. Pascal's triangle is the recurrence scheme specified by

$$\binom{n+1}{k} = \binom{n}{k-1} + \binom{n}{k} \tag{4.1}$$

together with the boundary conditions $\binom{n}{0} = \binom{n}{n} = 1$. To derive equation (4.1) we take a set of size $n+1$ and divide it into a set of size n and a set of size 1. We can either choose $k-1$ elements from the n-set and combine them with the single element from the 1-set or choose all k elements from the n-set. The first choice can be made in $\binom{n}{k-1}$ ways and the second in $\binom{n}{k}$ ways.

As indicated by its name, we visualize Pascal's triangle as an infinite lower triangular matrix with n as row index and k as column index. The boundary values specify the first column and the diagonal as the constant 1. The recurrence proceeds row by row. If one desires only the binomial coefficients for a single final row, it is advantageous in coding Pascal's triangle to proceed from right to left along the current row. This minimizes computer storage by making it possible to overwrite safely the contents of the previous row with the contents of the current row. Pascal's triangle also avoids the danger of computer overflows caused by computing binomial coefficients via factorials. ■

Example 4.2.2 *Even and Odd Parity Subsets*

There are a host of identities connecting binomial coefficients. Here is one

$$\sum_{i=0}^{j}(-1)^i\binom{k}{i} = (-1)^j\binom{k-1}{j} \tag{4.2}$$

that has an interesting relationship to subset parity. Let e_j (o_j) denote the number of subsets of the set $S = \{1, \ldots, k\}$ with even (odd) cardinality not exceeding j. Equation (4.2) can be restated as the two equations

$$\begin{aligned} e_{2j} &= o_{2j} + \binom{k-1}{2j} \\ o_{2j+1} &= e_{2j+1} + \binom{k-1}{2j+1}. \end{aligned}$$

The bijection method provides an easy proof of these identities. Suppose T is a subset of S that contains k and has cardinality $|T| \leq j$. Then we

map T to $T \setminus \{k\}$. If T does not contain k and has cardinality $|T| < j$, then we map T to $T \cup \{k\}$. This construction maps subsets with even cardinality not exceeding j into subsets with odd cardinality not exceeding j and vice versa. The subsets not containing k with cardinality exactly j are not taken into account in this partial correspondence. There are $\binom{k-1}{j}$ such subsets, and they each have the same parity as j. The binomial coefficient $\binom{k-1}{j}$ provides the correction term to the partial correspondence. As an alternative to this combinatorial proof, the reader can verify equation (4.2) algebraically by induction on j. The right-tail identity

$$\begin{aligned} \sum_{i=j+1}^{k} (-1)^i \binom{k}{i} &= (1-1)^k - \sum_{i=0}^{j} (-1)^i \binom{k}{i} \\ &= (-1)^{j+1} \binom{k-1}{j} \end{aligned} \tag{4.3}$$

is a direct consequence of equation (4.2) and will prove useful later. ■

Example 4.2.3 *Bell Numbers and Set Partitions*

The Bell number B_n denotes the number of partitions of a set with n elements. By a partition we mean a division of the set into disjoint blocks. A partition induces an equivalence relation on the set in the sense that two elements are equivalent if and only if they belong to the same block. Two partitions are the same if and only if they induce the same equivalence relation.

Starting with $B_0 = 1$, the B_n satisfy the recurrence relation

$$B_{n+1} = \sum_{k=0}^{n} \binom{n}{k} B_{n-k} = \sum_{k=0}^{n} \binom{n}{k} B_k.$$

The reasoning leading to equation (4.4) is basically the same as in Example 4.2.1. We divide our set with $n+1$ elements into an n-set and a 1-set. The 1-set can form a block by itself, and the n-set can be partitioned in B_n ways. Or we can choose $k \geq 1$ elements from the n-set in $\binom{n}{k}$ ways and form a block consisting of these elements and the single element from the 1-set. The remaining $n-k$ elements of the n-set can be partitioned in B_{n-k} ways. ■

Example 4.2.4 *Fibonacci Numbers*

Let s_n be the number of subsets of $\{1, \ldots, n\}$ that do not contain two consecutive integers. Because the empty set is a valid subset, it is obvious that $s_1 = 2$ and $s_2 = 3$. The recurrence $s_n = s_{n-1} + s_{n-2}$ generates the remaining elements of the sequence. To verify the recurrence, consider such a subset S. If n is not a member of S, then the other elements of S can

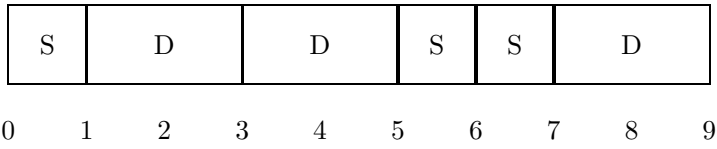


FIGURE 4.1. A Tiling of a 9-Row with 3 Square Pieces and 3 Dominoes

be chosen in s_{n-1} ways. If n is an element of S, then $n-1$ cannot be an element of S, and the other elements of S can be chosen in s_{n-2} ways.

The same recurrence relation $f_n = f_{n-1} + f_{n-2}$ generates the well-known Fibonacci sequence with initial conditions $f_1 = 1$ and $f_2 = 2$. The Fibonacci number f_n counts the number of tilings of a single row of an extended checkerboard by square pieces and dominoes. The row in question has n squares to be filled. The initial condition $f_2 = 2$ is true because a board with two squares can be filled by two consecutive square pieces or by one domino. The truth of the Fibonacci recurrence can be checked by conditioning on whether a square piece or a domino occupies square n.

The common recurrence and the shifted initial conditions $s_1 = f_2$ and $s_2 = f_3$ demonstrate that $s_n = f_{n+1}$. We can also prove this assertion by constructing a bijection between the collection of subsets of $\{1, \ldots, n\}$ with no consecutive integers and the collection of tilings of a row consisting of $n+1$ squares. A row with $n+1$ squares has $n+2$ boundaries, which we label $0, 1, \ldots, n+1$. Boundary 0 lies to the left of square 1, and boundary $n+1$ lies to the right of square $n+1$. Now consider a subset S of $\{1, \ldots, n\}$ with no consecutive elements. If i is in S, we insert a domino in the row so that it straddles boundary i. We fill the remaining squares of the row with square pieces. Figure 4.1 depicts a tiling constructed from the set $\{2, 4, 8\}$ along a row with $n+1 = 9$ squares. Square pieces are labeled "S" and dominoes "D." This construction creates the desired one-to-one correspondence. ■

4.3 Inclusion-Exclusion

The simplest inclusion-exclusion formula is

$$\Pr(A_1 \cup A_2) \quad = \quad \Pr(A_1) + \Pr(A_2) - \Pr(A_1 \cap A_2). \tag{4.4}$$

The probability $\Pr(A_1 \cap A_2)$ is subtracted from the sum $\Pr(A_1) + \Pr(A_2)$ to compensate for double counting of the intersection $A_1 \cap A_2$. As pointed

out in Chapter 1, we can derive formula (4.4) by taking expectations of indicator random variables. In general, let $1_{A_1}, \ldots, 1_{A_n}$ be the indicators of n events $A_1, \ldots, A_n$. We are interested in the probability $p_{[k]}$ that exactly k of these events occur. To track which events participate in the joint occurrence, we record the relevant event indices in a set R with cardinality k. The probability that the events indexed by R occur and no other events occur can be written as the expectation

$$\mathrm{E}\left[\prod_{i\in R} 1_{A_i} \prod_{i\in R^c} (1 - 1_{A_i})\right],$$

where R^c is the set complement $\{1, \ldots, n\} \setminus R$. Applying the distributive rule and the linearity of expectation, we find that

$$\mathrm{E}\left[\prod_{i\in R} 1_{A_i} \prod_{i\in R^c} (1 - 1_{A_i})\right] = \sum_{j=k}^{n} (-1)^{j-k} \sum_{\substack{S\supset R \\ |S|=j}} \Pr\left(\bigcap_{i\in S} A_i\right).$$

With this notation in hand, we calculate

$$\begin{aligned} p_{[k]} &= \sum_{|R|=k} \mathrm{E}\left[\prod_{i\in R} 1_{A_i} \prod_{i\in R^c} (1 - 1_{A_i})\right] \\ &= \sum_{|R|=k} \sum_{j=k}^{n} (-1)^{j-k} \sum_{\substack{S\supset R \\ |S|=j}} \Pr\left(\bigcap_{i\in S} A_i\right) \\ &= \sum_{j=k}^{n} (-1)^{j-k} \sum_{|R|=k} \sum_{\substack{S\supset R \\ |S|=j}} \Pr\left(\bigcap_{i\in S} A_i\right) \qquad (4.5) \\ &= \sum_{j=k}^{n} (-1)^{j-k} \binom{j}{k} \sum_{|S|=j} \Pr\left(\bigcap_{i\in S} A_i\right). \end{aligned}$$

The last equality in this string of equalities reflects the fact that there are $\binom{j}{k}$ subsets R of size k contained within a given set S of size j.

With a slight elaboration of this argument, we can calculate the probability $p_{(k)}$ that at least k of the events $A_1, \ldots, A_n$ occur. Indeed, this probability is

$$\begin{aligned} \sum_{l=k}^{n} p_{[l]} &= \sum_{l=k}^{n} \sum_{j=l}^{n} (-1)^{j-l} \binom{j}{l} \sum_{|S|=j} \Pr\left(\bigcap_{i\in S} A_i\right) \\ &= \sum_{j=k}^{n} \sum_{l=k}^{j} (-1)^{j-l} \binom{j}{l} \sum_{|S|=j} \Pr\left(\bigcap_{i\in S} A_i\right) \qquad (4.6) \end{aligned}$$

$$= \sum_{j=k}^{n}(-1)^{j-k}\binom{j-1}{k-1}\sum_{|S|=j}\Pr\Big(\bigcap_{i\in S}A_i\Big),$$

where the last equality in this string of equalities invokes the identity (4.3).

In many practical examples, the events $A_1, \ldots, A_n$ are exchangeable in the sense that

$$\Pr\Big(\bigcap_{i\in S}A_i\Big) = \Pr\Big(A_1\bigcap\cdots\bigcap A_j\Big)$$

for all subsets S of size j. (A similar definition holds for exchangeable random variables.) Because there are $\binom{n}{j}$ such subsets, formula (4.5) reduces to

$$p_{[k]} = \sum_{j=k}^{n}(-1)^{j-k}\binom{j}{k}\binom{n}{j}\Pr\Big(A_1\bigcap\cdots\bigcap A_j\Big) \tag{4.7}$$

in the presence of exchangeable events. Formula (4.6) for $p_{(k)}$ likewise simplifies to

$$p_{(k)} = \sum_{j=k}^{n}(-1)^{j-k}\binom{j-1}{k-1}\binom{n}{j}\Pr\Big(A_1\bigcap\cdots\bigcap A_j\Big)$$

in the presence of exchangeable events.

Example 4.3.1 *Fixed Points of a Random Permutation*

Resuming our investigation of Example 2.2.1, let us find the exact distribution of the number of fixed points X of a random permutation π. The pertinent events $A_i = \{\pi : \pi(i) = i\}$ for $1 \le i \le n$ are clearly exchangeable. Furthermore, $\Pr(A_1 \bigcap \cdots \bigcap A_j) = (n-j)!/n!$ because the movable integers $\{j+1, \ldots, n\}$ can be permuted in $(n-j)!$ ways. Hence, formula (4.7) gives

$$\begin{aligned} p_{[k]} &= \sum_{j=k}^{n}(-1)^{j-k}\binom{j}{k}\binom{n}{j}\frac{(n-j)!}{n!} \\ &= \frac{1}{k!}\sum_{j=k}^{n}(-1)^{j-k}\frac{1}{(j-k)!} \\ &= \frac{1}{k!}\sum_{i=0}^{n-k}(-1)^{i}\frac{1}{i!}. \end{aligned}$$

For $n-k$ reasonably large, the approximation $p_{[k]} \approx e^{-1}/k!$ holds, and this validates our earlier claim that X follows an approximate Poisson distribution with mean 1. ■

Example 4.3.2 *Euler's Totient Function*

Let n be a positive integer with prime factorization $n = p_1^{m_1} \cdots p_q^{m_q}$. (See Appendix A.1 and the introduction to Chapter 15 for a brief review of number theory.) For instance, if $n = 20$, then $n = 2^2 \cdot 5$. If we impose the uniform distribution on the set $\{1, \ldots, n\}$, then it makes sense to ask for the probability $p_{[0]}$ that a random integer N shares no common prime factors with n. Euler considered this problem and gave a lovely formula for his totient function $\varphi(n) = np_{[0]}$. To calculate $\varphi(n)$ via inclusion-exclusion, let A_i be the set of integers between 1 and n divisible by the ith prime p_i in the prime decomposition of n. A little reflection shows that $\Pr(A_i) = \frac{1}{p_i}$ and that in general

$$\Pr\left(\bigcap_{i \in S} A_i\right) = \prod_{i \in S} \frac{1}{p_i}.$$

Hence, equation (4.5) implies

$$\begin{aligned} \frac{\varphi(n)}{n} &= 1 - \sum_i \frac{1}{p_i} + \sum_{i<j} \frac{1}{p_i p_j} - \sum_{i<j<k} \frac{1}{p_i p_j p_k} + \cdots \\ &= \left(1 - \frac{1}{p_1}\right)\left(1 - \frac{1}{p_2}\right) \cdots \left(1 - \frac{1}{p_q}\right). \end{aligned}$$

■

Example 4.3.3 *0-1 Matrices*

Consider an $m \times n$ random matrix M with entries restricted to the values 0 and 1 [139]. Each entry is determined by flipping an unbiased coin. If the coin lands heads up, then the entry is set to 1; otherwise, it is set to 0. We now ask for the probability $p_{[0]}$ that M possesses no row or column filled entirely with 0's. This problem yields to inclusion-exclusion if we let R_i be the event that row i consists entirely of 0's and C_j be the event that column j consists entirely of 0's. When we intersect s different R_i with t different C_j, the resulting event has probability $1/2^{sn+tm-st}$. Note in this regard that s rows and t columns overlap in st entries; therefore, we must subtract st from $sn + tm$ to avoid double counting of entries. The inclusion-exclusion formula (4.5) now boils down to

$$\begin{aligned} p_{[0]} &= \sum_{s=0}^{m} \sum_{t=0}^{n} (-1)^{s+t} \binom{m}{s} \binom{n}{t} \frac{1}{2^{sn+tm-st}} \\ &= \frac{1}{2^{mn}} \sum_{s=0}^{m} (-1)^s \binom{m}{s} \sum_{t=0}^{n} \binom{n}{t} (-1)^t 2^{(m-s)(n-t)} \\ &= \frac{1}{2^{mn}} \sum_{s=0}^{m} (-1)^s \binom{m}{s} \left(2^{m-s} - 1\right)^n \end{aligned}$$

because s events R_i can be chosen in $\binom{m}{s}$ ways and t events C_j can be chosen in $\binom{n}{t}$ ways. ■

In practice, it is often cumbersome to calculate all of the terms in the alternating series (4.5). Fortunately, the partial sums

$$\sum_{j=k}^{k+m}(-1)^{j-k}\binom{j}{k}\sum_{|S|=j}\Pr\left(\bigcap_{i\in S}A_i\right)$$

overestimate $p_{[k]}$ for m even and underestimate $p_{[k]}$ for m odd. When $k=0$, the first two of these Bonferroni inequalities are

$$\Pr\left(\bigcup_{i=1}^{n}A_i\right) \;=\; 1-p_{[0]} \;\le\; \sum_{i=1}^{n}\Pr(A_i)$$

and

$$\Pr\left(\bigcup_{i=1}^{n}A_i\right) \;\ge\; \sum_{i=1}^{n}\Pr(A_i)-\sum_{i<j}\Pr\left(A_i\bigcap A_j\right). \tag{4.8}$$

In exactly the same manner, the partial sums

$$\sum_{j=k}^{k+m}(-1)^{j-k}\binom{j-1}{k-1}\sum_{|S|=j}\Pr\left(\bigcap_{i\in S}A_i\right)$$

overestimate $p_{(k)}$ for m even and underestimate $p_{(k)}$ for m odd [59, 67].

To prove these claims, let us re-examine the derivation (4.5). Suppose that all of the events A_i occur for $i\in R$ and $q>0$ of the events A_i occur for $i\in R^c$. If we truncate the expanded product

$$\prod_{i\in R}1_{A_i}\prod_{i\in R^c}(1-1_{A_i}) \;=\; 1-\binom{q}{1}+\binom{q}{2}-\binom{q}{3}+\cdots$$

after its first $m+1$ terms, identity (4.3) implies that the error committed is

$$\sum_{i=m+1}^{q}(-1)^i\binom{q}{i} \;=\; (-1)^{m+1}\binom{q-1}{m}.$$

For m even we therefore find that

$$\sum_{j=k}^{n}(-1)^{j-k}\sum_{\substack{S\supset R\\|S|=j}}\prod_{i\in S}1_{A_i} \;\le\; \sum_{j=k}^{k+m}(-1)^{j-k}\sum_{\substack{S\supset R\\|S|=j}}\prod_{i\in S}1_{A_i},$$

and for m odd that

$$\sum_{j=k}^{n}(-1)^{j-k}\sum_{\substack{S\supset R\\|S|=j}}\prod_{i\in S}1_{A_i} \;\geq\; \sum_{j=k}^{k+m}(-1)^{j-k}\sum_{\substack{S\supset R\\|S|=j}}\prod_{i\in S}1_{A_i}.$$

These inequalities are preserved by the expectation operator. If none of the events A_i with $i \in R^c$ occurs, then $\prod_{i\in R}1_{A_i}\prod_{i\in R^c}(1-1_{A_i})$ is perfectly approximated by any of its truncated expansions. The remaining steps in the derivations of $p_{[k]}$ and $p_{(k)}$ are valid provided we replace equalities by inequalities throughout.

4.4 Applications to Order Statistics

In many practical problems, it is convenient to rearrange n random variables $X_1, X_2, \ldots, X_n$ so that they appear in increasing order. Understanding the marginal distributions and moments of the resulting order statistics $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$ is difficult even when the original random variables are independent and identically distributed. Surprisingly, the principle of inclusion-exclusion sheds considerable light on the subject [9, 17, 45]. To make this claim precise, we need some notation. Denote the distribution function of $X_{(i)}$ by $F_{(i)}(t)$. For an arbitrary subset $S \subset \{1, 2, \ldots, n\}$, let $X_S = \min\{X_j : j \in S\}$ and $X^S = \max\{X_j : j \in S\}$, and designate the corresponding distribution functions $F_S(t)$ and $F^S(t)$, respectively. The following proposition is then true.

Proposition 4.4.1 *The distribution functions of the order statistics $X_{(i)}$ can be expressed as*

$$F_{(i)}(t) = \sum_{j=i}^{n}(-1)^{j-i}\binom{j-1}{i-1}\sum_{|S|=j}F^S(t) \tag{4.9}$$

$$F_{(n-i+1)}(t) = \sum_{j=i}^{n}(-1)^{j-i}\binom{j-1}{i-1}\sum_{|S|=j}F_S(t), \tag{4.10}$$

where the sum on S extends over all subsets of $\{1, 2, \ldots, n\}$ with j elements. Consequently, if each X_j possesses a kth absolute moment, then

$$\mathrm{E}\left[X_{(i)}^k\right] = \sum_{j=i}^{n}(-1)^{j-i}\binom{j-1}{i-1}\sum_{|S|=j}\mathrm{E}[(X^S)^k] \tag{4.11}$$

$$\mathrm{E}\left[X_{(n-i+1)}^k\right] = \sum_{j=i}^{n}(-1)^{j-i}\binom{j-1}{i-1}\sum_{|S|=j}\mathrm{E}[(X_S)^k]. \tag{4.12}$$

All of these formulas simplify in the obvious manner if the X_i are exchangeable random variables.

Proof: If we define the events $A_j = \{X_j \leq t\}$, then $F_{(i)}(t)$ is the probability that at least i of the n events A_j occur. Hence, equation (4.9) is just a restatement of equation (4.6). To prove equation (4.10), let $Y_j = -X_j$ and note that $Y_{(i)} = -X_{(n-i+1)}$. Now apply equation (4.6) to the events

$$A_j \;=\; \{Y_j < -t\} \;=\; \{X_j > t\}$$

and deduce that

$$\begin{aligned}
\Pr(X_{(n-i+1)} > t) &= \Pr(Y_{(i)} < -t) \qquad (4.13)\\
&= \sum_{j=i}^{n} (-1)^{j-i} \binom{j-1}{i-1} \sum_{|S|=j} \Pr\left(\max_{k\in S} -Y_k < -t\right)\\
&= \sum_{j=i}^{n} (-1)^{j-i} \binom{j-1}{i-1} \sum_{|S|=j} \Pr\left(\min_{k\in S} X_k > t\right).
\end{aligned}$$

Subtracting the extreme sides of this equation from the constant

$$1 \;=\; \sum_{j=i}^{n} (-1)^{j-i} \binom{j-1}{i-1} \sum_{|S|=j} 1 \qquad (4.14)$$

gives the final result (4.10). Note that equation (4.14) follows from equation (4.9) by sending t to ∞.

The two moment identities (4.11) and (4.12) are valid because two finite measures that share an identical distribution function also share identical moments. (Problem 12 asks the reader to check that all moments in sight exist.) Alternatively, if the X_j are nonnegative, then we can prove identity (4.12) by multiplying both sides of equality (4.13) by kt^{k-1} and integrating as discussed in Example 2.5.1. Finally, identity (4.11) is proved in similar fashion. ■

4.5 Catalan Numbers

The Catalan numbers c_n have numerous combinatorial interpretations [78, 119, 177, 207]. One of the most natural involves consistent arrangements of parentheses. Consider a string of n open parentheses and n closed parentheses. In the string each open parenthesis is paired with a closed parenthesis on its right. Thus, as one scans from left to right, the count of closed parentheses always trails or equals the corresponding count of open parentheses. By definition the number of such strings is c_n. For instance with $n = 3$, the $c_3 = 5$ possible strings are

$$()()(), \quad ()(()), \quad (())(), \quad (()()), \quad ((())).$$

Direct enumeration of legal strings soon becomes tedious. Fortunately under the convention $c_0 = 1$, the recurrence relation

$$c_{n+1} = \sum_{k=0}^{n} c_k c_{n-k} \tag{4.15}$$

enables straightforward evaluation of c_n for all n of moderate size. The rationale for equation (4.15) requires noting the position of the closed parenthesis balancing the first open parenthesis. For instance with the string (())(), balance is achieved at position 4 from the left. Stripping off the first open parenthesis and its corresponding closed parenthesis leaves a left legal substring of length $2k$ and a right legal substring of length $2(n-k)$ for some k between 0 and n. The product $c_k c_{n-k}$ counts the number of such legal pairs for a given k. Summing on k then gives the total number of legal strings of length $2n+2$.

The recurrence (4.15) yields the generating function $c(x) = \sum_{n=0}^{\infty} c_n x^n$. Indeed, if we multiply the recurrence by x^{n+1} and sum on n, then we get

$$c(x) - 1 = x \sum_{n=0}^{\infty} \sum_{k=0}^{n} c_k x^k c_{n-k} x^{n-k} = x c(x)^2.$$

This quadratic can be solved in the form

$$c(x) = \frac{1 - \sqrt{1-4x}}{2x}. \tag{4.16}$$

The other root is rejected because it has a singularity at $x = 0$ where $c(x)$ has the well-behaved value 1. Extracting the nth coefficient of the expression (4.16) by Newton's binomial formula leads to

$$c_n = -\frac{1}{2}\binom{\frac{1}{2}}{n+1}(-4)^{n+1} = \frac{1}{n+1}\binom{2n}{n}. \tag{4.17}$$

As a check on this calculation, observe that

$$\frac{1}{n+1}\binom{2n}{n} = \binom{2n}{n} - \binom{2n}{n-1}.$$

is a positive integer.

A simple random walk offers another interpretation of the Catalan numbers. Such a walk begins at 0 and moves up (+1) or down (−1) at each step with equal probability. If we identify an open parenthesis with +1 and a closed parenthesis with −1, then c_n counts the number of walks with $2n$ steps that end at 0 and remain at or above level 0 over all steps. Because there are 2^{2n} possible walks, the probability of a random walk satisfying these conditions is $c_n 2^{-2n}$.

4.6 Stirling Numbers

There are two kinds of Stirling numbers [22, 26, 78, 139, 167, 207]. Stirling numbers $\left\{ {n \atop k} \right\}$ of the second kind count the number of possible partitions of a set of n objects into k disjoint blocks. For instance, $\left\{ {3 \atop 2} \right\} = 3$ because the set $\{1, 2, 3\}$ can be partitioned into two disjoint blocks in the three ways $\{1, 2\} \cup \{3\}$, $\{1, 3\} \cup \{2\}$, and $\{2, 3\} \cup \{1\}$. The identity $B_n = \sum_k \left\{ {n \atop k} \right\}$ connects the Stirling numbers $\left\{ {n \atop k} \right\}$ to the Bell number B_n.

We can generate the numbers $\left\{ {n \atop k} \right\}$ recursively from the boundary conditions $\left\{ {n \atop 1} \right\} = 1$ and $\left\{ {n \atop k} \right\} = 0$ for $k > n$ and the recurrence relation

$$\left\{ {n \atop k} \right\} \quad = \quad \left\{ {n-1 \atop k-1} \right\} + k \left\{ {n-1 \atop k} \right\}. \tag{4.18}$$

To prove the recurrence (4.18), imagine adding n to an existing partition of $\{1, \ldots, n-1\}$. If the existing partition has $k-1$ blocks, then we must create a new block for n in order to achieve k blocks. If the existing partition has k blocks, then we must add n to one of the k existing blocks. This can be done in k ways. Since none of the other partitions of $\{1, \ldots, n-1\}$ can be successfully modified to form k blocks, formula (4.18) is true.

As an application of Stirling numbers of the second kind, consider the problem of throwing n symmetric dice with r faces each. To calculate the probability that k different faces appear when the n dice are thrown, we must first take into account the number of ways $\left\{ {n \atop k} \right\}$ that the n dice can be partitioned into k blocks. Once these blocks are chosen, then top-side faces can be assigned to the k blocks in $r(r-1)\cdots(r-k+1)$ ways. Thus, the probability in question amounts to $r^{-n} \left\{ {n \atop k} \right\} r(r-1)\cdots(r-k+1)$, which vanishes when $k > \min\{n, r\}$.

In the dice problem, the possible probabilities sum to 1. This establishes the polynomial identity

$$\sum_{k=1}^{n} \left\{ {n \atop k} \right\} x(x-1)\cdots(x-k+1) \quad = \quad x^n \tag{4.19}$$

for all positive integers $x = r$ and therefore for all real numbers x. Substituting $-x$ for x in (4.19) gives the similar identity

$$\sum_{k=1}^{n} (-1)^{n-k} \left\{ {n \atop k} \right\} x(x+1)\cdots(x+k-1) \quad = \quad x^n. \tag{4.20}$$

Finally, substituting a random variable X for x in equations (4.19) and (4.20) and taking expectations leads to the relations

$$\sum_{k=1}^{n} \left\{ {n \atop k} \right\} \mathrm{E}[X(X-1)\cdots(X-k+1)] \quad = \quad \mathrm{E}(X^n)$$

$$\sum_{k=1}^{n}(-1)^{n-k}\left\{{n \atop k}\right\}\mathrm{E}[X(X+1)\cdots(X+k-1)] \;=\; \mathrm{E}(X^n) \quad (4.21)$$

connecting falling and rising factorial moments to ordinary moments. The former relation is obviously pertinent when we calculate moments by differentiating a probability generating function.

We now turn to Stirling numbers $\left[{n \atop k}\right]$ of the first kind. These have a combinatorial interpretation in terms of the cycles of a permutation. Consider the permutation π of $\{1,\ldots,6\}$ that carries the top row of the matrix

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 6 & 5 & 4 & 1 & 2 \end{pmatrix} \quad (4.22)$$

to its bottom row. We can also represent π by the cycle decomposition $(1,3,5),(2,6),(4)$. The first cycle $(1,3,5)$ indicates that π satisfies $\pi(1)=3$, $\pi(3)=5$, and $\pi(5)=1$; the second cycle that $\pi(2)=6$ and $\pi(6)=2$; and the third cycle that $\pi(4)=4$. Note that the order of the cycles is irrelevant in representing π. Also within a cycle, only rotational order is relevant. Thus, the three cycles $(1,3,5)$, $(5,1,3)$, and $(3,5,1)$ are all equivalent. In the preferred or canonical cycle representation of a permutation, the first entry of each cycle is the largest entry of the cycle. The cycles are then ordered so that the successive first entries appear in increasing order. For example, the canonical representation of our given permutation is $(4),(5,1,3),(6,2)$.

The Stirling number $\left[{n \atop k}\right]$ counts the number of permutations of $\{1,\ldots,n\}$ with k cycles. These numbers satisfy the boundary conditions $\left[{n \atop k}\right]=0$ for $k>n$ and $\left[{n \atop 1}\right]=(n-1)!$. The former condition is obvious, and the latter condition follows once we recall our convention of putting n at the left of the cycle in the canonical representation. All remaining numbers $\left[{n \atop k}\right]$ can be generated via the recurrence relation

$$\left[{n \atop k}\right] \;=\; \left[{n-1 \atop k-1}\right]+(n-1)\left[{n-1 \atop k}\right]. \quad (4.23)$$

The proof of (4.23) parallels that of (4.18). The first term on the right counts the number of ways of adding n as a separate cycle to a permutation π of $\{1,\ldots,n-1\}$ with $k-1$ cycles. The second term on the right counts the number of ways of adding n to an existing cycle of a permutation π of $\{1,\ldots,n-1\}$ with k cycles. If π is such a permutation, then we extend π to n by taking $\pi(n)=i$ for $1\le i\le n-1$. This action conflicts with a current assignment $\pi(j)=i$, so we have to patch things up by defining $\pi(j)=n$. The cycle containing i and j is left intact except for these changes.

We now investigate the number of cycles Y_n in a random permutation π of $\{1,\ldots,n\}$. Clearly, Y_1 is identically 1. If we divide the recurrence (4.23) by $n!$, then we get the recurrence

$$\Pr(Y_n=k) \;=\; \frac{1}{n}\Pr(Y_{n-1}=k-1)+\left(1-\frac{1}{n}\right)\Pr(Y_{n-1}=k).$$

This convolution equation says that $Y_n = Y_{n-1} + Z_n$, where Z_n is independent of Y_{n-1} and follows a Bernoulli distribution with success probability $1/n$. Proceeding inductively, we conclude that in a distributional sense Y_n can be represented as the sum $Z_1 + \cdots + Z_n$ of n independent Bernoulli random variables with decreasing success probabilities. Problem 20 provides a concrete interpretation of Z_k.

We are now in a position to extract useful information about Y_n. For example, the mean number of cycles is

$$\mathrm{E}(Y_n) \;=\; \sum_{k=1}^{n} \frac{1}{k} \;\approx\; \ln n + \gamma,$$

where $\gamma \approx .5772$ is Euler's constant. Because the probability generating function of Z_n is $\mathrm{E}(x^{Z_n}) = 1 - \frac{1}{n} + \frac{x}{n} = \frac{x+n-1}{n}$, the probability generating function of Y_n is

$$\mathrm{E}(x^{Y_n}) \;=\; \frac{1}{n!} x(x+1)\cdots(x+n-1). \tag{4.24}$$

In view of the definition of Y_n, we get the interesting identity

$$\sum_{k=1}^{n} \begin{bmatrix} n \\ k \end{bmatrix} x^k \;=\; n!\,\mathrm{E}(x^{Y_n}) \;=\; x(x+1)\cdots(x+n-1). \tag{4.25}$$

This polynomial identity in $x \in [0,1]$ persists for all real x. Substituting $-x$ for x in equation (4.25) yields the dual identity

$$\sum_{k=1}^{n} (-1)^{n-k} \begin{bmatrix} n \\ k \end{bmatrix} x^k \;=\; x(x-1)\cdots(x-n+1). \tag{4.26}$$

Finally, substituting a random variable X for x in equations (4.25) and (4.26) and taking expectations lead to the relations

$$\begin{aligned} \sum_{k=1}^{n} \begin{bmatrix} n \\ k \end{bmatrix} \mathrm{E}(X^k) &= \mathrm{E}[X(X+1)\cdots(X+n-1)] \\ \sum_{k=1}^{n} (-1)^{n-k} \begin{bmatrix} n \\ k \end{bmatrix} \mathrm{E}(X^k) &= \mathrm{E}[X(X-1)\cdots(X-n+1)] \end{aligned}$$

connecting ordinary moments to rising and falling factorial moments.

We close this section by giving another combinatorial interpretation of Stirling numbers of the first kind. This interpretation is important in the theory of record values for i.i.d. sequences of random variables [5]. A permutation π is said to possess a left-to-right maximum at i provided $\pi(j) < \pi(i)$ for all $j < i$. The Stirling number $\begin{bmatrix} n \\ k \end{bmatrix}$ also counts the number of permutations of $\{1, \ldots, n\}$ with k left-to-right maxima. To prove this assertion,

it suffices to construct a one-to-one correspondence between permutations with k cycles and permutations with k left-to-right maxima. The canonical representation of a permutation achieves precisely this end since the leading number in each cycle is a left-to-right maximum. For example, the permutation (4.22) with canonical representation $(4), (5, 1, 3), (6, 2)$ and three cycles is mapped to the permutation

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 5 & 1 & 3 & 6 & 2 \end{pmatrix}$$

with three left-to-right maxima.

4.7 Application to an Urn Model

The family planning model discussed in Example 2.3.3 is a kind of urn model in which sampling is done with replacement. In other models, sampling without replacement is more appropriate. Such sampling can be realized by starting with n urns and b_j balls in urn j. At each trial a ball is randomly selected from one of the balls currently available and extracted. Under sampling without replacement, this process gradually depletes the supply of balls within the urns. By analogy with the family planning model, it is interesting to set a quota for each urn. When q_j balls from urn j have been drawn, then urn j is said to have reached its quota. Sampling continues until exactly i urns reach their quotas. The trial N_i at which this occurs is a waiting time random variable. Calculation of the moments N_i is challenging. Fortunately, we can apply Proposition 4.4.1 and the magical method of probabilistic embedding [25].

As a concrete example, consider laundering n pairs of dirty socks. Suppose one removes clean socks from the washing machine one by one. If each pair of socks is distinguishable, let N_1 be the number of clean socks extracted until a pair is found. Blom et al. [26] compute the mean of N_1, a problem originally posed and solved by Friedlen [66]. In the current context, pairs of socks correspond to urns and socks to balls. The quota for the jth pair of socks is $q_j = 2$.

The embedding argument works by imagining the balls in urn j as drawn in turn at times determined by a random sample of size b_j from the uniform distribution on $[0, 1]$. We will refer to such a sample as a uniform process. If the uniform processes for the n different urns are independent, then superimposing them creates a uniform process of size $b = \sum_{j=1}^{n} b_j$ on $[0, 1]$. The original discrete-time urn sampling process is said to be embedded in the superposition process.

It is helpful to retain the source of each point in the superposition process. This preserves the time X_i at which the quota q_i for the ith urn is reached. The independence of these attainment times can be used to good effect. In

the superposition process, the order statistic $X_{(i)}$ is the waiting time until i urns reach their quotas. If N_i is the number of trials until the occurrence of this event, then we need to relate the moments of N_i and $X_{(i)}$. The order statistic $X_{(i)}$ can be represented as

$$X_{(i)} = \sum_{j=1}^{N_i} Y_j,$$

where $Y_1, Y_2, \ldots, Y_{b+1}$ are the spacings between adjacent points in the superposition process.

It is straightforward to show that the random distance $Z_m = \sum_{j=1}^{m} Y_j$ to the mth point in the superposition process follows a beta distribution with parameters m and $b-m+1$. Indeed, Z_m is found in the interval $(z, z+dz)$ when one of the b points falls within $(z, z+dz)$, $m-1$ random points fall to the left of z, and $b-m$ random points fall to the right of $z+dz$. This composite event occurs with approximate probability

$$b\binom{b-1}{m-1} z^{m-1}(1-z)^{b-m+1-1}dz.$$

Dividing this probability by dz and letting dz tend to 0 gives the requisite beta density. In view of this fact, conditioning and Example 2.3.1 yield

$$\begin{aligned} \mathrm{E}[X_{(i)}^k] &= \mathrm{E}(Z_{N_i}^k) \\ &= \mathrm{E}\left[\mathrm{E}\left(Z_{N_i}^k \mid N_i\right)\right] \\ &= \frac{1}{(b+1)\cdots(b+k)} \mathrm{E}[N_i \cdots (N_i+k-1)]. \end{aligned} \tag{4.27}$$

Equation (4.27) gives, for instance, $\mathrm{E}(N_i) = (b+1)\,\mathrm{E}[X_{(i)}]$. Furthermore, we can recover all of the ordinary moments $\mathrm{E}(N_i^k)$ from the ascending factorial moments $\mathrm{E}[N_i \cdots (N_i+k-1)]$ via equation (4.21).

Formula (4.12) provides a means of computing $\mathrm{E}[X_{(i)}^k]$ exactly. The independence of the urn-specific uniform processes implies

$$\Pr(X_S > t) = \prod_{l \in S} \left[\sum_{m_l=0}^{q_l-1} \binom{b_l}{m_l} t^{m_l}(1-t)^{b_l-m_l}\right].$$

To calculate the expectations $\mathrm{E}[(X_S)^k]$ required by formula (4.12), we define $\mathbf{m} = (m_l)$ to be a multi-index ranging over the Cartesian product set $R = \{\mathbf{m} : 0 \le m_l \le q_l - 1, l \in S\}$. Then with $|\mathbf{m}| = \sum_{l \in S} m_l$, it follows that

$$\mathrm{E}[(X_S)^k] = k \int_0^\infty t^{k-1} \Pr(X_S > t)\, dt$$

$$\begin{aligned} &= k \sum_{\mathbf{m} \in R} \prod_{l \in S} \binom{b_l}{m_l} \int_0^1 t^{k+|\mathbf{m}|-1} (1-t)^{b-|\mathbf{m}|} \, dt \quad (4.28) \\ &= k \sum_{\mathbf{m} \in R} \prod_{l \in S} \binom{b_l}{m_l} \frac{\Gamma(k+|\mathbf{m}|)\Gamma(b-|\mathbf{m}|+1)}{\Gamma(k+b+1)}, \end{aligned}$$

where $\Gamma(u)$ is the gamma function. In the socks in the laundry problem, formula (4.12) for $\mathrm{E}(X_{(1)}^k)$ reduces to the single term $\mathrm{E}[(X_S)^k]$ with S equal to the full set $\{1, \ldots, n\}$. Equation (4.28) therefore produces the exact solution

$$\begin{aligned} \mathrm{E}[X_{(1)}^k] &= k \sum_{m_1=0}^{1} \cdots \sum_{m_n=0}^{1} \prod_{l=1}^{n} \binom{2}{m_l} \frac{(k+|\mathbf{m}|-1)!(2n-|\mathbf{m}|)!}{(k+2n)!} \\ &= k \sum_{i=0}^{n} 2^i \binom{n}{i} \frac{(k+i-1)!(2n-i)!}{(k+2n)!}, \end{aligned} \quad (4.29)$$

which is most useful for small n. We will revisit this problem from an asymptotic perspective later in Example 12.3.4.

4.8 Application to Faà di Bruno's Formula

Faà di Bruno's formula is an explicit expression for the nth derivative of a composite function $f \circ g(t)$. The formula reads

$$[f \circ g]^{(n)}(t) = \sum \frac{n!}{b_1! \cdots b_n!} f^{(k)}[g(t)] \left[\frac{g^{(1)}(t)}{1!}\right]^{b_1} \cdots \left[\frac{g^{(n)}(t)}{n!}\right]^{b_n},$$

where the sum ranges over all solutions to the equations $\sum_{m=1}^n m b_m = n$ and $\sum_{m=1}^n b_m = k$ in nonnegative integers. For instance, the formula

$$\begin{aligned} [f \circ g]^{(3)}(t) &= f^{(3)}[g(t)]g^{(1)}(t)^3 + 3f^{(2)}[g(t)]g^{(1)}(t)g^{(2)}(t) \\ &\quad + f^{(1)}[g(t)]g^{(3)}(t) \end{aligned}$$

involves the possible values $(3,0,0)$, $(1,1,0)$, and $(0,0,1)$ for the triple (b_1, b_2, b_3). It turns out that Faà di Bruno's formula can be easily deduced by considering the partitions of the set $\{1, \ldots, n\}$ [63]. As we have seen, the Stirling number $\left\{ {n \atop k} \right\}$ counts the number of possible partitions of this set into k disjoint blocks. Let us ask the more nuanced question of how many partitions have exactly b_1 blocks of size 1, b_2 blocks of size 2, and so forth down to b_n blocks of size n. We can relate this problem to permutations π presented in the form $[\pi(1), \ldots, \pi(n)]$ by dividing each such sequence into blocks defined from left to right, with smaller blocks coming before larger blocks. The value $n!$ for the number of permutations overcounts the number

of partitions with the given block sizes in two senses. First, the order of the b_m blocks with m integers each is immaterial. Second, the order of the integers within each block is also immaterial. Hence, the total number of partitions with b_m blocks of size m, $1 \leq m \leq n$, is

$$\frac{n!}{b_1!(1!)^{b_1} \cdots b_n!(n!)^{b_n}}, \tag{4.30}$$

which is precisely the mysterious coefficient appearing in Faà di Bruno's formula.

We now prove Faà di Bruno's formula by induction on n. It sharpens our argument to omit the coefficient (4.30) and extend the sum over all possible set partitions. With this change, the formula becomes

$$[f \circ g]^{(n)}(t) = \sum_{\text{partitions}} f^{(k)}[g(t)]g^{(1)}(t)^{b_1} \cdots g^{(n)}(t)^{b_n}. \tag{4.31}$$

Because it obviously holds for $n = 1$, suppose it is true for $n-1$. The recurrence (4.18) depended on constructing an arbitrary partition of $\{1, \ldots, n\}$ by appending n to an arbitrary partition of $\{1, \ldots, n-1\}$. Let us exploit the same reasoning here. The induction hypothesis implies that each partition of $\{1, \ldots, n-1\}$ corresponds to a term in Faà di Bruno's formula for the derivative $[f \circ g]^{(n-1)}(t)$. With this fact in mind, we apply the product rule of differentiation. Appending n as a singleton block to a given partition corresponds to differentiating the factor $f^{(k)}[g(t)]$. The result $f^{(k+1)}[g(t)]g^{(1)}(t)$ increases the number of blocks of size 1 by 1. Appending n to an existing block of size m corresponds to differentiating the factor $g^{(m)}(t)^{b_m}$. The result $b_m g^{(m)}(t)^{b_m - 1} g^{(m+1)}(t)$ decreases the number of blocks of size m by 1 and increases the number of blocks of size $m+1$ by 1. The factor b_m takes into account the b_m possible blocks of size m to which n can be appended. Thus, the correspondence between partitions and terms in the derivative formula (4.31) carries over from $n-1$ to n.

As an example of Faà di Bruno's formula, let us take $f(t) = e^{\theta t}$ and $g(t) = -\ln(1-t)$. It follows that $h(t) = g \circ f(t) = (1-t)^{-\theta}$. Furthermore, straightforward calculations show that

$$\begin{aligned} f^{(n)}(t) &= \theta^n e^{\theta t} \\ g^{(n)}(t) &= \frac{(n-1)!}{(1-t)^n} \\ h^{(n)}(t) &= \frac{\theta(\theta+1)\cdots(\theta+n-1)}{(1-t)^{\theta+n}}. \end{aligned}$$

Substitution in Faà di Bruno's formula therefore yields

$$\frac{\theta(\theta+1)\cdots(\theta+n-1)}{(1-t)^{\theta+n}} = \sum \frac{n!}{b_1! \cdots b_n!} \frac{\theta^k}{(1-t)^\theta} \prod_{i=1}^{n} \Big[\frac{1}{i(1-t)^i}\Big]^{b_i}.$$

This can be rearranged to give the suggestive identity

$$1 \;=\; \sum \frac{n!\,\theta^k}{\theta(\theta+1)\cdots(\theta+n-1)} \prod_{i=1}^{n} \frac{1}{i^{b_i} b_i!}$$

connecting Faà di Bruno's formula to Ewens' sampling distribution in population genetics. The reader can consult the article of Hoppe [98] for full details.

4.9 Pigeonhole Principle

The pigeonhole principle is an elementary technique of great beauty and utility [32]. In its simplest form, it deals with p pigeons and b boxes (holes), where $p > b$. If we assign the pigeons to boxes, then some box gets more than one pigeon. A stronger form of the pigeonhole principle deals with n numbers $r_1, \ldots, r_n$. If the sum of $r_1 + \cdots + r_n > mn$, then at least one of the r_i satisfies $r_i > m$.

Example 4.9.1 *Longest Increasing Subsequence*

As an application of the principle, consider a sequence $a_1, \ldots, a_{mn+1}$ of $mn+1$ distinct real numbers. We claim that a_i contains an increasing subsequence of length $n+1$ or a decreasing subsequence of length $m+1$. To apply the pigeonhole principle, we suppose the contrary and label each a_i by the length l_i of the longest increasing subsequence commencing with a_i. For example, if $8, 1, 6, 2, 5, 4, 3$ is the sequence, then a longest increasing subsequence beginning with 1 is $1, 2, 5$. Thus, we label 1 with the number $l_2 = 3$. By assumption, no label can exceed n. Let r_l be the number of l_i satisfying $l_i = l$. Because $r_1 + \cdots + r_n = mn + 1$, at least one of the n summands r_l must exceed m. If $r_l = s > m$, then there are s numbers $a_{i_1}, \ldots, a_{i_s}$ such that each a_{i_j} is the start of a maximal increasing subsequence of length l. Now suppose that $a_{i_j} < a_{i_{j+1}}$ for some j. If this is the case, then by appending a_{i_j} to a maximal increasing subsequence beginning with $a_{i_{j+1}}$, we get an increasing subsequence of length $l+1$. This contradicts the label of a_{i_j}. Thus, the $s > m$ numbers $a_{i_1}, \ldots, a_{i_s}$ are in decreasing order. In other words, if an increasing subsequence of length $n+1$ does not exist, then a decreasing subsequence of length $m+1$ does.

We can put this result to good use in a probabilistic version of the longest increasing subsequence problem. Let $X_1, \ldots, X_n$ be independent random variables uniformly distributed on the interval [0,1]. Let I_n and D_n be the random lengths of the longest increasing and decreasing subsequences of $X_1, \ldots, X_n$. These random variables satisfy

$$\sqrt{n} \;\leq\; \max\{I_n, D_n\} \;\leq\; I_n + D_n.$$

Because $\mathrm{E}(I_n) = \mathrm{E}(D_n)$, we therefore conclude that $\mathrm{E}(I_n) \geq \sqrt{n}/2$.

To show that this lower bound is of the correct order of magnitude, we supplement it by a comparable upper bound. We first observe that

$$\Pr(I_n \geq k) \quad \leq \quad \frac{\binom{n}{k}}{k!}.$$

Indeed, there are $\binom{n}{k}$ subsequences of length k of $X_1, \ldots, X_n$, and each has probability $1/k!$ of being increasing. This inequality comes in handy in the estimate

$$\begin{aligned} \mathrm{E}(I_n) &\leq k + n \Pr(I_n \geq k) \\ &\leq k + \frac{n\binom{n}{k}}{k!} \\ &\leq k + \frac{n^{k+1}}{(k!)^2}. \end{aligned} \tag{4.32}$$

The latter estimate can be improved by employing Stirling's approximation

$$k! \quad \asymp \quad \sqrt{2\pi} k^{k+1/2} e^{-k}$$

and choosing k appropriately. The good choice $k = \alpha\sqrt{n}$ yields

$$\begin{aligned} k + \frac{n^{k+1}}{(k!)^2} &\approx k + \frac{n^{k+1}}{(\sqrt{2\pi} k^{k+1/2} e^{-k})^2} \\ &= \alpha\sqrt{n} + \frac{e^{2\alpha\sqrt{n}}\sqrt{n}}{2\pi\alpha^{2(\alpha\sqrt{n}+1/2)}} \\ &= \alpha\sqrt{n} + \frac{e^{-2\alpha\sqrt{n}(\ln\alpha - 1)}\sqrt{n}}{2\pi\alpha}. \end{aligned} \tag{4.33}$$

If we take $\alpha > e$, then $\ln\alpha > 1$, and inequality (4.32) and approximate equality (4.33) together produce $\mathrm{E}(I_n) \leq c\sqrt{n}$ for some $c > \alpha$. Thus, $\mathrm{E}(I_n)$ is on the order of $\sqrt{n}$ in magnitude. Refinements of these arguments show that $\mathrm{E}(I_n)/\sqrt{n}$ tends to a limit as n tends to ∞ [186]. ■

4.10 Problems

1. Prove the following binomial coefficient identities by constructing an appropriate bijection:

$$\begin{aligned} \binom{n}{k} &= \binom{n}{n-k} \\ k\binom{n}{k} &= n\binom{n-1}{k-1} \end{aligned}$$

$$\begin{aligned}
\binom{n}{k}\binom{k}{m} &= \binom{n}{m}\binom{n-m}{k-m} \\
\binom{m+n}{k} &= \sum_{j=0}^{k}\binom{m}{j}\binom{n}{k-j} \\
\sum_{k=0}^{n}\binom{n}{k} &= 2^n \\
\sum_{k=0}^{n}k\binom{n}{k} &= n2^{n-1}.
\end{aligned}$$

Avoid algebra as much as possible, but impose reasonable restrictions on the integers k, m, and n. (Hint: Think of forming a committee of a given size from a class of a given size. You may have to select a subcommittee or committee chair.)

2. Prove the identity

$$\sum_{m=1}^{n-1} m\, m! = n! - 1$$

by a counting argument. (Hint: Let $n - m$ be the first integer not fixed by a permutation π. Thus, $\pi(i) = i$ for $1 \leq i \leq n - m - 1$.)

3. Suppose you select k balls randomly from a box containing n balls labeled 1 through n. Let X_{nk} be the lowest label selected and Y_{nk} be the highest label selected. Demonstrate the mean recurrences

$$\begin{aligned}
\mathrm{E}(X_{nk}) &= \frac{k}{n}\,\mathrm{E}(X_{n-1,k-1}) + \left(1 - \frac{k}{n}\right)\mathrm{E}(X_{n-1,k}) \\
\mathrm{E}(Y_{nk}) &= k + \left(1 - \frac{k}{n}\right)\mathrm{E}(Y_{n-1,k})
\end{aligned}$$

and initial conditions $\mathrm{E}(X_{kk}) = 1$ and $\mathrm{E}(Y_{kk}) = k$. Prove that these recurrences have the unique solutions

$$\mathrm{E}(X_{nk}) = \frac{n+1}{k+1}, \quad \mathrm{E}(Y_{nk}) = \frac{k(n+1)}{k+1}.$$

4. Prove Pascal's identity

$$\sum_{j=1}^{n}\sum_{k=0}^{l-1}\binom{l}{k}b^{l-k}[a + (j-1)b]^k = (a+nb)^l - a^l$$

for a and b positive reals and l and n positive integers. Note that the special case $a = b = 1$ amounts to

$$\sum_{k=0}^{l-1} \binom{l}{k} \sum_{j=1}^{n} j^k = (n+1)^l - 1.$$

(Hints: Drop l random points on the interval $[0, a + nb]$. Divide the interval into $n + 1$ subintervals, the last n of which have length b. If at least one random point falls outside the first subinterval, then let j be the last subinterval of length b containing a random point.)

5. You have 10 pairs of shoes jumbled together in your closet [59]. Show that if you reach in and randomly pull out four shoes, then the probability of extracting at least one pair is 99/323.

6. Prove that there are $8! \sum_{k=0}^{8} \frac{(-1)^k}{k!}$ ways of placing eight rooks on a chessboard so that none can take another and none stands on a white diagonal square [59]. (Hint: Think of the rook positions as a random permutation π, and let A_i be the event $\{\pi(i) = i\}$.)

7. A permutation that satisfies the equation $\pi(\pi(i)) = i$ for all i is called an involution [139]. Prove that a random permutation of $\{1, \ldots, n\}$ is an involution with probability

$$\sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} \frac{1}{2^k k!(n-2k)!}.$$

(Hint: An involution has only fixed points and two-cycles. Count the number of involutions and divide by $n!$.)

8. Define q_r to be the probability that in r tosses of two dice each pair $(1, 1), \ldots, (6, 6)$ appears at least once [59]. Show that

$$q_r = \sum_{k=0}^{6} \binom{6}{k} (-1)^k \Big(\frac{36-k}{36}\Big)^r.$$

9. Calculate the probability $p_{[k]}$ that exactly k suits are missing in a poker hand [59]. To a good approximation $p_{[0]} = .264$, $p_{[1]} = .588$, $p_{[2]} = .146$, and $p_{[3]} = .002$. (Hint: Each hand has probability $1/\binom{52}{5}$.)

10. Give an inclusion-exclusion proof of the identity

$$\left\{ {n \atop k} \right\} = \frac{1}{k!} \sum_{j=0}^{k} (-1)^j \binom{k}{j} (k-j)^n.$$

What is the probability measure and what are the events? (Hint: Consider n labeled balls falling into k labeled boxes.)

11. Let $C_1, \ldots, C_n$ be independent events. Define A_n^m to be the event that at least m of these events occur and B_n^m be the event that exactly m of these events occur. Demonstrate the recurrence relations

$$\begin{aligned} \Pr(A_n^m) &= \Pr(B_{n-1}^{m-1}) \Pr(C_n) + \Pr(A_{n-1}^m) \\ \Pr(B_n^m) &= \Pr(B_{n-1}^{m-1}) \Pr(C_n) + \Pr(B_{n-1}^m)[1 - \Pr(C_n)] \end{aligned}$$

for $m < n$ and $\Pr(A_n^n) = \Pr(B_n^n) = \Pr(B_{n-1}^{n-1}) \Pr(C_n)$, starting from $\Pr(A_1^1) = \Pr(B_1^1) = \Pr(C_1)$, $\Pr(A_1^0) = 1$, and $\Pr(B_1^0) = 1 - \Pr(C_1)$. These recurrences provide an alternative to the method of inclusion-exclusion [157]. Show how they can be applied to find the distribution function of the order statistic $X_{(m)}$ from a sample $X_1, \ldots, X_n$ of independent, not necessarily identically distributed, random variables.

12. Suppose that each of the random variables $X_1, \ldots, X_n$ of Proposition 4.4.1 satisfies $\mathrm{E}(|X_i|^k) < \infty$. Show that all of the expectations $\mathrm{E}(|X_{(i)}|^k)$, $\mathrm{E}(|X_S|^k)$, and $\mathrm{E}(|X^S|^k)$ are finite. (Hints: Bound each random variable in question by $(\sum_{j=1}^n |X_j|)^k$ and apply Minkowski's inequality given in Problem 23 of Chapter 3.)

13. Suppose the n random variables $X_1, \ldots, X_n$ are independent and share the common distribution function $F(x)$. Prove that the jth order statistic $X_{(j)}$ has distribution function

$$F_{(j)}(x) = \sum_{k=j}^{n} \binom{n}{k} F(x)^k [(1 - F(x)]^{n-k}.$$

If $F(x)$ has density $f(x)$, then show that $X_{(j)}$ has density function

$$f_{(j)}(x) = n \binom{n-1}{j-1} F(x)^{j-1} [1 - F(x)]^{n-j} f(x).$$

(Hint: The event $X_{(j)} \leq x$ occurs if and only if at least j of the X_i satisfy $X_i \leq x$ while the remaining X_i satisfy $X_i > x$.)

14. Show that the Catalan numbers satisfy the recurrence

$$c_{n+1} = \frac{2(2n+1)}{n+2} c_n,$$

which is consistent with expression (4.17).

15. Consider a simple random walk of $2n$ steps. Conditional on the event that the walk returns to 0 at step $2n$, show that this is the first return with probability $1/(2n-1)$. (Hint: The first and last steps are in opposite directions. Between these two steps, the walk stays at or above 1 or at or below -1.)

16. List in canonical form the 11 permutations of $\{1, 2, 3, 4\}$ with 2 cycles.

17. Show that $\left[{n \atop n}\right] = \left\{{n \atop n}\right\} = 1$, $\left[{n \atop n-1}\right] = \left\{{n \atop n-1}\right\} = \binom{n}{2}$, and $\left\{{n \atop 2}\right\} = 2^{n-1} - 1$.

18. Demonstrate that the harmonic number $H_n = 1 + \frac{1}{2} + \cdots + \frac{1}{n}$ satisfies $H_n = \frac{1}{n!}\left[{n+1 \atop 2}\right]$. (Hint: Apply equation (4.23) and $H_n = H_{n-1} + \frac{1}{n}$.)

19. Prove the inequality $\left[{n \atop k}\right] \geq \left\{{n \atop k}\right\}$ for all n and k by invoking the definitions of the two kinds of Stirling numbers.

20. In our discussion of Stirling numbers of the first kind, we showed that the number of left-to-right maxima Y_n of a random permutation has the decomposition $Y_n = Z_1 + \cdots + Z_n$, where the Z_k are independent Bernoulli variables with decreasing success probabilities. Prove that Z_k can be interpreted as the indicator of the event that position $n-k+1$ is a left-to-right maximum. In other words, Z_k is the indicator of the event $\{\pi(n-k+1) > \pi(j) \text{ for } 1 \leq j < n-k+1\}$. Why are the Z_k independent [26]?

21. Suppose the random variable X is nonnegative, bounded, and integer valued. Show that the probability generating function $G(u) = \mathrm{E}(u^X)$ of X can be expressed as

$$G(u) = \sum_{j=0}^{\infty} \mathrm{E}\left[\binom{X}{j}\right](u-1)^j$$

using the binomial moments

$$\mathrm{E}\left[\binom{X}{j}\right] = \frac{1}{j!}\mathrm{E}[X(X-1)\cdots(X-j+1)] = \frac{1}{j!}\frac{d^j}{du^j}G(1).$$

Now consider the special case where $X = 1_{A_1} + \cdots + 1_{A_n}$ is a sum of indicator random variables. In view of the trivial identity

$$u^{1_{A_i}} = 1 + 1_{A_i}(u-1),$$

demonstrate that

$$G(u) = \sum_{j=0}^{n} \sum_{|S|=j} \Pr\left(\bigcap_{i \in S} A_i\right)(u-1)^j,$$

where $|S|$ is the number of elements of the subset S of $\{1, \ldots, n\}$. Equating coefficients of $(u-1)^j$ in these two representations of $G(u)$ yields the identity

$$\mathrm{E}\left[\binom{X}{j}\right] = \sum_{|S|=j} \Pr\left(\bigcap_{i \in S} A_i\right). \tag{4.34}$$

22. Let X be the number of fixed points of a random permutation of $\{1, \ldots, n\}$. Demonstrate that

$$\mathrm{E}(X^j) \;=\; \sum_{k=1}^{\min\{j,n\}} \left\{ {j \atop k} \right\}.$$

(Hint: Find the binomial moments of X via equation (4.34), convert these to factorial moments, and then convert these to ordinary moments [139].)

23. Suppose π is a random permutation of $\{1, \ldots, n\}$. Show that

$$\mathrm{E}\left\{ \sum_{j=1}^{n-1} [\pi(j) - \pi(j+1)]^2 \right\} \;=\; \binom{n+1}{3}.$$

(Hint: Each term has the same expectation [139].)

24. Balls are randomly extracted one by one from a box containing b black balls and w white balls. Show that the expected number of black balls left when the last white ball is extracted equals $\frac{b}{w+1}$. (Hint: Let the extraction times of the black balls and the white balls constitute two independent uniform processes on $[0, 1]$. Condition on the time when the last white ball is extracted.)

25. Consider a random graph with n nodes. Between every pair of nodes, we independently introduce an edge with probability p. A trio of nodes forms a triangle if each of its three pairs is connected by an edge. If N counts the number of triangles, then demonstrate that $\mathrm{E}(N) = \binom{n}{3}p^3$ and $\mathrm{Var}(N) = \binom{n}{3}p^3(1-p^3) + \binom{n}{3}3(n-3)(p^5 - p^6)$.

26. Consider the n-dimensional unit cube $[0, 1]^n$. Suppose that each of its $n2^{n-1}$ edges is independently assigned one of two equally likely orientations. Let S be the number of vertices at which all neighboring edges point toward the vertex. Show that S has mean $\mathrm{E}(S) = 1$ and variance $\mathrm{Var}(S) = 1 - (n+1)2^{-n}$. When n is large, S follows an approximate Poisson distribution. (Hint: Let X_α be the indicator that vertex α has all of its edges directed toward α. Note that X_α is independent of X_β unless α and β share an edge. If α and β share an edge, then $X_\alpha X_\beta = 0$.)

27. Let X be a random variable with moment generating function

$$M(t) \;=\; \mathrm{E}\left(e^{tX}\right) \;=\; \sum_{n=0}^{\infty} \frac{\mu_n}{n!} t^n$$

defined in some neighborhood of the origin. Here $\mu_n = \mathrm{E}(X^n)$ is the nth moment of X. The function

$$\ln M(t) = \sum_{n=1}^{\infty} \frac{\kappa_n}{n!} t^n$$

is called the cumulant generating function, and its nth coefficient κ_n is called the nth cumulant. Based on Faà di Bruno's formula, show that

$$\begin{aligned} \mu_n &= \sum \frac{n!}{\prod_{m=1}^n b_m!(m!)^{b_m}} \kappa_1^{b_1} \cdots \kappa_n^{b_n} \\ \kappa_n &= \sum \frac{n!(-1)^{k-1}(k-1)!}{\prod_{m=1}^n b_m!(m!)^{b_m}} \mu_1^{b_1} \cdots \mu_n^{b_n}, \end{aligned}$$

where the sum ranges over solutions to the equations $\sum_{m=1}^n m b_m = n$ and $\sum_{m=1}^n b_m = k$ in nonnegative integers. In particular verify the relationships

$$\begin{aligned} \mu_1 &= \kappa_1 \\ \mu_2 &= \kappa_2 + \kappa_1^2 \\ \mu_3 &= \kappa_3 + 3\kappa_1\kappa_2 + \kappa_1^3 \\ \mu_4 &= \kappa_4 + 4\kappa_1\kappa_3 + 3\kappa_2^2 + 6\kappa_1^2\kappa_2 + \kappa_1^4 \end{aligned}$$

and

$$\begin{aligned} \kappa_1 &= \mu_1 \\ \kappa_2 &= \mu_2 - \mu_1^2 \\ \kappa_3 &= \mu_3 - 3\mu_1\mu_2 + 2\mu_1^3 \\ \kappa_4 &= \mu_4 - 4\mu_1\mu_3 - 3\mu_2^2 + 12\mu_1^2\mu_2 - 6\mu_1^4. \end{aligned}$$

28. Continuing Problem 27, show that cX has nth cumulant $c^n\kappa_n$ and that $X + c$ has first cumulant $\kappa_1 + c$ and subsequent cumulants κ_n. If Y is independent of X and has nth cumulant η_n, then demonstrate that $X + Y$ has nth cumulant $\kappa_n + \eta_n$.

29. Five points are chosen from an equilateral triangle with sides of length 1. Demonstrate that there exist two points separated by a distance of at most $1/2$.

30. Suppose $n + 1$ numbers are chosen from the set $\{1, 2, \ldots, 2n\}$. Show that there is some pair having no common factor other than 1. Show that there is another pair such that one member of the pair is divisible by the other.

31. Consider a graph with more than one node. Let d_i be the degree of node i. Prove that at least two d_i coincide.

32. Given n integers $a_1, \ldots, a_n$, demonstrate that there is some sum $\sum_{i=j+1}^{k} a_i$ that is a multiple of n. (Hint: Map each of the $n+1$ partial sums $s_j = \sum_{i=1}^{j} a_i$ into its remainder after division by n.)

5
Combinatorial Optimization

5.1 Introduction

Combinatorial averaging is a supple tool for understanding the solutions of discrete optimization problems. Computer scientists have designed many algorithms to solve such problems. Traditionally, these algorithms have been classified by their worst-case performance. Such an analysis can lead to undue pessimism. The average behavior of an algorithm is usually more relevant. Of course, to evaluate the average complexity of an algorithm, we must have some probability model for generating typical problems on which the algorithm operates. The examples in this chapter on sorting, data compression, and graph coloring illustrate some of the underlying models and the powerful techniques probabilists have created for analyzing algorithms.

Not only is combinatorial averaging helpful in understanding the complexity of algorithms, but it can also yield nonconstructive existence proofs and verify that a proposed solution of a discrete optimization problem is optimal [1, 4, 78, 206]. The former role is just the probabilistic method of combinatorics initiated by Erdös and Rényi. In the probabilistic method, we take a given set of objects, embed it in a probability space, and show that the subset of objects lacking a certain property has probability less than 1. The subset of objects possessing the property must therefore be nonempty. Alternatively, if the property is determined by some number X assigned to each object, then we can view X as a random variable and calculate its expectation. If the property holds for $X \leq c$ and $\mathrm{E}(X) \leq c$,

K. Lange, *Applied Probability,*
DOI 10.1007/978-1-4419-7165-4_5, © Springer Science+Business Media, LLC 2010

then some object with the property exists. Our treatment of Sperner's theorem illustrates the role of probability in discrete optimization. Finally, we discuss in the current chapter subadditive and superadditive sequences and their application to the longest common subsequence problem. The linear growth in complexity seen in this problem does not always occur, as our concluding example on the Euclidean traveling salesman problem shows.

5.2 Quick Sort

Sorting lists of items such as numbers or words is one of the most thoroughly studied tasks in computer science. It is a pleasant fact that the fastest sorting algorithm can be explained by a probabilistic argument [206]. At the heart of this argument is a recurrence relation specifying the average number of operations encountered in sorting n numbers. In this problem, we can explicitly solve the recurrence relation and estimate the rate of growth of its solution as a function of n.

The quick sort algorithm is based on the idea of finding a splitting entry x_i of a sequence $x_1, \ldots, x_n$ of n distinct numbers in the sense that $x_j < x_i$ for $j < i$ and $x_j > x_i$ for $j > i$. In other words, a splitter x_i is already correctly ordered relative to the rest of the entries of the sequence. Finding a splitter reduces the computational complexity of sorting because it is easier to sort both of the subsequences $x_1, \ldots, x_{i-1}$ and $x_{i+1}, \ldots, x_n$ than it is to sort the original sequence. At this juncture, one can reasonably object that no splitter need exist, and even if one does, it may be difficult to locate. The quick sort algorithm avoids these difficulties by randomly selecting a splitting value and then slightly rearranging the sequence so that this splitting value occupies the correct splitting location.

In the background of quick sort is the probabilistic assumption that all $n!$ permutations of the n values are equally likely. The algorithm begins by randomly selecting one of the n values and moving it to the leftmost or first position of the sequence. Through a sequence of exchanges, this value is then promoted to its correct location. In the probabilistic setting adopted, the correct location of the splitter is uniformly distributed over the n positions of the sequence.

The promotion process works by exchanging or swapping entries to the right of the randomly chosen splitter x_1, which is kept in position 1 until a final swap. Let j be the current position of the sequence as we examine it from left to right. In the sequence up to position j, a candidate position i for the insertion of x_1 must satisfy the conditions $x_k < x_1$ for $1 < k \leq i$ and $x_k > x_1$ for $i < k \leq j$. At position $j = 1$, we are forced to put $i = 1$. This choice works because then the set $\{k : 1 < k \leq i \text{ or } i < k \leq j\}$ is empty. Now suppose we have successfully advanced to a general position j and identified a corresponding candidate position i. To move from position

j to position $j+1$, we examine x_{j+1}. If $x_{j+1} > x_1$, then we keep the current candidate position i. On the other hand, if $x_{j+1} < x_1$, then we swap x_{i+1} and x_{j+1} and replace i by $i+1$. In either case, the two required conditions imposed on i continue to hold in moving from position j to position $j+1$. It is now clear that we can inductively march from the left end to the right end of the sequence, carrying out a few swaps in the process, so that when $j = n$, the value i marks the correct position to insert x_1. Once this insertion is made, the subsequences $x_1, \ldots, x_{i-1}$ and $x_{i+1}, \ldots, x_n$ can be sorted separately by the same splitting procedure.

Now let e_n be the expected number of operations involved in quick sorting a sequence of n numbers. By convention $e_0 = 0$. If we base our analysis only on how many positions j must be examined at each stage and not on how many swaps are involved, then we can write the recurrence relation

$$\begin{aligned} e_n &= n - 1 + \frac{1}{n}\sum_{i=1}^{n}(e_{i-1} + e_{n-i}) \\ &= n - 1 + \frac{2}{n}\sum_{i=1}^{n} e_{i-1} \end{aligned} \tag{5.1}$$

by conditioning on the correct position i of the first splitter.

The recurrence relation (5.1) looks formidable, but a few algebraic maneuvers render it solvable. Multiplying equation (5.1) by n produces

$$ne_n = n(n-1) + 2\sum_{i=1}^{n} e_{i-1}.$$

If we subtract from this the corresponding expression for $(n-1)e_{n-1}$, then we get

$$ne_n - (n-1)e_{n-1} = 2n - 2 + 2e_{n-1},$$

which can be rearranged to give

$$\frac{e_n}{n+1} = \frac{2(n-1)}{n(n+1)} + \frac{e_{n-1}}{n}. \tag{5.2}$$

Equation (5.2) can be iterated to yield

$$\begin{aligned} \frac{e_n}{n+1} &= 2\sum_{k=1}^{n}\frac{(k-1)}{k(k+1)} \\ &= 2\sum_{k=1}^{n}\left(\frac{2}{k+1} - \frac{1}{k}\right) \\ &= 2\sum_{k=1}^{n}\frac{1}{k} - \frac{4n}{n+1}. \end{aligned}$$

Because $\sum_{k=1}^{n} \frac{1}{k}$ approximates $\int_1^n \frac{1}{x}\,dx = \ln n$, it follows that

$$\lim_{n\to\infty} \frac{e_n}{2n \ln n} = 1.$$

Quick sort is, indeed, a very efficient algorithm on average. Press et al. [163] provide good computer code implementing it.

5.3 Data Compression and Huffman Coding

Huffman coding is an algorithm for data compression without loss of information [163, 168, 177]. In this section we present the algorithm and prove its optimality in an average sense. To motivate Huffman coding, it is useful to think of an alphabet $\mathcal{A}$ with typical letter $l \in \mathcal{A}$. From previous experience with the alphabet, we can assign a usage probability p_l to l. Inside a computer, we represent l using a bit string s_l, each bit having the value 0 or 1. One possibility is to use bit strings of fixed length to represent all letters. This is an inefficient allocation of memory if there is wide variation in the probabilities p_l. Huffman coding uses bit strings of varying length, with frequent letters assigned short strings and infrequent letters assigned long strings.

In addition to the obvious requirement that no two assigned bit strings coincide, we require instantaneous decoding. This is motivated by the necessity of recording words and a sequence of words. Words are separated by spaces, so we enlarge our alphabet to contain a space symbol if necessary. Thus, if we want to encode a message, we do so letter by letter and concatenate the corresponding bit strings. This tactic leads to confusion if we fail to design the bit strings properly. For example, if the alphabet is the ordinary alphabet, we could conceivably assign the letter e the bit string 111 and the letter a the bit string 1110. When we encounter the three bits 111 in the encoded message, we then face the ambiguity of whether we have an e or the start of an a. Consequently, we impose the further constraint that no prefix of a bit string representing a letter coincides with a bit string representing a different letter. We interpret "prefix" to mean either a beginning portion of a bit string or the whole bit string.

Huffman coding solves the instantaneous decoding problem by putting all letters at the bottom of a binary tree. For example, Figure 5.1 shows the Huffman tree corresponding to the alphabet $\mathcal{A} = \{\text{a,e,i,o,u}\}$ consisting of the vowels. To construct the bit string for a given letter, we just traverse the tree from the root at its top to the corresponding letter node at its bottom. Each new edge encountered adds a 0 or 1 to the bit string for the letter. Every left edge taken adds a 0, and every right edge taken adds a 1. Thus, we represent the letter o by the bit string 00 and the letter u by the bit string 100 in Figure 5.1.

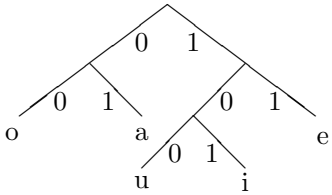


FIGURE 5.1. The Huffman Vowel Tree

In an arbitrary encoding of the alphabet $\mathcal{A}$, we also assign each letter l a bit string. The number of bits (or length) in a bit string s is denoted by $\text{len}(s)$. We can view a code S as a random map taking the random letter l to its bit string s_l. Huffman coding minimizes the average letter length

$$\text{E}[\text{len}(S)] \quad = \quad \sum_{l \in \mathcal{A}} \text{len}(s_l) p_l.$$

A Huffman code is constructed recursively. Consider an alphabet $\mathcal{A}_n$ of n letters $l_1, \ldots, l_n$ arranged so that $p_{l_1} \geq \cdots \geq p_{l_n}$. In the event of one or more ties $p_m = p_{m+1}$, there will be multiple Huffman codings with the same average length. Because we want the most infrequent letters to reside at the bottom of the tree, we build a minitree by joining l_n on the left and l_{n-1} on the right to a parent node above designated $m_{n-1} = \{l_n, l_{n-1}\}$. To node m_{n-1} we attribute probability $p_{l_n} + p_{l_{n-1}}$. We then proceed to construct the Huffman tree for the alphabet $\mathcal{A}_{n-1} = \{l_1, \ldots, l_{n-2}, m_{n-1}\}$. At the final stage of the Huffman algorithm, we have a single node, which becomes the root of the tree.

For example, the vowels have approximate usage probabilities $p_a = .207$, $p_e = .332$, $p_i = .185$, $p_o = .203$, and $p_u = .073$ in English [168]. Huffman coding first combines u on the left with i on the right into the node $\{u, i\}$ with probability .258. Second, it combines o on the left with a on the right into the node $\{o, a\}$ with probability .410. Third, it combines $\{u, i\}$ on the left with e on the right into the node $\{u, i, e\}$ with probability .590. Finally, it combines $\{o, a\}$ on the left with $\{u, i, e\}$ on the right into the root.

In proving that Huffman coding is optimal, let us simplify notation by identifying the n letters of the alphabet $\mathcal{A}_n$ with the integers $1, \ldots, n$. Under the innocuous assumption $p_1 \geq \cdots \geq p_n$, there are two general methods for improving any instantaneous coding S. First, we can assume that $\text{len}(s_j) \leq \text{len}(s_k)$ whenever $j < k$. If this is not the case, then we can improve S by interchanging s_j and s_k. Second, we can represent any bit string s_j in S by $s_j = (x, y)$, where x is the longest prefix of s_j coinciding with a prefix of any other bit string s_k. If y contains more than just its initial bit y_1, then we can improve S by truncating s_j to (x, y_1).

String truncation has implications for the length of the longest bit string s_n. Suppose that $s_n = (x, y)$ and $\text{len}(s_n) > \text{len}(s_{n-1})$. The longest matching prefix x satisfies $\text{len}(x) < \text{len}(s_{n-1})$; otherwise, x coincides with s_{n-1} or some other bit string having the same length as s_{n-1}. Once we replace s_n by (x, y_1), then we can assume that $\text{len}(s_n) \leq \text{len}(s_{n-1})$. If strict inequality prevails, then we interchange the new s_n with s_{n-1}. If we continue truncating the longest bit string, eventually we reach the point where $s_n = (x, y_1)$ and $\text{len}(s_n) = \text{len}(s_{n-1})$. By definition of x, the bit string $(x, y_1 + 1 \bmod 2)$ also is in S. Performing a final interchange if necessary, we can consequently assume that s_n and s_{n-1} have the same length and differ only in their last bit.

Now consider the Huffman coding H_n of $\mathcal{A}_n$ with h_m denoting the bit string corresponding to m. Huffman's construction replaces the letters n and $n-1$ by a single letter with probability $p_n + p_{n-1}$. If we let h denote the bit string assigned to this new letter, then $\text{len}(h)+1 = \text{len}(h_n) = \text{len}(h_{n-1})$. The old and new Huffman trees therefore satisfy

$$\begin{aligned} \text{E}[\text{len}(H_n)] &= \text{E}[\text{len}(H_{n-1})] - \text{len}(h)(p_n + p_{n-1}) \\ &\quad + [\text{len}(h) + 1]p_n + [\text{len}(h) + 1]p_{n-1} \\ &= \text{E}[\text{len}(H_{n-1})] + (p_n + p_{n-1}). \end{aligned} \tag{5.3}$$

Now consider an alternative coding S_n of $\mathcal{A}_n$. As just demonstrated, we can assume that s_n and s_{n-1} have the same length and differ only in their last bit. The changes necessary to achieve this goal can only decrease the average length of S_n. Without loss of generality, suppose that $s_n = (x, 0)$ and $s_{n-1} = (x, 1)$. This assumption places $n-1$ and n next to each other at the bottom of the tree for S_n. Assigning the amalgamation of n and $n-1$ the bit string x and the probability $p_n + p_{n-1}$ leads to a code S_{n-1} on $\mathcal{A}_{n-1}$ satisfying

$$\begin{aligned} \text{E}[\text{len}(S_n)] &= \text{E}[\text{len}(S_{n-1})] - \text{len}(x)(p_n + p_{n-1}) \\ &\quad + [\text{len}(x) + 1]p_n + [\text{len}(x) + 1]p_{n-1} \\ &= \text{E}[\text{len}(S_{n-1})] + (p_n + p_{n-1}). \end{aligned} \tag{5.4}$$

If we assume by induction on n that $\text{E}[\text{len}(H_{n-1})] \leq \text{E}[\text{len}(S_{n-1})]$, then equations (5.3) and (5.4) prove that $\text{E}[\text{len}(H_n)] \leq \text{E}[\text{len}(S_n)]$. Given the obvious optimality of Huffman coding when $n = 2$, this finishes our inductive argument that Huffman coding minimizes average code length.

5.4 Graph Coloring

In the graph coloring problem, we are given a graph with n nodes and asked to color each node with one color from a palette of k colors [207].

Two adjacent nodes must be colored with different colors. For the sake of convenience, let us label the nodes $1, \ldots, n$ and the colors $1, \ldots, k$. The solution to the well-known four-color problem states that any planar graph can be colored with at most four colors. Roughly speaking, a graph is planar if it can be drawn in two dimensions in such a way that no edges cross. For example, if we wish to color a map of contiguous countries, then countries are nodes, and edges connect adjacent countries. In the general graph coloring problem, the graphs need not be planar.

It turns out that a standard computer science technique called backtracking will solve every graph coloring problem. The catch is that backtracking is extremely inefficient for large n on certain worst-case graphs. Nonetheless, the average behavior of backtracking is surprisingly good as n increases. The reason for this good performance is that we can reject the possibility of k-coloring most graphs with n nodes.

To illustrate the backtracking algorithm, consider the simple graph of Figure 5.2 with $n = 4$ nodes. This graph can be colored in several ways with three colors. For instance, one solution is the coloring 1213 that assigns color 1 to node 1, color 2 to node 2, color 1 to node 3, and color 3 to node 4. If carried to completion, the backtracking algorithm will construct all possible colorings. To commence the backtracking algorithm, we assign color 1 to node 1. We then are forced to assign colors 2 or 3 to node 2 to avoid a conflict. In the former case, we represent the partial coloring of the first two nodes by 12. In backtracking we keep growing a partial solution until we reach a full solution or a forbidden color match between two neighboring nodes. When either of these events occur, we backtrack to the first available full or partial solution that we have not previously encountered.

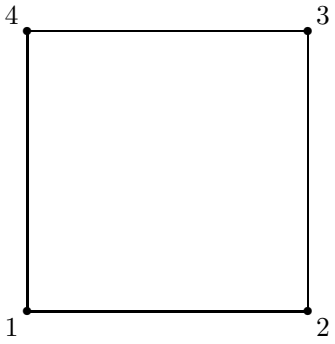


FIGURE 5.2. A Graph with Four Nodes

In our simple example, we extend the partial solution 12 to the larger partial solution 121 by choosing the first available color (color 1) for node 3. From there we extend to the full solution 1212 by choosing the first available color (color 2) for node 4. We then substitute the next available color (color

3) for node 4 to reach the next full solution 1213. At this point, we have exhausted full solutions and are forced to backtrack to node 3 and assign the next available color (color 3) to it. This gives the partial solution 123, which can be grown to the full solution 1232 before backtracking. Figure 5.3 depicts the family of partial and full solutions generated by backtracking with color 1 assigned to node 1. The full backtracking tree with node 1 assigned any of the three available colors is similar to Figure 5.3 but too large to draw.

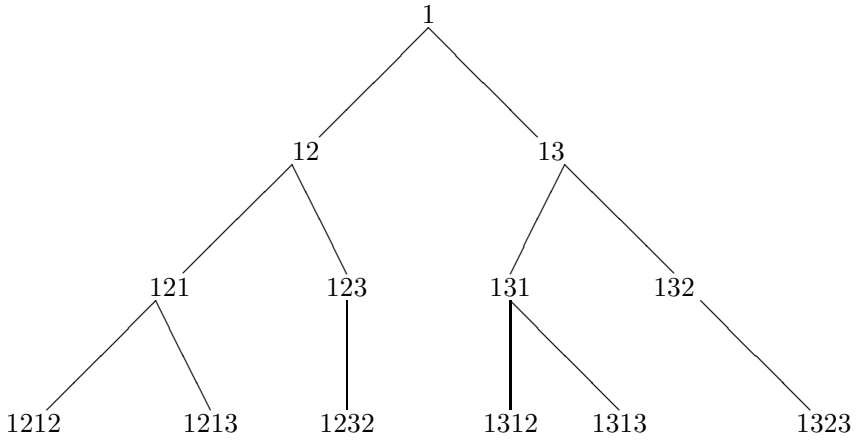


FIGURE 5.3. A Backtracking Tree

In summary, we can depict the functioning of the backtracking algorithm by drawing a backtracking tree with $n+1$ levels. Level 0 (not shown in Figure 5.3) is a root connected to the k partial solutions $1, \ldots, k$ involving node 1 at level 1. Level l of the backtracking tree contains partial solutions involving nodes 1 through l. Each partial solution is a sequence of length l with entries chosen from the integers 1 to k. The backtracking algorithm finds at least one full solution if and only if the backtracking tree descends all the way to level n.

We now assess the average computational complexity of the backtracking algorithm. To do so, we set up a simple probability model for random graphs with n nodes. There are $\binom{n}{2}$ possible edges in a graph with n nodes and a total of $2^{\binom{n}{2}}$ possible graphs. The uniform distribution on this sample space can be achieved by considering each pair of nodes in turn and independently introducing an edge between the nodes with probability $\frac{1}{2}$. For each given graph G with n nodes, we imagine constructing the corresponding backtracking tree and recording the number $X_l(G)$ of partial solutions at level l. (If $l = n$, the partial solutions are full solutions.) The amount of work done in backtracking is proportional to $\sum_{l=1}^{n} X_l$. With this notation, our goal is to demonstrate the remarkable fact that $\sum_{l=1}^{n} \mathrm{E}(X_l)$ is bounded above by a constant that does not depend on n.

To prove this claim, we need to bound $\mathrm{E}(X_l)$. Now each partial solution in the backtracking tree at level l represents a proper coloring of the first l nodes of G. There are, of course, k^l possible colorings of l nodes. Instead of trying to estimate the number of colorings compatible with each subgraph on l nodes, let us try to estimate the probability that a random subgraph on l nodes is compatible with a given coloring of the l nodes. Under the uniform distribution, each of the $2^{\binom{l}{2}}$ possible subgraphs on l nodes is equally likely. Many subgraphs can be eliminated as inconsistent with the given coloring because they involve forbidden edges.

Suppose m_i nodes have color i. Because we can draw an edge only between nodes of different colors, the total number of permitted edges is

$$\begin{aligned}\sum_{i<j} m_i m_j &= \frac{1}{2}\sum_{i=1}^{k}\sum_{j\neq i} m_i m_j \\ &= \frac{1}{2}\left(\sum_{i=1}^{k} m_i\right)^2 - \frac{1}{2}\sum_{i=1}^{k} m_i^2. \qquad (5.5)\end{aligned}$$

The variance inequality $\frac{1}{k}\sum_{i=1}^{k} m_i^2 - \left(\frac{1}{k}\sum_{i=1}^{k} m_i\right)^2 \geq 0$ and the counting identity $\sum_{i=1}^{k} m_i = l$ together imply $-\sum_{i=1}^{k} m_i^2 \leq -\frac{l^2}{k}$. Substituting this inequality in equation (5.5) produces the upper bound

$$\sum_{i<j} m_i m_j \leq \frac{l^2}{2} - \frac{l^2}{2k}$$

on the total number of possible edges. Thus, the maximum number of graphs compatible with a given coloring is $2^{l^2/2 - l^2/(2k)}$, and the probability that a random graph is compatible is at most

$$\frac{2^{\frac{l^2}{2}-\frac{l^2}{2k}}}{2^{\binom{l}{2}}} = 2^{\frac{l}{2}-\frac{l^2}{2k}}. \qquad (5.6)$$

We now write

$$X_l = \sum_c 1_{A_c},$$

where c is a coloring of the first l nodes and A_c is the event that the underlying random graph is consistent at the first l nodes with the coloring. Taking expectations in this identity and invoking inequality (5.6) yield

$$\mathrm{E}(X_l) \leq k^l 2^{\frac{l}{2}-\frac{l^2}{2k}}$$

and therefore the upper bound

$$\sum_{l=1}^{n} \mathrm{E}(X_l) \leq \sum_{l=1}^{n} k^l 2^{\frac{l}{2}-\frac{l^2}{2k}} \qquad (5.7)$$

on the average number of partial solutions in the backtracking tree. The limit of the series on the right-hand side of inequality (5.7) exists as n tends to ∞ by the ratio test. Indeed, the ratio of term $l+1$ to term l is

$$\frac{k^{l+1}2^{\frac{l+1}{2}-\frac{(l+1)^2}{2k}}}{k^l 2^{\frac{l}{2}-\frac{l^2}{2k}}} = k2^{\frac{1}{2}-\frac{2l+1}{2k}}, \tag{5.8}$$

which tends to 0 as l tends to ∞. When $k = 3$, the limit of the series (5.7) is about 197. In other words, the average backtracking tree contains fewer than 197 partial solutions regardless of the size n of the graph. Once again, let us stress that this result is simply a manifestation of the fact that most graphs with n nodes are quickly eliminated as colorable with k colors. ■

5.5 Point Sets with Only Acute Angles

Consider a finite set of points S in some Euclidean space R^d. Any three points $x, y, z \in S$ determine three angles, depending on which point is taken as apex. For example, taking x as apex produces two vectors $y-x$ and $z-x$ with an angle between them that is obtuse, right, or acute when the inner product $(y-x)^t(z-x)$ is negative, zero, or positive, respectively. It seems likely that at least some of these angles will be obtuse if the number of points is large. For example, Problem 13 asks the reader to check that any five noncollinear points in the plane determine at least one obtuse angle. Asking for only acute angles would seem to diminish the possibilities. Nonetheless, for high-dimensional spaces, it is possible to construct large sets of points with only acute angles [55].

In approaching this problem, we will limit ourselves to sets S contained in the vertex set $\{0,1\}^d$ of the d-dimensional unit cube. This has the advantage of eliminating the possibility of obtuse angles. Indeed, if we express

$$(y-x)^t(z-x) = \sum_{i=1}^{d}(y_i-x_i)(z_i-x_i),$$

then all of the products in the indicated sum are 0 or 1 because the coordinates x_i, y_i, z_i are chosen from $\{0,1\}$. When we take $S = \{0,1\}^d$, we attain a set of maximal size with no obtuse angles [1]. However, many of the angles are right. Let us consider smaller sets $S \subset \{0,1\}^d$ of size m, where m is to be decided later. Instead of picking S directly, we first construct a set T containing $2m$ random points chosen independently and uniformly from $\{0,1\}^d$. Some of these points may coincide.

Now consider three points $x, y, z \in T$. Let us call the triple (x, y, z) with apex x a bad triple whenever $(y-x)^t(z-x) = 0$. What is the probability of this happening? For each coordinate i we must have $(y_i - x_i)(z_i - x_i) = 0$. This occurs if either $y_i = x_i$ or $z_i = x_i$ and therefore has probability

$\frac{3}{4}$. Because the coordinates are chosen independently, the inner product $(y-x)^t(z-x)$ vanishes with probability $(\frac{3}{4})^d$. We now calculate the expected number of bad triples. The total number of triples is $\binom{2m}{3}$. Each triple has three possible choices for its apex. Thus the expected number of bad triples is

$$3\binom{2m}{3}\Big(\frac{3}{4}\Big)^d \quad < \quad m(2m)^2\Big(\frac{3}{4}\Big)^d.$$

We now choose m so that the right-hand side of this inequality is less than m. For example, we can take

$$m \quad = \quad \Big\lfloor \frac{1}{2}\Big(\frac{2}{\sqrt{3}}\Big)^d \Big\rfloor.$$

With this choice, there must be at least one configuration T with m or fewer bad triples. For such a T, we throw out the apex of any bad triple. This creates a set S with m or more points and no bad triples. The points of S define only acute angles, no right angles. For example, if $d = 35$, then there is some set S with at least $m = 76$ points defining only acute angles.

5.6 Sperner's Theorem

For a positive integer n, consider a family $\mathcal{F}$ of nonempty subsets of the set $\{1, \ldots, n\}$. In Sperner's theorem, we impose the condition that two distinct subsets A and B in $\mathcal{F}$ satisfy neither $A \subset B$ nor $B \subset A$. With this restriction, how many subsets can $\mathcal{F}$ contain?

One extreme case is to take $\mathcal{F}$ to consist of all subsets of $\{1, \ldots, n\}$ having exactly $\lfloor \frac{n}{2} \rfloor$ elements. Because two subsets of the same size either coincide or satisfy the Sperner restriction, it is clear that $\mathcal{F}$ qualifies as a Sperner family. This family contains

$$|\mathcal{F}| \quad = \quad \binom{n}{\lfloor \frac{n}{2} \rfloor}$$

subsets. Following Lubell [139], we now show that this special family contains the maximum possible number of subsets.

Our line of attack proceeds through random permutations. To a given permutation π of $\{1, \ldots, n\}$, there correspond n subsets of the form

$$S(\pi, k) \quad = \quad \{\pi(1), \pi(2), \ldots, \pi(k)\}.$$

These satisfy $S(\pi, k) \subset S(\pi, k+1)$ for $1 \leq k \leq n-1$. Now consider the random variable

$$X(\pi) \quad = \quad \sum_{k=1}^{n} 1_{\{S(\pi,k) \in \mathcal{F}\}}$$

defined relative to a Sperner family $\mathcal{F}$. Because at most one of the events $\{S(\pi,k) \in \mathcal{F}\}$ can occur, X must equal either 0 or 1, and $0 \leq \mathrm{E}(X) \leq 1$.

In view of the fact that $X(\pi)$ counts the number of $S(\pi,k)$ in $\mathcal{F}$, we can also write

$$X(\pi) \quad = \quad \sum_{A \in \mathcal{F}} 1_{\{S(\pi,|A|)=A\}},$$

where $|A|$ denotes the number of elements of A. Taking into account the facts that each $S(\pi,|A|)$ is a randomly chosen subset of size $|A|$ and that $\binom{n}{k}$ is maximized by $k = \lfloor \frac{n}{2} \rfloor$, we calculate

$$\begin{aligned} \mathrm{E}[X] &= \sum_{A \in \mathcal{F}} \Pr[S(\pi,|A|) = A] \\ &= \sum_{A \in \mathcal{F}} \frac{1}{\binom{n}{|A|}} \\ &\geq \frac{|\mathcal{F}|}{\binom{n}{\lfloor \frac{n}{2} \rfloor}}, \end{aligned}$$

where $|\mathcal{F}|$ is the size of $\mathcal{F}$. Combining this inequality with our earlier inequality $\mathrm{E}(X) \leq 1$ leads to the desired conclusion $|\mathcal{F}| \leq \binom{n}{\lfloor \frac{n}{2} \rfloor}$.

5.7 Subadditivity and Expectations

Many solutions to hard discrete optimization problems involve complicated random variables whose distributions and moments are nearly impossible to calculate exactly. In such situations, probabilists attempt to pin down the asymptotic behavior of the random variables as the problem size increases. The theory of subadditive sequences constitutes one of the most powerful tools for understanding mean behavior.

A sequence $\{a_n\}_{n \geq 1}$ is said to be subadditive if

$$a_{m+n} \quad \leq \quad a_m + a_n \tag{5.9}$$

for all positive integers m and n [53, 130, 186]. If the opposite inequality

$$a_{m+n} \quad \geq \quad a_m + a_n$$

holds, then the sequence is superadditive. Subadditive and superadditive sequences arise in many combinatorial optimization problems. For example, let X_n denote the minimum effort it takes to solve a random problem of size n. Now suppose that we can decompose a problem of size $m+n$ into two subproblems of size m and n and patch together optimal solutions of these to derive a suboptimal solution of the problem of size $m+n$. If the effort

for the concatenated solution is the sum of the efforts for the subsolutions, then the minimal efforts satisfy

$$X_{m+n} \quad \leq \quad X_m + X_n. \tag{5.10}$$

In other words, the random sequence X_n is subadditive. Taking expectations in inequality (5.10) demonstrates that the sequence $\mathrm{E}(X_n)$ is also subadditive, provided the expectations exist.

We now prove the remarkable fact that inequality (5.9) implies that

$$\lim_{n\to\infty} \frac{a_n}{n} \quad = \quad \inf_n \frac{a_n}{n} \quad = \quad \gamma. \tag{5.11}$$

The possibility $\gamma = -\infty$ is not ruled out. Consider first the case $\gamma > -\infty$, and set $a_0 = 0$. For any $\epsilon > 0$, we can find a k such that $a_k \leq (\gamma + \epsilon)k$. Because any $m > 0$ can be written as $m = nk + j$ with $0 \leq j < k$, it follows that

$$a_m \quad = \quad a_{nk+j} \quad \leq \quad na_k + a_j \quad \leq \quad (\gamma + \epsilon)nk + \max_{0\leq l<k} a_l$$

and consequently that

$$\limsup_m \frac{a_m}{m} \quad \leq \quad \gamma + \epsilon \quad \leq \quad \liminf_m \frac{a_m}{m} + \epsilon.$$

By virtue of the arbitrariness of ϵ, this shows that the limit (5.11) exists. The easier case of $\gamma = -\infty$ is left to the reader. A similar result holds for a superadditive sequence.

Example 5.7.1 *Longest Common Subsequence*

A string is a finite sequence of letters taken from some alphabet. For instance in DNA sequence analysis, the relevant alphabet consists of the four letters A, C, T, and G. Two DNA strings sharing an evolutionary history will have long subsequences in common. If we represent two strings of length n by $u_1, \ldots, u_n$ and $v_1, \ldots, v_n$, then the subsequences $u_{i_1}, u_{i_2}, \ldots, u_{i_m}$ and $v_{j_1}, v_{j_2}, \ldots, v_{j_m}$ are shared provided $u_{i_k} = v_{j_k}$ for $1 \leq k \leq m$. Now consider two random strings whose letters are independently and identically distributed. It is important to characterize the random length $M_n = m$ of the longest common subsequence.

Considerable effort has gone into finding $\mathrm{E}(M_n)$. We now show that $\mathrm{E}(M_n) \asymp \gamma n$ for large n and a constant $\gamma \in [0, 1]$ depending on the letter distribution imposed on the alphabet [186]. This follows readily from the superadditivity property $M_{r+s} \geq M_r + M_s^*$ derived by concatenating a longest common subsequence drawn from the first block of r pairs of letters with a longest common subsequence drawn from the last block of s pairs of letters. Since M_s and M_s^* have the same distribution, the mean inequality $\mathrm{E}(M_{r+s}) \geq \mathrm{E}(M_r) + \mathrm{E}(M_s)$. Unfortunately, this argument fails to identify

the constant γ. This difficulty plagues all applications of subadditivity and superadditivity. Further problem-specific information must be brought to bear to find γ [186]. For example, Problem 20 provides bounds on γ for the problem of calculating self-avoidance probabilities in a symmetric random walk. ■

Example 5.7.2 *Euclidean Traveling Salesman Problem*

The average complexity of many combinatorial optimization problems grows at a slower than linear rate, thus falling outside the domain of application of subadditivity. A probabilistic version of the traveling salesman problem furnishes a case in point. In the classical version of the traveling salesman problem, the salesman must visit n towns, starting and ending in his hometown. To minimize his travel time and expense, the salesman takes an optimal route. We defer to Example 7.8.1 the question of how to find such a route.

In the Euclidean, probabilistic version of the problem, n points (sites) $Y_1, \ldots, Y_n$ are drawn uniformly and independently from the unit square [186]. The shortest circuit that the salesman can make through the points is given by the random variable

$$D_n = \min_{\sigma} \sum_{i=1}^{n} \|Y_{\sigma(i)} - Y_{\sigma(i+1)}\|,$$

where σ denotes a generic permutation of $\{1, \ldots, n\}$, $\sigma(n+1) = \sigma(1)$, and $\|\cdot\|$ is the Euclidean norm. We now demonstrate that the average distance $\mathrm{E}(D_n)$ the salesman travels is roughly proportional to $\sqrt{n}$.

One obvious upper bound on $\mathrm{E}(D_n)$ is furnished by $M_n = \sup D_n$. It is difficult to calculate M_n, so we will be content with bounding it. We can attack this easier problem by choosing $m = \max\{k \geq 1 : k^2 < n\}$ and dividing the unit square into m^2 nonoverlapping subsquares having sides of length $1/m$. Any two points within one of these subsquares are separated by a distance of at most $\sqrt{2}/m$. Furthermore, $\sqrt{2}/m \leq 2/\sqrt{n}$, as Problem 21 asks the reader to check. Because $m^2 < n$, the pigeonhole principle requires that two of the points, say Y_j and Y_k, fall in the same subsquare. Now consider a minimum-length tour of the $n-1$ points excluding Y_k. If $Y_i \to Y_j$ in this tour, then we can extend the tour to a tour of all n points by replacing the path $Y_i \to Y_j$ by the two paths $Y_i \to Y_k$ and $Y_k \to Y_j$. In view of the triangle inequality

$$\begin{aligned} ||Y_i - Y_k|| + ||Y_k - Y_j|| &\leq ||Y_i - Y_j|| + 2||Y_k - Y_j|| \qquad (5.12) \\ &\leq ||Y_i - Y_j|| + \frac{4}{\sqrt{n}}, \end{aligned}$$

the bounds

$$D_n \leq D_{n-1} + \frac{4}{\sqrt{n}}$$

and

$$M_n \leq M_{n-1} + \frac{4}{\sqrt{n}}$$

hold. If we iterate the last bound and employ $M_0 = 0$, then it is clear that

$$M_n \leq \sum_{i=1}^{n} \frac{4}{\sqrt{i}} \leq 4 \int_0^n \frac{1}{\sqrt{x}}\, dx = 8\sqrt{n}$$

and therefore that $\mathrm{E}(D_n) \leq 8\sqrt{n}$.

One can supplement this upper bound with a lower bound of the same order of magnitude. In this case we begin with

$$\begin{aligned} \mathrm{E}(D_n) &\geq \sum_{i=1}^{n} \mathrm{E}\left(\min_{j \neq i} ||Y_j - Y_i|| \right) \\ &= n\, \mathrm{E}\left(\min_{j \neq n} ||Y_j - Y_n|| \right) \\ &= n\, \mathrm{E}\left[\mathrm{E}\left(\min_{j \neq n} ||Y_j - y|| \mid Y_n = y \right) \right]. \end{aligned}$$

To calculate the conditional expectation

$$\mathrm{E}\left(\min_{j \neq n} ||Y_j - y|| \mid Y_n = y \right) = \mathrm{E}\left(\min_{j \neq n} ||Y_j - y|| \right),$$

we use the right-tail probability bound

$$\Pr\left(\min_{j \neq n} ||Y_j - y|| \geq r \right) \geq (1 - \pi r^2)^{n-1}$$

valid for any y in the unit square. Thus, Example 2.5.1 implies that

$$\begin{aligned} \mathrm{E}(\min_{j \neq n} ||Y_j - y||) &\geq \int_0^{\frac{1}{\sqrt{\pi}}} (1 - \pi r^2)^{n-1} dr \\ &\approx \int_0^{\frac{1}{\sqrt{\pi}}} e^{-\pi(n-1)r^2} dr \\ &\approx \frac{1}{2} \int_{-\infty}^{\infty} e^{-\pi(n-1)r^2} dr \\ &= \frac{1}{2\sqrt{n-1}}. \end{aligned}$$

In summary, we conclude that $\mathrm{E}(D_n)/\sqrt{n} \geq \sqrt{n}/(2\sqrt{n-1}) \approx 1/2$ to a good approximation for large n. A combination of further theoretical work and numerical experimentation [186] suggests that $\lim_{n \to \infty} D_n/\sqrt{n} \to \gamma$ for some constant $\gamma \in [0.70, 0.73]$. ■

5.8 Problems

1. What is the probability that a random permutation of n distinct numbers contains at least one preexisting splitter? What are the mean and variance of the number of preexisting splitters?

2. Show that the worst case of quick sort takes on the order of n^2 operations.

3. Consider the problem of finding an order statistic $x_{(k)}$ from an unsorted array $\{x_1, \ldots, x_n\}$ of n distinct numbers. This can be accomplished in $O(n)$ operations based on the quick sort strategy. After the initial partitioning step, one can tell which of the two subarrays contains $x_{(k)}$ just by looking at their sizes. If the left array has $k-1$ entries, then the splitting value is $x_{(k)}$. If the left array has k or more entries, then it contains $x_{(k)}$. Otherwise, the right array contains $x_{(k)}$. Now let T_{nk} denote the expected number of operations to find $x_{(k)}$, and put $T_n = \max_k T_{nk}$. One can prove that $T_n \leq 4n$. In view of the fact that it takes $n-1$ comparisons to create the left and right subarrays, show that

$$T_n \quad \leq \quad n - 1 + \frac{2}{n} \sum_{k=\lfloor \frac{n}{2} \rfloor}^{n-1} T_k,$$

   and argue by induction that $T_n \leq 4n$.

4. Consider the uniform distribution $p_l = n^{-1}$ on an alphabet $\mathcal{A}$ with n letters. Let $\text{len}(s_l)$ be the number of bits in the bit string s_l representing l under Huffman coding. If $m = \max\{\text{len}(s_l) : l \in \mathcal{A}\}$, then show that $\text{len}(s_l) = m$ or $m-1$ for all l. If $n = \alpha 2^k$ for $1 < \alpha \leq 2$, then determine the number of letters l with $\text{len}(s_l) = j$ for $j = m-1$ and $j = m$. Use these numbers to calculate $\text{E}[\text{len}(H)]$, where H is a random Huffman bit string.

5. A sequence $X_1, \ldots, X_n$ of independent random variables uniformly distributed over the set $S_n = \{1, 2, \ldots, n\}$ defines a random function from S_n into itself. Prove that the number $F_n = \sum_{j=1}^n 1_{\{X_j = j\}}$ of fixed points satisfies $\text{E}(F_n) = 1$, $\text{Var}(F_n) = 1 - \frac{1}{n}$, and

$$\text{E}\left(e^{itF_n}\right) \quad = \quad \left(1 - \frac{1}{n} + \frac{1}{n} e^{it}\right)^n.$$

   Use the last identity in conjunction with Proposition 12.6.1 to show that F_n is approximately Poisson distributed with mean 1.

6. Read Example 4.2.4 on Fibonacci numbers. Consider a probability model whose sample space is the collection of tilings of a checkerboard

row by square pieces and dominoes. If we assign equal probability to each of the f_n possible tilings of a row of length n, then the most pertinent random variable is the number of dominoes D_n in a random tiling. Prove that

$$\begin{aligned} \mathrm{E}(D_n) &= \frac{1}{f_n}\sum_{i=1}^{n-1} f_{i-1}f_{n-i-1} \\ &= \frac{1}{f_n}\sum_j f_j f_{n-2-j} \end{aligned}$$

using the conventions $f_0 = 1$ and $f_i = 0$ for $i < 0$ and the representation

$$D_n = \sum_{i=1}^{n-1} C_i,$$

where C_i is the indicator of the event that a domino occupies squares i and $i+1$. Similarly prove that

$$\mathrm{Var}(D_n) = \frac{2}{f_n}\sum_j\sum_k f_j f_k f_{n-4-j-k} + \mathrm{E}(D_n) - \mathrm{E}(D_n)^2.$$

How can you use $\mathrm{E}(D_n)$ and $\mathrm{Var}(D_n)$ to calculate the mean and variance of the number of square pieces used?

7. Continuing Problem 6, show that the Fibonacci sequence has generating function

$$F(s) = \sum_{n=0}^{\infty} f_n x^n = \frac{1}{1-x-x^2}.$$

Use this to prove the convolution identity

$$\sum_{k=0}^{n} f_k f_{n-k} = \frac{(n+1)f_{n+1} + 2(n+2)f_n}{5}$$

leading to the simplification

$$\mathrm{E}(D_n) = \frac{(n-1)f_{n-1} + 2nf_{n-2}}{5f_n}.$$

8. Consider a probability model under which all partitions of a set with n elements are equally likely. Let X_n be the number of blocks in a random partition of the n-set. Show that

$$\begin{aligned} \mathrm{E}(X_n) &= \frac{B_{n+1}}{B_n} - 1 \\ \mathrm{Var}(X_n) &= \frac{B_{n+2}}{B_n} - \left(\frac{B_{n+1}}{B_n}\right)^2 - 1 \end{aligned}$$

using the Bell numbers. (Hint: Review Example 4.2.3 and apply recurrence relation (4.18).)

9. Let Y_n be the number of cycles in a random permutation. Demonstrate that

$$\begin{aligned}\text{Var}(Y_n) &= \sum_{k=1}^{n}\frac{1}{k}-\sum_{k=1}^{n}\frac{1}{k^2}\\ &\approx \ln n+\gamma-\frac{\pi^2}{6}.\end{aligned}$$

(Hint: Equation (4.24) provides the generating function of Y_n.)

10. Euler's combinatorial number e_{nk} denotes the number of permutations π of $\{1,\ldots,n\}$ with k ascents $\pi(j)<\pi(j+1)$. Prove that

$$\begin{aligned}e_{nk} &= e_{n,n-1-k}\\ e_{nk} &= (k+1)e_{n-1,k}+(n-k)e_{n-1,k-1}.\end{aligned}$$

11. Continuing Problem 10, let A_n be the number of ascents in a random permutation of $\{1,\ldots,n\}$. Show that

$$\begin{aligned}\text{E}(A_n) &= \frac{n-1}{2}\\ \text{Var}(A_n) &= \frac{n-1}{4}+2(n-2)\Big(\frac{1}{6}-\frac{1}{2^2}\Big)1_{\{n\geq 2\}}\\ &= \begin{cases}\frac{n+1}{12} & n\geq 2\\ 0 & n=1\,.\end{cases}\end{aligned}$$

(Hint: Write A_n as a sum of indicator random variables.)

12. Consider the set of $n\times n$ matrices M whose entries are drawn independently and uniformly from the set $\{-1,1\}$. Thus, each such matrix has probability 2^{-n^2}. Show that $\text{E}(\det M)=0$ and $\text{Var}(\det M)=n!$. It follows that some matrix exists in the set with $|\det M|\geq\sqrt{n!}$. The maximum value of $|\det M|$ is unknown [205]. (Hint: Express $\det M=\sum_\pi \text{sgn}(\pi)m_{1\pi(1)}\cdots m_{n\pi(n)}$ as a sum over all permutations π of $\{1,\ldots,n\}$.)

13. Check that any five noncollinear points in the plane R^2 determine at least one obtuse angle.

14. Let $\|x\|$ be the standard Euclidean norm of a vector $x\in\mathsf{R}^n$. A vector x with $\|x\|=1$ is said to be a unit vector. For any sequence $x_1,\ldots,x_n$ of n unit vectors from R^n, it is possible to find n numbers $\epsilon_1,\ldots,\epsilon_n$ drawn from $\{-1,1\}$ such that $\|\epsilon_1x_1+\cdots+\epsilon_nx_n\|\leq\sqrt{n}$. A different choice of $\epsilon_1,\ldots,\epsilon_n$ yields the reverse inequality [4]. Prove this striking

result by setting up a simple probability model. (Hints: Choose the ϵ_i independently from $\{-1, 1\}$ in such a way that $\mathrm{E}(\epsilon_i) = 0$. Now show that the random variable $X = \|\epsilon_1 x_1 + \cdots + \epsilon_n x_n\|^2$ has expectation $\mathrm{E}(X) = n$.)

15. Exactly 10% of the surface of a sphere in R^3 is colored black, and 90% is colored white. Show that it is possible to inscribe a cube in the sphere with all of its vertices colored white [80].

16. Consider a graph with m nodes and n edges. For any set of nodes S, let $X(S)$ be the number of edges with exactly one endpoint in S. Show that $\max_S X(S) \geq n/2$. (Hints: Generate S randomly by independently sampling each node with probability $1/2$. Decompose X as a sum of indicators indexed by the edges.)

17. Consider a family of subsets $\mathcal{F}$ of a set S with the property that each $A \in \mathcal{F}$ has exactly d elements. The family $\mathcal{F}$ is said to be 2-colorable if we can assign one of two colors, say black and white, to each element of S in such a manner that each $A \in \mathcal{F}$ possesses at least one element of each color. Prove that $\mathcal{F}$ is 2-colorable if it contains fewer than 2^{d-1} subsets [1]. (Hints: Randomly color each element of S with one of the two equally likely colors black and white. Let C_A be the event that all elements of $A \in \mathcal{F}$ receive the same color. Show that $\cup_A C_A$ has probability strictly less than 1.)

18. Let $f(t)$ be a nonnegative function on $(0, \infty)$ with $\lim_{t \to 0} f(t) = 0$. If $f(t)$ is subadditive in the sense that $f(s + t) \leq f(s) + f(t)$ for all positive s and t, then show that

$$\lim_{t \downarrow 0} \frac{f(t)}{t} \quad = \quad \sup_{t > 0} \frac{f(t)}{t} \quad = \quad q.$$

(Hint: Demonstrate that

$$p \quad \leq \quad \liminf_{t \downarrow 0} \frac{f(t)}{t} \quad \leq \quad \limsup_{t \downarrow 0} \frac{f(t)}{t}$$

for all $p \in [0, q)$.)

19. A random walk X_n on the integer lattice in R^m is determined by the transition probabilities $\Pr(X_{n+1} = j \mid X_n = i) = q_{j-i}$ and the initial value $X_0 = \mathbf{0}$. Let p_n be the probability that $X_i \neq X_j$ for all pairs $0 \leq i < j \leq n$. In other words, p_n is the probability that the walk avoids itself during its first n steps. Prove that either $p_n = 0$ for all sufficiently large n or $\lim_{n \to \infty} \frac{1}{n} \ln p_n = \gamma$ for some $\gamma \leq 0$. (Hint: Argue that $p_{m+n} \leq p_m p_n$.)

20. Continuing Problem 19, suppose the random walk is symmetric in the sense that $q_i = 1/(2m)$ if and only if $\|i\| = 1$. Prove that

$$\frac{m^n}{(2m)^n} \;\le\; p_n \;\le\; \frac{2m(2m-1)^{n-1}}{(2m)^n}.$$

Use these inequalities to prove that the constant γ of Problem 19 satisfies $\ln(1/2) \le \gamma \le \ln[1 - 1/(2m)]$. (Hints: A random walk is self-avoiding if all its steps are in the positive direction. After its first step, a self-avoiding walk cannot move from its current position back to its previous position.)

21. Suppose $m = \max\{k \ge 1 : k^2 < n\}$. Prove that $\sqrt{2}/m \le 2/\sqrt{n}$ for all sufficiently large n.

6
Poisson Processes

6.1 Introduction

The Poisson distribution rivals the normal distribution in importance. It occupies this position of eminence because of its connection to Poisson processes [59, 60, 80, 96, 106, 114, 170]. A Poisson process models the formation of random points in space or time. Most textbook treatments of Poisson processes stress one-dimensional processes. This is unfortunate because many of the important applications occur in higher dimensions, and the underlying theory is about as simple there. In this chapter, we emphasize multidimensional Poisson processes, their transformation properties, and computational tools for extracting information about them.

The number of applications of Poisson processes is truly amazing. To give just a few examples, physicists use them to describe the emission of radioactive particles, astronomers to account for the distribution of stars, communication engineers to model the arrival times of telephone calls at an exchange, radiologists to reconstruct medical images in emission and transmission tomography, and ecologists to test for the random location of plants. Almost equally important, Poisson processes can provide a theoretical perspective helpful in complex probability calculations that have no obvious connection to random points. We will visit a few applications of both types as we proceed.

K. Lange, *Applied Probability,*
DOI 10.1007/978-1-4419-7165-4_6, © Springer Science+Business Media, LLC 2010

6.2 The Poisson Distribution

It is helpful to begin our exposition of Poisson processes with a brief review of the Poisson distribution. Readers will recall that a Poisson random variable $Z \geq 0$ has discrete density $\Pr(Z = k) = e^{-\mu}\frac{\mu^k}{k!}$ with mean and variance μ and probability generating function $\mathrm{E}(t^Z) = e^{\mu(t-1)}$. Furthermore, if $Z_1, \ldots, Z_m$ are independent Poisson random variables, then the sum $Z = \sum_{k=1}^m Z_k$ is also a Poisson random variable. Less well known is the next proposition.

Proposition 6.2.1 *Suppose a Poisson random variable Z with mean μ represents the number of outcomes from some experiment. Let each outcome be independently classified in one of m categories, the kth of which occurs with probability p_k. Then the number of outcomes Z_k falling in category k is Poisson distributed with mean $\mu_k = p_k\mu$. Furthermore, the random variables $Z_1, \ldots, Z_m$ are independent. Conversely, if $Z = \sum_{k=1}^m Z_k$ is a sum of independent Poisson random variables Z_k with means $\mu_k = p_k\mu$, then conditional on $Z = n$, the vector $(Z_1, \ldots, Z_m)$ follows a multinomial distribution with n trials and cell probabilities $p_1, \ldots, p_m$.*

Proof: If $n = n_1 + \cdots + n_m$, then

$$\begin{aligned}
\Pr(Z_1 = n_1, \ldots, Z_m = n_m) &= e^{-\mu}\frac{\mu^n}{n!}\binom{n}{n_1, \ldots, n_m}\prod_{k=1}^m p_k^{n_k} \\
&= \prod_{k=1}^m e^{-\mu_k}\frac{\mu_k^{n_k}}{n_k!} \\
&= \prod_{k=1}^m \Pr(Z_k = n_k).
\end{aligned}$$

To prove the converse, divide the last string of equalities by the probability $\Pr(Z = n) = e^{-\mu}\frac{\mu^n}{n!}$. ■

In practice, it is useful to extend the definition of a Poisson random variable to include the limiting cases $X \equiv 0$ and $X \equiv \infty$ with corresponding means 0 and ∞.

6.3 Characterization and Construction

A Poisson process involves points randomly scattered in some measurable region S of m-dimensional space R^m. To formalize the notion that the points are completely random but concentrated on average more in some regions rather than in others, we introduce four postulates involving an intensity function $\lambda(x) \geq 0$ on S. Postulate (d) in the following list uses the notation $o(\mu)$ to signify a generic error term satisfying $\lim_{\mu\to 0}\frac{o(\mu)}{\mu} = 0$.

(a) There exists a sequence of disjoint subregions S_n satisfying $S = \bigcup_n S_n$ and $\int_{S_n} \lambda(x)\,dx < \infty$ for all n.

(b) The probability $p_k(\mu)$ that a region $T \subset S$ contains k random points depends only on the mass $\mu = \int_T \lambda(x)\,dx$ of T. If $\mu = \infty$, then all $p_k(\mu) = 0$, and the given region possesses an infinite number of points.

(c) The numbers of random points in disjoint regions are independent.

(d) The first two probabilities $p_0(\mu)$ and $p_1(\mu)$ have the asymptotic values

$$\begin{aligned} p_0(\mu) &= 1 - \mu + o(\mu) \\ p_1(\mu) &= \mu + o(\mu) \end{aligned} \tag{6.1}$$

as μ tends to 0. Thus, $p_k(\mu) = o(\mu)$ for all $k > 1$.

Proposition 6.3.1 *Based on postulates (a) through (d), the number of random points in a region with mass μ has the Poisson distribution*

$$p_k(\mu) = e^{-\mu}\frac{\mu^k}{k!}.$$

Proof: We first remark that for any region T with $\int_T \lambda(x)\,dx < \infty$ and any μ in the interval $[0, \int_T \lambda(x)\,dx]$, there exists a region $R \subset T$ with mass $\mu = \int_R \lambda(x)\,dx$. To verify this assertion, we need only consider regions of the form $R = T \cap B_t$, where B_t is the ball $\{x \in \mathsf{R}^m : \|x\| \leq t\}$. The function $t \mapsto \int_{T\cap B_t} \lambda(x)\,dx$ is continuous from the right by the dominated convergence theorem. It is continuous from the left because the surface of B_t has volume 0 and therefore $\int_{\{x:\|x\|=t\}} \lambda(x)\,dx = 0$. The intermediate value theorem now gives the desired conclusion.

Let μ and $d\mu$ represent the masses of two nonoverlapping regions. By the preceding comments, $d\mu$ can be made as small as we please. The equality

$$\begin{aligned} p_0(\mu + d\mu) &= p_0(\mu)p_0(d\mu) \\ &= p_0(\mu)[1 - d\mu + o(d\mu)] \end{aligned}$$

follows immediately from the postulates and can be rearranged to give the difference quotient

$$\frac{p_0(\mu + d\mu) - p_0(\mu)}{d\mu} = -p_0(\mu) + o(1).$$

Taking limits as $d\mu$ tends to 0 produces the ordinary differential equation $p_0'(\mu) = -p_0(\mu)$ with solution $p_0(\mu) = e^{-\mu}$ satisfying the initial condition $p_0(0) = 1$.

For $k \geq 1$, we again invoke the postulates and execute the expansion

$$\begin{aligned} p_k(\mu + d\mu) &= p_k(\mu)p_0(d\mu) + p_{k-1}(\mu)p_1(d\mu) + \sum_{j=2}^{k} p_{k-j}(\mu)p_j(d\mu) \\ &= p_k(\mu)[1 - d\mu + o(d\mu)] + p_{k-1}(\mu)[d\mu + o(d\mu)] \\ &\quad + \sum_{j=2}^{k} p_{k-j}(\mu)o(d\mu). \end{aligned}$$

Rearrangement of this approximation yields the difference quotient

$$\frac{p_k(\mu + d\mu) - p_k(\mu)}{d\mu} = -p_k(\mu) + p_{k-1}(\mu) + o(1)$$

and ultimately the ordinary differential equation $p_k'(\mu) = -p_k(\mu) + p_{k-1}(\mu)$ with initial condition $p_k(0) = 0$. The transformed function $q_k(\mu) = p_k(\mu)e^{\mu}$ satisfies the simpler ordinary differential equation $q_k'(\mu) = q_{k-1}(\mu)$ with initial condition $q_k(0) = 0$. If the $p_k(\mu)$ are Poisson, then $q_k(\mu) = \frac{\mu^k}{k!}$ should hold. This formula is certainly true for $q_0(\mu) = 1$. Assuming that it is true for $q_{k-1}(\mu)$, we see that

$$\begin{aligned} q_k(\mu) &= \int_0^{\mu} q_{k-1}(u)\, du \\ &= \int_0^{\mu} \frac{u^{k-1}}{(k-1)!}\, du \\ &= \frac{\mu^k}{k!} \end{aligned}$$

has the necessary value to advance the inductive argument and complete the proof. ■

At this junction, several remarks are in order. First, the proposition is less than perfectly rigorous because we have only considered derivatives from the right. A better proof under less restrictive conditions is given in reference [114]. Second, $\mu = \int_T \lambda(x)\, dx$ is the expected number of random points in the region T. Third, only a finite number of random points can occur in any S_n. Fourth, if $\int_T \lambda(x)\, dx = \infty$, then an infinite number of random points occur in T. Fifth, because every $y \in S$ has mass $\int_{\{y\}} \lambda(x)\, dx = 0$, the probability that a random point coincides with y is 0. Sixth, no two random points ever coincide. Seventh, the approximations (6.1) are consistent with the final result $p_k(\mu) = e^{-\mu}\frac{\mu^k}{k!}$. Eighth and finally, if the intensity function $\lambda(x)$ is constant, then the Poisson process is said to be homogeneous.

We now turn the question of Proposition 6.3.1 around and ask how one can construct a Poisson process with a given intensity function $\lambda(x)$. This

question has more than theoretical interest because we often need to simulate a Poisson process on a computer. Briefly, we attack the problem by independently generating random points in each of the disjoint regions S_n. The number of points N_{S_n} to be scattered in S_n follows a Poisson distribution with mean $\mu_n = \int_{S_n} \lambda(x)\, dx$. Once we sample N_{S_n}, then we independently distribute the corresponding N_{S_n} points X_{ni} one by one over the region S_n according to the probability measure

$$\Pr(X_{ni} \in R) \quad = \quad \frac{1}{\mu_n} \int_R \lambda(x)\, dx.$$

This procedure incorporates the content of Proposition 6.2.1. The resulting union of random points $\Pi = \bigcup_n \bigcup_{i=1}^{N_{S_n}} \{X_{ni}\}$ constitutes one realization from the required Poisson process.

6.4 One-Dimensional Processes

When a Poisson process occurs on a subset of the real line, it is often convenient to refer to time instead of space and events instead of points. Consider a homogeneous Poisson process on $[0, \infty)$ with intensity λ. Let T_k be the waiting time until the kth event after time 0. The interarrival time between events $k-1$ and k equals $W_k = T_k - T_{k-1}$ for $k > 1$. By convention $W_1 = T_1$.

Proposition 6.4.1 *The random waiting time T_k has a gamma distribution with density $\lambda \frac{(\lambda t)^{k-1}}{(k-1)!} e^{-\lambda t}$. Furthermore, the interarrival times W_k are independent and exponentially distributed with intensity λ.*

Proof: The event $T_k > t$ is equivalent to the event that $k-1$ or fewer random points fall in $[0, t]$. Hence,

$$\Pr(T_k \leq t) \quad = \quad 1 - \Pr(T_k > t) \quad = \quad 1 - \sum_{j=0}^{k-1} e^{-\lambda t} \frac{(\lambda t)^j}{j!}.$$

Differentiating this distribution function with respect to t gives the density function

$$-\sum_{j=0}^{k-1} j\lambda e^{-\lambda t} \frac{(\lambda t)^{j-1}}{j!} + \sum_{j=0}^{k-1} \lambda e^{-\lambda t} \frac{(\lambda t)^j}{j!} \quad = \quad \lambda \frac{(\lambda t)^{k-1}}{(k-1)!} e^{-\lambda t}.$$

This proves the first claim.

Assume for the moment that the second claim is true. Because the matrix of the linear transformation

$$T_1 \quad = \quad W_1$$

$$\begin{aligned} T_2 &= W_1 + W_2 \\ &\vdots \\ T_n &= W_1 + W_2 + \cdots + W_n \end{aligned}$$

taking $(W_1, \ldots, W_n)$ to $(T_1, \ldots, T_n)$ is lower triangular with 1's down its diagonal, it has Jacobian 1. The change of variables formula (1.12) therefore implies that the random vector $(T_1, \ldots, T_n)$ has density

$$f_n(t_1, \ldots, t_n) = \prod_{i=1}^{n} \lambda e^{-\lambda w_i} = \lambda^n e^{-\lambda t_n}$$

on the region $\Gamma_n = \{0 \leq t_1 \leq \cdots \leq t_n\}$. Conversely, if $(T_1, \ldots, T_n)$ possesses the density $f_n(t_1, \ldots, t_n)$, then applying the inverse transformation shows that the W_k are independent and exponentially distributed with intensity λ.

Now let $F_n(t_1, \ldots, t_n)$ be the distribution function corresponding to $f_n(t_1, \ldots, t_n)$. Integrating the obvious identity

$$\begin{aligned} f_n(s_1, \ldots, s_n) &= \lambda^n e^{-\lambda s_n} \\ &= \lambda e^{-\lambda s_1} f_{n-1}(s_2 - s_1, \ldots, s_n - s_1) \end{aligned}$$

over the intersection $\Gamma_n \cap \{s_1 \leq t_1, \ldots, s_n \leq t_n\}$ yields the identity

$$F_n(t_1, \ldots, t_n) = \int_0^{t_1} \lambda e^{-\lambda s_1} F_{n-1}(t_2 - s_1, \ldots, t_n - s_1)\, ds_1 \quad (6.2)$$

recursively determining the distribution functions $F_n(t_1, \ldots, t_n)$ starting with $F_1(t_1) = 1 - e^{-\lambda t_1}$. If $G_n(t_1, \ldots, t_n)$ denotes the actual distribution function of $(T_1, \ldots, T_n)$, then our strategy is to show that $G_n(t_1, \ldots, t_n)$ satisfies identity (6.2). Given the fact that $G_1(t_1) = 1 - e^{-\lambda t_1}$, induction on n then shows that $G_n(t_1, \ldots, t_n)$ and $F_n(t_1, \ldots, t_n)$ coincide.

To verify that $G_n(t_1, \ldots, t_n)$ satisfies identity (6.2), we first note that

$$\Pr(T_1 \leq t_1, \ldots, T_n \leq t_n) = \Pr(\cap_{i=1}^{n} \{N_{t_i} \geq i\}). \quad (6.3)$$

We also note the identity

$$\begin{aligned} \frac{(\lambda t)^j}{j!} e^{-\lambda t} &= \frac{\lambda^j}{(j-1)!} e^{-\lambda t} \int_0^t s^{j-1} ds \\ &= \frac{\lambda^j}{(j-1)!} e^{-\lambda t} \int_0^t (t-s)^{j-1} ds \\ &= \int_0^t \lambda e^{-\lambda s} \frac{[\lambda(t-s)]^{j-1}}{(j-1)!} e^{-\lambda(t-s)} ds. \end{aligned}$$

Consequently, if $1 \leq j_1 \leq \cdots \leq j_n$, then

$$\begin{aligned}
\Pr(\cap_{i=1}^n \{N_{t_i} = j_i\}) &= \Pr(N_{t_1} = j_1) \Pr(\cap_{i=2}^n \{N_{t_i} - N_{t_{i-1}} = j_i - j_{i-1}\}) \\
&= \int_0^{t_1} \lambda e^{-\lambda s} \frac{[\lambda(t_1 - s)]^{j_1 - 1}}{(j_1 - 1)!} e^{-\lambda(t_1 - s)} \\
&\quad \times \Pr(\cap_{i=2}^n \{N_{t_i} - N_{t_{i-1}} = j_i - j_{i-1}\})\, ds \\
&= \int_0^{t_1} \lambda e^{-\lambda s} \frac{[\lambda(t_1 - s)]^{j_1 - 1}}{(j_1 - 1)!} e^{-\lambda(t_1 - s)} \\
&\quad \times \Pr(\cap_{i=2}^n \{N_{t_i - s} - N_{t_{i-1} - s} = j_i - j_{i-1}\})\, ds \\
&= \int_0^{t_1} \lambda e^{-\lambda s} \Pr(\cap_{i=1}^n \{N_{t_i - s} = j_i - 1\})\, ds.
\end{aligned}$$

Summing this equality over the intersection of the sets $\{j_1 \leq \cdots \leq j_n\}$ and $\{j_1 \geq 1, \ldots, j_n \geq n\}$ produces

$$\Pr(\cap_{i=1}^n \{N_{t_i} \geq i\}) = \int_0^{t_1} \lambda e^{-\lambda s} \Pr(\cap_{i=2}^n \{N_{t_i - s} \geq i - 1\})\, ds \quad (6.4)$$

because the event $N_{t_1} \geq 0$ is certain. Taking into account representation (6.3), identity (6.4) is just a disguised form of identity (6.2). ■

This proposition implies that generating a sequence of exponentially distributed interarrival times $W_1, W_2, \ldots$ and extracting the corresponding waiting times $T_k = \sum_{j=1}^k W_j$ from it provides another method of constructing a homogeneous Poisson process on $[0, \infty)$.

The exponential distribution has an important "lack of memory" property. If X is exponentially distributed with intensity λ, then

$$\begin{aligned}
\Pr(X > t + h \mid X > t) &= \frac{\Pr(X > t + h)}{\Pr(X > t)} \\
&= \frac{e^{-\lambda(t+h)}}{e^{-\lambda t}} \\
&= e^{-\lambda h} \\
&= \Pr(X > h).
\end{aligned}$$

Lack of memory characterizes the exponential.

Proposition 6.4.2 *Suppose X is a random variable with values in $(0, \infty)$ and satisfying $\Pr(X > t + h) = \Pr(X > t) \Pr(X > h)$ for all positive h and t. Then X is exponentially distributed.*

Proof: If we let $g(t) = \Pr(X > t)$, then $g(0) = 1$ and $g(t)$ satisfies the familiar functional equation (2.6). Given differentiability of $g(t)$, the reasoning in Example 2.4.7 leads to the solution $g(t) = e^{-\lambda t}$. Here $\lambda > 0$ because $g(t)$ tends to 0 as t tends to ∞. ■

Example 6.4.1 *Waiting Time Paradox*

Buses arrive at a bus stop at random times according to a Poisson process with intensity λ. I arrive at time t and ask how much time $\mathrm{E}(W)$ on average I will have to wait for the next bus. To quote Feller [60], there are two mutually contradictory responses:

(a) "The lack of memory of the Poisson process implies that the distribution of my waiting time should not depend on the epoch of my arrival. In this case, $\mathrm{E}(W) = 1/\lambda$."

(b) "The epoch of my arrival is chosen at random in the interval between two consecutive buses, and for reasons of symmetry my expected waiting time should be half the expected time between two consecutive buses, that is $\mathrm{E}(W) = 1/(2\lambda)$."

Answer (a) is correct; answer (b) neglects the fact that I am more likely to arrive during a long interval than a short interval. This is the paradox of length-biased sampling. In fact, the random length of the interval capturing my arrival is distributed as the sum of two independent exponentially distributed random variables with intensity λ. This assertion is clear if I arrive at time 0, and it continues to hold for any other time t because a homogeneous Poisson process is stationary and possesses no preferred time origin. ■

Example 6.4.2 *Order Statistics from an Exponential Sample*

The lack of memory property of the exponential distribution makes possible an easy heuristic derivation of a convenient representation of the order statistics $X_{(1)} \leq \cdots \leq X_{(n)}$ from an independent sample $X_1, \ldots, X_n$ of exponentially distributed random variables with common intensity λ [60]. From the calculation $\Pr(X_{(1)} \geq x) = \prod_{j=1}^{n} \Pr(X_j \geq x) = e^{-n\lambda x}$, we find that $X_{(1)}$ is exponentially distributed with intensity $n\lambda$. Because of the lack of memory property of the exponential, the $n-1$ random points to the right of $X_{(1)}$ provide an exponentially distributed sample of size $n-1$ starting at $X_{(1)}$. Duplicating our argument for $X_{(1)}$, we find that the difference $X_{(2)} - X_{(1)}$ is independent of $X_{(1)}$ and exponentially distributed with intensity $(n-1)\lambda$. Arguing inductively, we now see that $Z_1 = X_{(1)}$ and the differences $Z_k = X_{(k)} - X_{(k-1)}$ are independent and that Z_k is exponentially distributed with intensity $(n-k+1)\lambda$. Problem 5 proves the representation $X_{(j)} = \sum_{k=1}^{j} Z_k$ rigorously by transforming the relevant probability densities; Problem 6 provides the moments of $X_{(j)}$ based on it. ■

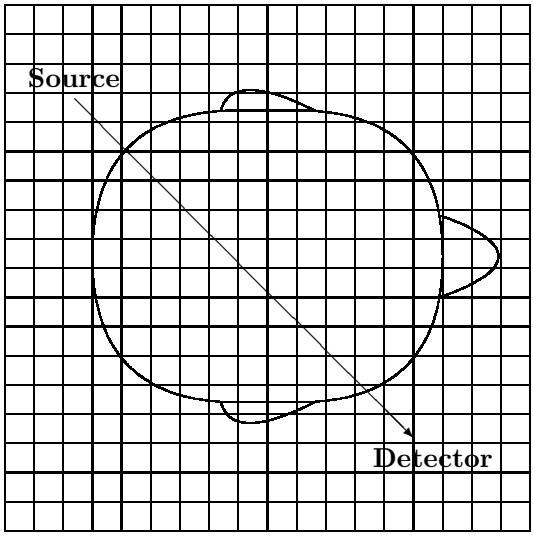


FIGURE 6.1. Cartoon of Transmission Tomography

6.5 Transmission Tomography

The purpose of transmission tomography is to reconstruct the local attenuation properties of the object being imaged. Attenuation is to be roughly equated with density. In medical applications, material such as bone is dense and stops or deflects X-rays better than soft tissue. With enough radiation, even small gradations in soft tissue can be detected. The traditional method of image reconstruction in transmission tomography relies on Fourier analysis and the Radon transform [88]. An alternative to this deterministic approach is to pose an explicitly Poisson process model that permits parameter estimation by maximum likelihood [125]. The MM algorithm presented in Chapter 3 immediately suggests itself in this context.

The stochastic model depends on dividing the object of interest into small nonoverlapping regions of constant attenuation called pixels. Typically the pixels are squares. The attenuation attributed to pixel j constitutes parameter θ_j of the model. Since there may be thousands of pixels, implementation of maximum likelihood algorithms such as scoring or Newton's method is out of the question. Each observation Y_i is generated by beaming a stream of X-rays or high-energy photons from an X-ray source toward some detector on the opposite side of the object. The observation (or projection) Y_i counts the number of photons detected along the ith line of flight. Figure 6.1 shows one such projection line beamed through a cartoon of the human head. Naturally, only a fraction of the photons are successfully transmitted from source to detector. If l_{ij} is the length of the segment of projection line i intersecting pixel j, then we claim that the

probability of a photon escaping attenuation along projection line i is the exponentiated line integral $\exp(-\sum_j l_{ij}\theta_j)$.

This result can be demonstrated by considering a Poisson process along projection i, starting with the source as origin. Each random point corresponds to a possible attenuation event. The first attenuation event encountered stops or deflects the photon and thus prevents it from being detected. The intensity of the attenuation process is determined locally by the attenuation coefficient of the surrounding pixel. It follows that a photon escapes attenuation with Poisson probability $\exp(-\sum_j l_{ij}\theta_j)$. Example 8.7.2 continues this discussion from the perspective of continuous-time Markov chains.

Of course, a second Poisson process is lurking in the background. In the absence of the intervening object, the number of photons generated and ultimately detected follows a Poisson distribution. Let the mean of this distribution be d_i for projection line i. Since Proposition 6.2.1 implies that random thinning of a Poisson random variable gives a Poisson random variable, the number Y_i is Poisson distributed with mean $d_i \exp(-\sum_j l_{ij}\theta_j)$. Owing to the Poisson nature of X-ray generation, the different projections will be independent even if collected simultaneously. This fact enables us to write the loglikelihood of the observed data $Y_i = y_i$ as the finite sum

$$L(\theta) = \sum_i \left[-d_i e^{-\sum_j l_{ij}\theta_j} - y_i \sum_j l_{ij}\theta_j + y_i \ln d_i - \ln y_i! \right]. \quad (6.5)$$

Omitting irrelevant constants, we can rewrite the loglikelihood (6.5) more succinctly as

$$L(\theta) = -\sum_i f_i(l_i^t\theta),$$

where $f_i(s) = d_i e^{-s} + y_i s$ and $l_i^t\theta = \sum_j l_{ij}\theta_j$ is the inner product of the attenuation parameter vector θ and the vector of intersection lengths l_i for projection i.

Following the lead of De Pierro [47] in emission tomography, one can devise an MM algorithm based on a convexity argument [126]. First define admixture constants

$$\alpha_{ij} = \frac{l_{ij}\theta_j^n}{l_i^t\theta^n}. \quad (6.6)$$

Since $\sum_j \alpha_{ij} = 1$ and each $f_i(s)$ is strictly convex, the inequality

$$\begin{aligned} L(\theta) &= -\sum_i f_i\left(\sum_j \alpha_{ij} \frac{\theta_j}{\theta_j^n} l_i^t\theta^n \right) \\ &\geq -\sum_i \sum_j \alpha_{ij} f_i\left(\frac{\theta_j}{\theta_j^n} l_i^t\theta^n \right) \quad (6.7) \\ &= Q(\theta \mid \theta^n) \end{aligned}$$

holds. Furthermore, equality occurs when $\theta_j = \theta_j^n$ for all j. Thus, the surrogate function $Q(\theta \mid \theta^n)$ minorizes $L(\theta)$. By construction, maximizing $Q(\theta \mid \theta^n)$ separates into a sequence of one-dimensional maximization problems, each of which can be solved approximately by one step of Newton's method as noted in Problem 13.

The images produced by maximum likelihood estimation in transmission tomography look grainy. Geman and McClure [72] recommend incorporating a Gibbs prior that enforces image smoothness. A Gibbs prior $\pi(\theta)$ can be written as

$$\ln \pi(\theta) \;=\; -\gamma \sum_{\{j,k\}\in N} w_{jk}\psi(\theta_j - \theta_k),$$

where γ and the weights w_{jk} are positive constants, N is a set of unordered pairs $\{j,k\}$ defining a neighborhood system, and $\psi(r)$ is called a potential function. For instance, if the pixels are squares, we might define the weights by $w_{jk} = 1$ for orthogonal nearest neighbors and $w_{jk} = 1/\sqrt{2}$ for diagonal nearest neighbors. The constant γ scales the overall strength assigned to the prior. To achieve a smooth image with good resolution, we maximize the log posterior function $L(\theta) + \ln \pi(\theta)$ rather than $L(\theta)$.

Choice of the potential function $\psi(r)$ is the most crucial feature of the Gibbs prior. It is convenient to assume that $\psi(r)$ is even and strictly convex. Strict convexity leads to strict concavity of the log posterior function $L(\theta) + \ln \pi(\theta)$ and permits simple modification of the MM algorithm based on the $Q(\theta \mid \theta^n)$ function defined by inequality (6.7). Many potential functions exist satisfying these conditions. One simple example is $\psi(r) = r^2$. Because this choice tends to deter the formation of boundaries, Green [79] has suggested the gentler alternative $\psi(r) = \ln[\cosh(r)]$, which grows linearly for large $|r|$ rather than quadratically.

One adverse consequence of introducing a prior is that it couples the parameters in the maximization step of the MM algorithm for finding the posterior mode. One can decouple the parameters by exploiting the convexity and evenness of the potential function $\psi(r)$ through the inequality

$$\begin{aligned}\psi(\theta_j - \theta_k) &= \psi\Big(\frac{1}{2}\Big[2\theta_j - \theta_j^n - \theta_k^n\Big] + \frac{1}{2}\Big[-2\theta_k + \theta_j^n + \theta_k^n\Big]\Big)\\ &\le \frac{1}{2}\psi(2\theta_j - \theta_j^n - \theta_k^n) + \frac{1}{2}\psi(2\theta_k - \theta_j^n - \theta_k^n),\end{aligned}$$

which is strict unless $\theta_j + \theta_k = \theta_j^n + \theta_k^n$ [47]. This inequality allows us to redefine the minorizing function as

$$\begin{aligned}Q(\theta \mid \theta^n) \;=\; & -\sum_i \sum_j \alpha_{ij} f_i\left(\frac{\theta_j}{\theta_j^n} l_i^t \theta^n\right)\\ & -\frac{\gamma}{2}\sum_{\{j,k\}\in N} w_{jk}[\psi(2\theta_j - \theta_j^n - \theta_k^n) + \psi(2\theta_k - \theta_j^n - \theta_k^n)],\end{aligned}$$

where $f_i(s) = d_i e^{-s} + y_i s$ and the admixture constants a_{ij} are given by equation (6.6). The parameters are once again separated in the M step, and maximizing $Q(\theta \mid \theta^n)$ drives the logposterior uphill and eventually leads to the posterior mode (maximum).

6.6 Mathematical Applications

Poisson processes not only offer realistic models for scientific phenomena, but they also provide devices for solving certain problems in combinatorics and probability theory [2, 15, 18, 26]. The Poisson strategies at the heart of the next two examples succeed by replacing dependent random variables by closely related independent random variables.

Example 6.6.1 *Schrödinger's Method*

Schrödinger's method is a technique for solving occupancy problems in multinomial sampling. Consider a multinomial sample with m equally likely categories and n trials. If we desire the probability of the event A_n that all m categories are occupied, then we can use an inclusion-exclusion argument. Alternatively in Schrödinger's method, we assume that the number of trials N is a Poisson random variable with mean λ. According to Proposition 6.2.1, this assumption decouples the categories in the sense that the numbers of outcomes falling in the different categories are independent Poisson random variables with common mean λ/m. In the Poisson setting, the probability of the event A that all m categories are occupied satisfies

$$\begin{aligned} e^{\lambda} \Pr(A) &= e^{\lambda}\left(1 - e^{-\frac{\lambda}{m}}\right)^m \\ &= \left(e^{\frac{\lambda}{m}} - 1\right)^m \\ &= \sum_{j=0}^{m} \binom{m}{j} (-1)^j e^{(m-j)\lambda/m} \\ &= \sum_{j=0}^{m} \binom{m}{j} (-1)^j \sum_{n=0}^{\infty} \left(1 - \frac{j}{m}\right)^n \frac{\lambda^n}{n!} \qquad (6.8) \\ &= \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} \sum_{j=0}^{m} \binom{m}{j} (-1)^j \left(1 - \frac{j}{m}\right)^n. \end{aligned}$$

On the other hand, conditioning on N produces

$$\Pr(A) = \sum_{n=0}^{\infty} \Pr(A_n) e^{-\lambda} \frac{\lambda^n}{n!}. \qquad (6.9)$$

We now multiply equation (6.9) by e^{λ} and equate the result to equation (6.8). Because the coefficients of λ^n must match, the conclusion

$$\Pr(A_n) = \sum_{j=0}^{m} \binom{m}{j} (-1)^j \left(1 - \frac{j}{m}\right)^n \tag{6.10}$$

follows immediately. This Poisson randomization technique extends to more complicated occupancy problems [15].

Despite the beauty of equation (6.10), it does not yield much insight into the size of the probability $\Pr(A_n)$. In statistical contexts, one is often interested in finding upper bounds on small probabilities. Thus, the inequality

$$\Pr(A_n) \leq 2\Pr(A) = 2\left(1 - e^{-\frac{n}{m}}\right)^m$$

with $\lambda = n$ is relevant. This bound is a special case of a more general result. Let X_i be the number of outcomes that fall in category i under multinomial sampling with n trials and N_i be the number under Poisson sampling with mean number of trials $\lambda = n$. If $f(x_1, \ldots, x_m)$ is a nonnegative function such that $\mathrm{E}[f(X_1, \ldots, X_m)]$ is monotonically increasing or decreasing in n, then Problem 18 asks the reader to prove that

$$\mathrm{E}[f(X_1, \ldots, X_m)] \leq 2\,\mathrm{E}[f(N_1, \ldots, N_m)]. \tag{6.11}$$

Because $\Pr(A_n)$ is obviously increasing in n, the bound applies in the current setting. ■

Example 6.6.2 *Poissonization in the Family Planning Model 2.3.3*

Poisson processes come into play in this model when we embed the births to the couple at the random times determined by a Poisson process on $[0, \infty)$ of unit intensity. Hence on average, n births occur during $[0, n]$ for any positive integer n. When births are classified by sex, then as suggested by Proposition 6.2.1 and discussed in more detail in the next section, male births and female births form two independent Poisson processes. Let T_s and T_d be the continuously distributed waiting times until the birth of s sons and d daughters, respectively. The waiting time until the quota of at least s sons and d daughters is reached is $T_{sd} = \max\{T_s, T_d\}$. Now independence of T_s and T_d and Example 2.5.1 entail

$$\begin{aligned}
\mathrm{E}(T_{sd}) &= \int_0^\infty [1 - \Pr(T_{sd} \leq t)]\, dt \\
&= \int_0^\infty [1 - \Pr(T_s \leq t) \Pr(T_d \leq t)]\, dt \\
&= \int_0^\infty \{1 - [1 - \Pr(T_s > t)][1 - \Pr(T_d > t)]\}\, dt
\end{aligned} \tag{6.12}$$

$$\begin{aligned} &= \int_0^\infty \Pr(T_s > t)\, dt + \int_0^\infty \Pr(T_d > t)\, dt \\ &\quad - \int_0^\infty \Pr(T_s > t) \Pr(T_d > t)\, dt. \end{aligned}$$

Proposition 6.4.1 implies that $\Pr(T_s > t) = \sum_{k=0}^{s-1} \frac{(pt)^k}{k!} e^{-pt}$ and similarly for $\Pr(T_d > t)$. Combining these facts with the identity

$$\int_0^\infty t^n e^{-rt} dt = \frac{n!}{r^{n+1}}$$

and equation (6.12) leads to the conclusion that

$$\begin{aligned} \mathrm{E}(T_{sd}) &= \int_0^\infty \sum_{k=0}^{s-1} \frac{(pt)^k}{k!} e^{-pt} dt + \int_0^\infty \sum_{l=0}^{d-1} \frac{(qt)^l}{l!} e^{-qt} dt \\ &\quad - \int_0^\infty \sum_{k=0}^{s-1} \sum_{l=0}^{d-1} \frac{p^k}{k!} \frac{q^l}{l!} t^{k+l} e^{-t} dt \qquad (6.13) \\ &= \sum_{k=0}^{s-1} \frac{p^k}{p^{k+1}} + \sum_{l=0}^{d-1} \frac{q^l}{q^{l+1}} - \sum_{k=0}^{s-1} \sum_{l=0}^{d-1} \binom{k+l}{k} p^k q^l \\ &= \frac{s}{p} + \frac{d}{q} - \sum_{k=0}^{s-1} \sum_{l=0}^{d-1} \binom{k+l}{k} p^k q^l. \end{aligned}$$

Having calculated $\mathrm{E}(T_{sd})$, we now show that $\mathrm{E}(N_{sd}) = \mathrm{E}(T_{sd})$ by considering the random sum

$$T_{sd} = \sum_{k=1}^{N_{sd}} W_k,$$

where the W_k are the independent exponential waiting times between successive births. Because $\mathrm{E}(W_1) = 1$ and N_{sd} is independent of the W_k, Example 2.4.4 implies $\mathrm{E}(T_{sd}) = \mathrm{E}(N_{sd})\,\mathrm{E}(W_1) = \mathrm{E}(N_{sd})$. Alternatively, readers can check the equality $\mathrm{E}(N_{sd}) = \mathrm{E}(T_{sd})$ by verifying that formula (6.13) for $\mathrm{E}(T_{sd})$ satisfies the same boundary conditions and the same recurrence relation as $\mathrm{E}(N_{sd})$. ■

6.7 Transformations

In this section, we informally discuss various ways of constructing new Poisson processes from old ones. (Detailed proofs of all assertions made here can be found in reference [114].) For instance, suppose Π is a Poisson process on the region $S \subset \mathsf{R}^m$. If T is a measurable subset of S, then the

random points Π_T falling in T clearly satisfy the postulates (a) through (d) of a Poisson process. The Poisson process Π_T is called the restriction of Π to T. Similarly, if Π_1 and Π_2 are two independent Poisson processes on S, then the union $\Pi = \Pi_1 \bigcup \Pi_2$ is a Poisson process called the superposition of Π_1 and Π_2. The intensity function $\lambda(x)$ of Π is the sum $\lambda_1(x) + \lambda_2(x)$ of the intensity functions of Π_1 and Π_2.

In some circumstances, one can create a new Poisson process $T(\Pi)$ from an existing Poisson process by transforming the underlying space $U \subset \mathsf{R}^m$ to a new space $V \subset \mathsf{R}^n$ via a measurable map $T : U \mapsto V$. In this paradigm, a random point $X \in \Pi$ is sent into the new random point $T(X)$. To prevent random points from piling up on one another, we must impose some restriction on the map $T(x)$. We can achieve this goal and avoid certain measure-theoretic subtleties by requiring the existence of an intensity function $\lambda_T(y)$ such that

$$\int_A \lambda_T(y)\, dy \quad = \quad \int_{T^{-1}(A)} \lambda(x)\, dx \tag{6.14}$$

for all measurable $A \subset V$. Equality (6.14) is just another way of stating that the expected number $\mathrm{E}(N_A)$ of transformed random points on A matches the expected number of random points $\mathrm{E}(N_{T^{-1}(A)})$ on the inverse image $T^{-1}(A) = \{x \in S : T(x) \in A\}$ of A. Because the inverse image operation sends disjoint regions B and C into disjoint inverse regions $T^{-1}(B)$ and $T^{-1}(C)$, the numbers of transformed random points N_B and N_C in B and C enjoy the crucial independence property of a Poisson process. A possible difficulty in the construction of $\lambda_T(y)$ lies in finding a sequence of subregions V_n such that $V = \bigcup_n V_n$ and $\int_{V_n} \lambda_T(y)\, dy < \infty$. The broader definition of a Poisson process adopted in reference [114] solves this apparent problem.

Two special cases of formula (6.14) cover most applications. In the first case, the transformation $T(x)$ is continuously differentiable and invertible. When this is true, the change of variables formula (1.12) implies that

$$\lambda_T(y) \quad = \quad \lambda \circ T^{-1}(y) |\det dT^{-1}(y)|.$$

Of course, invertibility presupposes that the dimensions m and n match. In the second case, suppose that $U = \mathsf{R}^m$ and $V = \mathsf{R}^n$ with $n < m$. Consider the projection $T(x_1, \ldots, x_m) = (x_1, \ldots, x_n)$ of a point x onto its first n coordinates. If a Poisson process on R^m has the intensity function $\lambda(x_1, \ldots, x_m)$, then the projected Poisson process on R^n has the intensity function

$$\lambda_T(x_1, \ldots, x_n) \quad = \quad \int \cdots \int \lambda(x_1, \ldots, x_m)\, dx_{n+1} \cdots dx_m$$

created by integrating over the last $m - n$ variables. Note that the multidimensional integral defining $\lambda_T(x_1, \ldots, x_n)$ can be infinite on a finite region $A \subset \mathsf{R}^n$. When this occurs, the projected Poisson process attributes an

infinite number of random points to A. This phenomenon crops up when $\lambda(x) \equiv 1$.

Example 6.7.1 *Polar Coordinates*

Let $T(x_1, x_2)$ be the map taking each point $(x_1, x_2) \in U = \mathsf{R}^2$ to its polar coordinates (r, θ). The change of variables formula

$$\int\int_A \lambda(r\cos\theta, r\sin\theta) r \, dr \, d\theta \quad = \quad \int\int_{T^{-1}(A)} \lambda(x_1, x_2) \, dx_1 \, dx_2$$

shows that the intensity function $\lambda(x_1, x_2)$ is transformed into the intensity function $\lambda(r\cos\theta, r\sin\theta)r$. If $\lambda(x_1, x_2)$ is a function of $r = \sqrt{x_1^2 + x_2^2}$ alone and we further project onto the r coordinate of (r, θ), then the doubly transformed Poisson process has intensity $2\pi r\lambda(r)$ on the interval $[0, \infty)$. Readers can exploit this fact in solving Problem 3. ■

6.8 Marking and Coloring

Our final construction involves coloring and marking. In coloring, we randomly assign a color to each random point in a Poisson process, with probability p_k attributed to color k. Expanding on Proposition 6.2.1, we can assert that the random points of different colors form independent Poisson processes. If $\lambda(x)$ is the intensity function of the overall process, then $p_k\lambda(x)$ is the intensity function of the Poisson process for color k.

In marking, we generalize this paradigm in two ways. First, we replace colors by points y in some arbitrary marking space M. Second, we allow the selection procedure assigning a mark $y \in M$ to a point $x \in S$ to depend on x. Thus, we select y according to the probability density $p(y \mid x)$, which we now assume to be a continuous density for the sake of consistency with our slightly restricted definition of a Poisson process. Marking is still carried out independently from point to point. The marking theorem [114] says that the pairs (X, Y) of random points X and their associated marks Y generate a Poisson process on the product space $S \times M$ with intensity function $\lambda(x)p(y \mid x)$. Thus, the expected number of pairs falling in the region $R \subset S \times M$ is $\int\int_R \lambda(x)p(y \mid x) \, dy \, dx$. If we combine marking with projection onto the marking space, then we can assert that the random marks Y constitute a Poisson process with intensity $\int \lambda(x)p(y \mid x) \, dx$.

Example 6.8.1 *New Cases of AIDS*

In a certain country, new HIV viral infections occur according to a Poisson process on $(-\infty, \infty)$ with intensity $\lambda(t)$ that varies with time t. Given someone is infected at t, he or she lives a random length of time $Y_t \geq 0$ until the onset of AIDS. Suppose that the latency period (mark) Y_t has a

density function $p(y \mid t)$. If the Y_t are assigned independently from person to person, then the pairs (T, Y_T), of random infection times T and associated latency periods Y_T constitute a Poisson process concentrated in the upper-half plane of R^2. The intensity function of this two-dimensional Poisson process is given by the product $\lambda(t)p(y \mid t)$.

If we apply the mapping procedure to the function $f(t, y) = t + y = u$, then we infer that the random onset times $U = T + Y_T$ of AIDS determine a Poisson process. It follows immediately that the numbers of new AIDS cases arising during disjoint time intervals are independent. The equation

$$\int \int_{\{t+y \in A\}} \lambda(t)p(y \mid t)\, dt\, dy \quad = \quad \int_A \int_{-\infty}^{u} \lambda(t)p(u-t \mid t)\, dt\, du$$

identifies $\int_{-\infty}^{u} \lambda(t)p(u-t \mid t)\, dt$ as the intensity function of the Poisson process. It is of some interest to calculate the expected number $\mathrm{E}(N_{[c,d]})$ of AIDS cases during $[c, d]$ given explicit models for $\lambda(t)$ and $p(y \mid t)$. Under exponential growth of the HIV epidemic, $\lambda(t) = \alpha e^{\beta t}$ for positive constants α and β. The model $p(y \mid t) = \gamma e^{-\gamma(y-\delta)} 1_{\{y \geq \delta\}}$ incorporates an absolute delay δ during which the immune system weakens before the onset of AIDS. After this waiting period, there is a constant hazard rate γ for the appearance of AIDS in an infected person. With these assumptions

$$\begin{aligned} \int_{-\infty}^{u} \lambda(t)p(u-t \mid t)\, dt &= \int_{-\infty}^{u-\delta} \alpha e^{\beta t} \gamma e^{-\gamma(u-t-\delta)} dt \\ &= \frac{\alpha\gamma}{\beta+\gamma} e^{\beta(u-\delta)}, \end{aligned}$$

and

$$\begin{aligned} \mathrm{E}(N_{[c,d]}) &= \frac{\alpha\gamma}{\beta+\gamma} \int_c^d e^{\beta(u-\delta)} du \\ &= \frac{\alpha\gamma}{\beta(\beta+\gamma)} [e^{\beta(d-\delta)} - e^{\beta(c-\delta)}]. \end{aligned}$$

Fortunately, HIV infections are no longer automatically lethal. ■

6.9 Campbell's Moment Formulas

In many Poisson process models, we are confronted with the task of evaluating moments of random sums of the type

$$S \quad = \quad \sum_{X \in \Pi} f(X), \tag{6.15}$$

where X ranges over the random points of a process Π and $f(x)$ is a deterministic measurable function. Campbell devised elegant formulas for

precisely this purpose [114]. It is easiest to derive Campbell's formulas when $f(x) = \sum_{j=1}^m c_j 1_{A_j}$ is a simple function defined by a partition $A_1, \ldots, A_m$ of the underlying space. If N_{A_j} counts the number of random points in A_j, then by virtue of the disjointness of the sets A_j, we can write

$$S = \sum_{j=1}^m c_j N_{A_j}.$$

This representation makes it clear that

$$\begin{aligned} \mathrm{E}(S) &= \sum_{j=1}^m c_j \mathrm{E}(N_{A_j}) \\ &= \sum_{j=1}^m c_j \int_{A_j} \lambda(x)\, dx \qquad (6.16) \\ &= \int f(x)\lambda(x)\, dx, \end{aligned}$$

where $\lambda(x)$ is the intensity function of Π.

Similar reasoning leads to the formulas

$$\mathrm{E}(e^{itS}) = \exp\left\{ \int [e^{itf(x)} - 1]\lambda(x)\, dx \right\} \qquad (6.17)$$

$$\mathrm{E}(u^S) = \exp\left\{ \int [u^{f(x)} - 1]\lambda(x)\, dx \right\} \qquad (6.18)$$

for the characteristic function of S and for the probability generating function of S when $f(x)$ is nonnegative and integer valued. The special value

$$\Pr(S = 0) = \exp\left\{ -\int_{\{x: f(x) > 0\}} \lambda(x)\, dx \right\} \qquad (6.19)$$

of the generating function is important in many applications. To prove formula (6.18), let $f(x) = \sum_{j=1}^m c_j 1_{A_j}$ be a simple function with nonnegative integer values. The steps

$$\begin{aligned} \mathrm{E}(u^S) &= \prod_{j=1}^m \mathrm{E}\left(u^{c_j N_{A_j}}\right) \\ &= \prod_{j=1}^m \exp\left[-\int_{A_j} \lambda(x)\, dx\, (1 - u^{c_j}) \right] \\ &= \exp\left[\sum_{j=1}^m \int_{A_j} \left(u^{c_j} - 1\right)\lambda(x)\, dx \right] \\ &= \exp\left\{ \int [u^{f(x)} - 1]\lambda(x)\, dx \right\} \end{aligned}$$

validate the result.

If we have a second random sum $T = \sum_{X \in \Pi} g(X)$ defined by a simple function $g(x) = \sum_{k=1}^{n} d_k 1_{B_k}$, then it is often useful to calculate $\mathrm{Cov}(S, T)$. Toward this end, note that

$$\begin{aligned} N_{A_j} &= N_{A_j \setminus B_k} + N_{A_j \cap B_k} \\ N_{B_k} &= N_{B_k \setminus A_j} + N_{A_j \cap B_k}. \end{aligned}$$

Because the numbers of random points occurring on disjoint sets are independent and Poisson distributed, these decompositions produce

$$\mathrm{Cov}(N_{A_j}, N_{B_k}) = \mathrm{Var}(N_{A_j \cap B_k}) = \mathrm{E}(N_{A_j \cap B_k}) = \int_{A_j \cap B_k} \lambda(x)\, dx.$$

It follows that

$$\begin{aligned} \mathrm{Cov}(S, T) &= \sum_{j=1}^{m} \sum_{k=1}^{n} c_j d_k \,\mathrm{Cov}(N_{A_j}, N_{B_k}) \\ &= \sum_{j=1}^{m} \sum_{k=1}^{n} c_j d_k \int_{A_j \cap B_k} \lambda(x)\, dx \qquad (6.20) \\ &= \int f(x) g(x) \lambda(x)\, dx. \end{aligned}$$

The special choice $g(x) = f(x)$ yields

$$\mathrm{Var}(S) = \int f(x)^2 \lambda(x)\, dx. \qquad (6.21)$$

Campbell's formulas (6.16), (6.17), (6.18), (6.20), and (6.21) extend to more general functions $f(x)$ and $g(x)$ by passing to appropriate limits [114].

Example 6.9.1 *An Astronomical Application*

Suppose stars occur in the universe $U \subset \mathsf{R}^3$ according to a Poisson process with intensity function $\lambda(x)$. Furthermore, assume each star radiates light at a level y chosen independently from a probability density $p(y)$. The marked Poisson process Π of pairs (X, Y) of random locations and radiation levels has intensity function $\lambda(x)p(y)$. At the center of the universe, the incoming radiation has level

$$S = \sum_{(X,Y) \in \Pi} \frac{Y}{\|X\|^2},$$

where $\|x\|$ is the Euclidean distance of x from the origin, and where we assume that the light radiated by different stars acts additively. Campbell's

formula (6.16) implies

$$\begin{aligned} \mathrm{E}(S) &= \int_U \int \frac{y}{\|x\|^2} \lambda(x) p(y)\, dy\, dx \\ &= \int y p(y)\, dy \int_U \frac{1}{\|x\|^2} \lambda(x)\, dx. \end{aligned}$$

Given this naive physical model and a constant $\lambda(x)$, passage to spherical coordinates shows that it is possible for the three-dimensional integral $\int_U \|x\|^{-2} \lambda(x)\, dx$ to diverge on an unbounded set such as $U = \mathsf{R}^3$. The fact that we are not blinded by light on a starlit night suggests that U is bounded. ■

6.10 Problems

1. Suppose the random variables X and Y have the joint probability generating function

$$\mathrm{E}(u^X v^Y) = e^{\alpha(u-1)+\beta(v-1)+\gamma(uv-1)}$$

for positive constants α, β, and γ. Show that X and Y are Poisson distributed but $X + Y$ is not Poisson distributed [80].

2. Consider a Poisson distributed random variable X whose mean λ is a positive integer. Demonstrate that

$$\Pr(X \geq \lambda) \geq \frac{1}{2}, \qquad \Pr(X \leq \lambda) \geq \frac{1}{2}.$$

(Hints: For the first inequality, show that

$$\Pr(X = \lambda + k) \geq \Pr(X = \lambda - k - 1)$$

when $0 \leq k \leq \lambda - 1$. For the second inequality, show that

$$\Pr(X \leq \lambda) = \int_\lambda^\infty \frac{y^\lambda}{\lambda!} e^{-y} dy$$

and argue that the integrand $f(y)$ satisfies $f(\lambda + y) \geq f(\lambda - y)$ for $y \in [0, \lambda]$.)

3. Consider a Poisson process in the plane with constant intensity λ. Find the distribution and density function of the distance from the origin of the plane to the nearest random point. What is the mean of this distribution?

4. Let $X_1, Y_1, X_2, Y_2, \ldots$ be independent exponentially distributed random variables with mean 1. Define

$$\begin{aligned} N_x &= \min\{n : X_1 + \cdots + X_n > x\} \\ N_y &= \min\{n : Y_1 + \cdots + Y_n > y\}. \end{aligned}$$

Demonstrate that

$$\Pr(N_x \leq N_y) = 1 - e^{-y} \int_0^x I_0(2\sqrt{yt}) e^{-t} dt,$$

where

$$I_0(z) = \sum_{n=0}^{\infty} \frac{1}{(n!)^2} \left(\frac{z}{2}\right)^{2n}$$

is a modified Bessel function of order 0. This problem can be motivated in various ways. For instance, two people X and Y take turns in drawing lengths X_k and Y_k. Person X starts the process and wins if his or her sum exceeds x before person Y's sum exceeds y. Alternatively, a particle travels at unit speed punctuated by random stops and starts at the times of a Poisson process of unit intensity. In this setting $\Pr(N_x \leq N_y)$ is the probability that the particle travels a distance x in time less than $x + y$ [190]. (Hints: First show that

$$\Pr(N_x \leq N_y) = 1 - \sum_{n=1}^{\infty} \Pr(N_y = n) \Pr(X_1 + \cdots + X_n \leq x).$$

Then simplify and invoke the Bessel function.)

5. Let $X_1, \ldots, X_n$ be independent exponentially distributed random variables with common intensity λ. Define the order statistics $X_{(i)}$ and the increments $Z_i = X_{(i)} - X_{(i-1)}$ and $Z_1 = X_{(1)}$. Show that the region $x_1 < x_2 < \cdots < x_n$ contributes the amount

$$f(z_1, \ldots, z_n) = \prod_{i=1}^{n} \lambda e^{-(n-i+1)\lambda z_i}$$

to the joint density of the Z_i. Since the same result holds when collectively $X_{(i)} = X_{\pi(i)}$ for any permutation π, conclude that the Z_i have joint density $n! f(z_1, \ldots, z_n)$. Hence, the Z_i are independent and exponentially distributed with the given intensities.

6. In the context of Example 6.4.2, show that the order statistics $X_{(j)}$ have means, variances, and covariances

$$\mathrm{E}(X_{(j)}) = \sum_{k=1}^{j} \frac{1}{\lambda(n-k+1)}$$

$$\begin{aligned}\mathrm{Var}(X_{(j)}) &= \sum_{k=1}^{j} \frac{1}{\lambda^2(n-k+1)^2} \\ \mathrm{Cov}(X_{(j)}, X_{(k)}) &= \sum_{i=1}^{\min\{j,k\}} \frac{1}{\lambda^2(n-i+1)^2}.\end{aligned}$$

7. Continuing Problem 6, prove that $X_{(j)}$ has distribution and density functions

$$\begin{aligned}F_{(j)}(x) &= \sum_{k=j}^{n} \binom{n}{k} (1-e^{-\lambda x})^k e^{-(n-k)\lambda x} \\ f_{(j)}(x) &= n\binom{n-1}{j-1}(1-e^{-\lambda x})^{j-1} e^{-(n-j)\lambda x} \lambda e^{-\lambda x}.\end{aligned}$$

(Hint: See Problem 13 of Chapter 4.)

8. Let $X_1, \ldots, X_n$ be independent exponentially distributed random variables with intensities $\lambda_1, \ldots, \lambda_n$. If $\lambda_j \neq \lambda_k$ for $j \neq k$, then show that $S = X_1 + \cdots + X_n$ has density

$$\begin{aligned}f(t) &= \sum_{j=1}^{n} c_j \lambda_j e^{-\lambda_j t} \\ c_j &= \prod_{k \neq j} \frac{\lambda_k}{\lambda_k - \lambda_j}\end{aligned}$$

for $t > 0$. In the particular case $\lambda_j = j\lambda$, show that S has density

$$f(t) = n\lambda \sum_{j=1}^{n} (-1)^{j-1} \binom{n-1}{j-1} e^{-j\lambda t}.$$

How does this result relate to Example 6.4.2? (Hint: Decompose the Laplace transform of S by partial fractions.)

9. In the context of Example 6.4.2, suppose you observe $X_{(1)}, \ldots, X_{(r)}$ and wish to estimate λ^{-1} by a linear combination $S = \sum_{i=1}^{r} \alpha_i X_{(i)}$. Demonstrate that $\mathrm{Var}(S)$ is minimized subject to $\mathrm{E}(S) = \lambda^{-1}$ by taking $\alpha_j = 1/r$ for $1 \leq j < r$ and $\alpha_r = (n-r+1)/r$ [60].

10. Let $X_1, X_2, \ldots$ be an i.i.d. sequence of exponentially random variables with common intensity 1. The observation X_i is said to be a record value if either $i = 1$ or $X_i > \max\{X_1, \ldots, X_{i-1}\}$. If R_j denotes the jth record value and I_j the jth record index, then I_j equals $\min\{i : X_i > R_{j-1}\}$. Argue that R_j has the gamma density $\frac{r^{j-1}}{(j-1)!} e^{-r}$

on the interval $(0, \infty)$. Using this fact, prove that

$$\begin{aligned}
\Pr(I_{j+1} - I_j > k) &= \int_0^\infty (1 - e^{-r})^k \frac{r^{j-1}}{(j-1)!} e^{-r} dr \\
\mathrm{E}(I_{j+1} - I_j) &= \sum_{k=0}^{\infty} \Pr(I_{j+1} - I_j > k) \\
&= \int_0^\infty \frac{r^{j-1}}{(j-1)!}\, dr \\
&= \infty
\end{aligned}$$

for $j \geq 1$. Also show that

$$\Pr(I_j = n) = \frac{1}{n!} \begin{bmatrix} n-1 \\ j-1 \end{bmatrix},$$

where $\begin{bmatrix} n-1 \\ j-1 \end{bmatrix}$ is a Stirling number of the first kind. Why do the formulas for $I_{j+1} - I_j$ and I_j generalize to any i.i.d. sequence $X_1, X_2, \ldots$ with a continuous distribution function? (Hints: To derive the distribution of R_j, invoke the lack of memory property of the exponential distribution. To find $\Pr(I_j = n)$, consider permutations of $\{1, \ldots, n\}$.)

11. For a fixed positive integer n, we define the generalized hyperbolic functions [146] ${}_n\alpha_j(x)$ of x as the finite Fourier transform coefficients

$${}_n\alpha_j(x) = \frac{1}{n} \sum_{k=0}^{n-1} e^{x u_n^k} u_n^{-jk},$$

where $u_n = e^{2\pi i/n}$ is the nth principal root of unity. These functions generalize the hyperbolic trigonometric functions $\cosh(x)$ and $\sinh(x)$. Prove the following assertions:

(a) ${}_n\alpha_j(x) = {}_n\alpha_{j+n}(x)$.

(b) ${}_n\alpha_j(x+y) = \sum_{k=0}^{n-1} {}_n\alpha_k(x)\,{}_n\alpha_{j-k}(y)$.

(c) ${}_n\alpha_j(x) = \sum_{k=0}^{\infty} \frac{x^{j+kn}}{(j+kn)!}$ for $0 \leq j \leq n-1$.

(d) $\frac{d}{dx}\big[{}_n\alpha_j(x)\big] = {}_n\alpha_{j-1}(x)$.

(e) $\lim_{x\to\infty} e^{-x}\,{}_n\alpha_j(x) = \frac{1}{n}$.

(f) In a Poisson process of intensity 1, $e^{-x}\,{}_n\alpha_j(x)$ is the probability that the number of random points on $[0, x]$ equals j modulo n.

(g) Relative to this Poisson process, let N_x count every nth random point on $[0, x]$. Then N_x has probability generating function

$$P(s) = e^{-x} \sum_{j=0}^{n-1} s^{-\frac{j}{n}}\, {}_n\alpha_j(s^{\frac{1}{n}} x).$$

(h) Furthermore, N_x has mean

$$\mathrm{E}(N_x) = \frac{x}{n} - \frac{e^{-x}}{n}\sum_{j=0}^{n-1} j_n \alpha_j(x).$$

(i) $\lim_{x\to\infty}\left[\mathrm{E}(N_x) - \frac{x}{n}\right] = -\frac{n-1}{2n}$.

12. Show that the loglikelihood (6.5) for the transmission tomography model is concave. State a necessary condition for strict concavity in terms of the number of pixels and the number of projections. Prove that the sufficient conditions mentioned in the text guarantee that the logposterior function $L(\theta) + \ln \pi(\theta)$ is strictly concave.

13. In the absence of a Gibbs smoothing prior, show that one step of Newton's method leads to the approximate MM update

$$\theta_j^{n+1} = \theta_j^n \frac{\sum_i l_{ij}[d_i e^{-l_i^t\theta^n}(1 + l_i^t\theta^n) - y_i]}{\sum_i l_{ij} l_i^t \theta^n d_i e^{-l_i^t\theta^n}}$$

in the transmission tomography model.

14. Under the assumptions of Problem 13, demonstrate that the exact solution of the one-dimensional equation

$$\frac{\partial}{\partial\theta_j} Q(\theta \mid \theta^n) = 0$$

exists and is positive when $\sum_i l_{ij} d_i > \sum_i l_{ij} y_i$. Why would this condition usually obtain in practical implementations of transmission tomography?

15. Prove that the function $\psi(r) = \ln[\cosh(r)]$ is even, strictly convex, infinitely differentiable, and asymptotic to $|r|$ as $|r| \to \infty$.

16. Suppose you randomly drop n balls into m boxes. Assume that a ball is equally likely to land in any box. Use Schrödinger's method to prove that each box receives an even number of balls with probability

$$e_n = \frac{1}{2^m}\sum_{j=0}^{m}\binom{m}{j}\left(1 - \frac{2j}{m}\right)^n$$

and an odd number of balls with probability

$$o_n = \frac{1}{2^m}\sum_{j=0}^{m}\binom{m}{j}(-1)^j\left(1 - \frac{2j}{m}\right)^n.$$

(Hint: The even terms of e^t sum to $\frac{1}{2}(e^t + e^{-t})$ and the odd terms to $\frac{1}{2}(e^t - e^{-t})$.)

17. Continuing Problem 16, show that the probability that exactly j boxes are empty is

$$\binom{m}{j}\sum_{k=0}^{m-j}\binom{m-j}{k}(-1)^{m-j-k}\Big(\frac{k}{m}\Big)^n.$$

18. Prove the upper bound (6.11) by calculating $\mathrm{E}[f(N_1,\ldots,N_m)]$. In the process, condition on the number of Poisson trials N, invoke the assumptions on $f(x_1,\ldots,x_m)$, and apply the bounds of Problem 2 [144].

19. In the family planning model of Example 2.3.3, we showed how to compute the probability R_{sd} that the couple reach their quota of s sons before their quota of d daughters. Deduce the formula

$$R_{sd} \;=\; \sum_{k=0}^{d-1}\binom{s+k-1}{s-1}p^s q^k$$

by viewing the births to the couple as occurring at the times of a Poisson process with unit intensity. Can you also derive this formula by counting all possible successful sequences of births? (Hint: The final birth in a successful sequence is a son.)

20. In the family planning model of Example 6.6.2, let M_{sd} be the number of children born when the family first attains either its quota of s sons or d daughters. Show that

$$\mathrm{E}(M_{sd}) \;=\; \mathrm{E}(\min\{T_s,T_d\}) \;=\; \sum_{k=0}^{s-1}\sum_{l=0}^{d-1}\binom{k+l}{k}p^k q^l.$$

Note that the formulas for $\mathrm{E}(\max\{T_s,T_d\})$ and $\mathrm{E}(\min\{T_s,T_d\})$ together yield

$$\mathrm{E}(\min\{T_s,T_d\})+\mathrm{E}(\max\{T_s,T_d\}) \;=\; \mathrm{E}(T_s)+\mathrm{E}(T_d). \quad (6.22)$$

Prove the general identity

$$\mathrm{E}(\min\{X,Y\})+\mathrm{E}(\max\{X,Y\}) \;=\; \mathrm{E}(X)+\mathrm{E}(Y)$$

for any pair of random variables X and Y with finite expectations. Finally, argue that

$$\mathrm{E}(M_{sd}) \;=\; d\sum_{k=0}^{s-1}\binom{d+k}{k}p^k q^d + s\sum_{l=0}^{d-1}\binom{s+l}{l}p^s q^l \quad (6.23)$$

by counting all possible successful sequences of births that lead to either the daughter quota or the son quota being fulfilled first. Combining equations (6.22) and (6.23) permits us to write

$$\mathrm{E}(N_{sd}) = \frac{s}{p} + \frac{d}{q} - d\sum_{k=0}^{s-1}\binom{d+k}{k}p^k q^d - s\sum_{l=0}^{d-1}\binom{s+l}{l}p^s q^l,$$

replacing a double sum with two single sums.

21. The motivation for the negative-multinomial distribution comes from multinomial sampling with $d+1$ categories assigned probabilities $p_1, \ldots, p_{d+1}$. Sampling continues until category $d+1$ accumulates β outcomes. At that moment we count the number of outcomes x_i falling in category i for $1 \le i \le d$. Demonstrate that the count vector $x = (x_1, \ldots, x_d)$ has probability

$$\begin{aligned}\Pr(X = x) &= \binom{\beta + |x| - 1}{|x|}\binom{|x|}{x_1 \ldots x_d}\prod_{i=1}^{d} p_i^{x_i}\, p_{d+1}^{\beta} \\ &= \frac{\beta(\beta+1)\cdots(\beta+|x|-1)}{x_1! \cdots x_d!}\prod_{i=1}^{d} p_i^{x_i}\, p_{d+1}^{\beta}. \quad (6.24)\end{aligned}$$

This formula continues to make sense even if the positive parameter β is not an integer. For arbitrary $\beta > 0$, the most straightforward way to construct the negative-multinomial distribution is to run d independent Poisson processes with intensities $p_1, \ldots, p_d$. Wait a gamma distributed length of time with shape parameter β and intensity parameter p_{d+1}. At the expiration of this waiting time, count the number of random events X_i of each type i among the first d categories. The random vector X has precisely the discrete density (6.24). Calculate the moments

$$\begin{aligned}\mathrm{E}(X_i) &= \beta\frac{p_i}{p_{d+1}} \\ \mathrm{Var}(X_i) &= \beta\frac{p_i}{p_{d+1}}\Big(1 + \frac{p_i}{p_{d+1}}\Big) \\ \mathrm{Cov}(X_i, X_j) &= \beta\frac{p_i}{p_{d+1}}\frac{p_j}{p_{d+1}}, \quad i \ne j.\end{aligned}$$

based on this Poisson process perspective.

22. Suppose we generate random circles in the plane by taking their centers (x, y) to be the random points of a Poisson process of constant intensity λ. Each center we independently mark with a radius r sampled from a probability density $g(r)$ on $[0, \infty)$. If we map each random triple (X, Y, R) to the point $U = \sqrt{X^2 + Y^2} - R$, then show that the

random points so generated constitute a Poisson process with intensity

$$\eta(u) \quad = \quad 2\pi\lambda \int_0^\infty (r+u)_+ g(r)\, dr.$$

Conclude from this analysis that the number of random circles that overlap the origin is Poisson with mean $\lambda\pi \int_0^\infty r^2 g(r)\, dr$ [196].

23. Continuing Problem 22, perform the same analysis in three dimensions for spheres. Conclude that the number of random spheres that overlap the origin is Poisson with mean $\frac{4\lambda\pi}{3} \int_0^\infty r^3 g(r)\, dr$ [196].

24. Consider a homogeneous Poisson process with intensity λ on the set $\{(x, y, t) \in \mathsf{R}^3 : t \geq 0\}$. The coordinate t is considered a time coordinate and the coordinates x and y spatial coordinates. If (x, y, t) is a random point, then a disc centered at (x, y) starts growing at position (x, y) at time t with radial speed v. Thus, at time $t + u$, the disc has radius uv. As time goes on more and more such discs appear. These discs may overlap. This process is called the Johnson-Mehl model. One question of obvious interest is the fraction of the (x, y) plane that is covered by at least one disc at time t. Show that this fraction equals the Poisson tail probability $1 - e^{-\lambda v^2 \pi t^3/3}$. (Hint: It suffices to consider the origin $\mathbf{0}$ in R^2. The region in R^3 that gives rise to circles overlapping $\mathbf{0}$ at time t is a cone.)

25. A one-way highway extends from 0 to ∞. Cars enter at position 0 at times s determined by a Poisson process on $[0, t]$ with constant intensity λ. Each car is independently assigned a velocity v from a density $g(v)$ on $[0, \infty)$. Demonstrate that the number of cars located in the interval (a, b) at time t has a Poisson distribution with mean $\lambda \int_0^t [G(\frac{b}{t-s}) - G(\frac{a}{t-s})]\, ds$, where $G(v)$ is the distribution function of $g(v)$ [170].

26. A train departs at time $t > 0$. During the interval $[0, t]$, passengers arrive at the depot at times T determined by a Poisson process with constant intensity λ. The total waiting time passengers spend at the depot is $W = \sum_T (t - T)$. Show that W has mean $\mathrm{E}(W) = \frac{\lambda t^2}{2}$ and variance $\mathrm{Var}(W) = \frac{\lambda t^3}{3}$ by invoking Campbell's formulas (6.16) and (6.21) [170].

27. Claims arrive at an insurance company at the times T of a Poisson process with constant intensity λ on $[0, \infty)$. Each time a claim arrives, the company pays S dollars, where S is independently drawn from a probability density $g(s)$ on $[0, \infty)$. Because of inflation and the ability of the company to invest premiums, the longer a claim is delayed, the less it costs the company. If a claim is discounted at rate β, then show

that the company's ultimate liability $L = \sum_T Se^{-\beta T}$ has mean and variance

$$\begin{aligned} \mathrm{E}(L) &= \frac{\lambda}{\beta} \int_0^\infty sg(s)\, ds \\ \mathrm{Var}(L) &= \frac{\lambda}{2\beta} \int_0^\infty s^2 g(s)\, ds. \end{aligned}$$

(Hints: The random pairs (T, S) constitute a marked Poisson process. Use Campbell's formulas (6.16) and (6.21).)

28. If $f(x)$ is a simple function and Π is a Poisson process with intensity function $\lambda(x)$, then demonstrate formula (6.17) for the characteristic function of the random sum S.

29. A random variable S is said to be infinitely divisible if for every positive integer n there exist independent and identically distributed random variables $X_1, \ldots, X_n$ such that the sum $X_1 + \cdots + X_n$ has the same distribution as S. Show that Campbell's sum (6.15) is infinitely divisible. (Hint: Superimpose independent Poisson processes.)

30. Consider a homogeneous Poisson process N_t on $[0, \infty)$ with intensity λ. Assign to the ith random point a real mark Y_i drawn independently from a density $p(y)$ that does not depend on the location of the point. The random sums $S_t = \sum_{i=1}^{N_t} Y_i$ constitute a compound Poisson process. Use Campbell's formulas to calculate the mean, variance, and characteristic function of S_t and the covariance $\mathrm{Cov}(S_{t_1}, S_{t_2})$ for $t_1 \neq t_2$.

7
Discrete-Time Markov Chains

7.1 Introduction

Applied probability thrives on models. Markov chains are one of the richest sources of good models for capturing dynamical behavior with a large stochastic component [23, 24, 59, 80, 106, 107, 118]. In this chapter we give a few examples and a quick theoretical overview of discrete-time Markov chains. The highlight of our theoretical development, Proposition 7.4.1, relies on a coupling argument. Because coupling is one of the most powerful and intuitively appealing tools available to probabilists, we examine a few of its general applications as well. We also stress reversible Markov chains. Reversibility permits explicit construction of the long-run or equilibrium distribution of a chain when such a distribution exists. Chapter 8 will cover continuous-time Markov chains.

7.2 Definitions and Elementary Theory

For the sake of simplicity, we will only consider chains with a finite or countable number of states [23, 59, 80, 106, 107]. The movement of such a chain from epoch to epoch (equivalently generation to generation) is governed by its transition probability matrix $P = (p_{ij})$. This matrix is infinite dimensional when the number of states is infinite. If Z_n denotes the state of the chain at epoch n, then $p_{ij} = \Pr(Z_n = j \mid Z_{n-1} = i)$. As a consequence, every entry of P satisfies $p_{ij} \geq 0$, and every row of P satisfies

K. Lange, *Applied Probability,*
DOI 10.1007/978-1-4419-7165-4_7, © Springer Science+Business Media, LLC 2010

$\sum_j p_{ij} = 1$. Implicit in the definition of p_{ij} is the fact that the future of the chain is determined by its present regardless of its past. This Markovian property is expressed formally by the equation

$$\Pr(Z_n = i_n \mid Z_{n-1} = i_{n-1}, \ldots, Z_0 = i_0) \quad = \quad \Pr(Z_n = i_n \mid Z_{n-1} = i_{n-1}).$$

The n-step transition probability $p_{ij}^{(n)} = \Pr(Z_n = j \mid Z_0 = i)$ is given by the entry in row i and column j of the matrix power P^n. This follows because the decomposition

$$p_{ij}^{(n)} \quad = \quad \sum_{i_1} \cdots \sum_{i_{n-1}} p_{ii_1} \cdots p_{i_{n-1}j}$$

over all paths $i \to i_1 \to \cdots \to i_{n-1} \to j$ of n steps corresponds to $n-1$ matrix multiplications. If the chain tends toward stochastic equilibrium, then the limit of $p_{ij}^{(n)}$ as n increases should exist independently of the starting state i. In other words, the matrix powers P^n should converge to a matrix with identical rows. Denoting the common limiting row by π, we deduce that $\pi = \pi P$ from the calculation

$$\begin{aligned} \begin{pmatrix} \pi \\ \vdots \\ \pi \end{pmatrix} &= \lim_{n\to\infty} P^{n+1} \\ &= \left(\lim_{n\to\infty} P^n \right) P \\ &= \begin{pmatrix} \pi \\ \vdots \\ \pi \end{pmatrix} P. \end{aligned}$$

Any probability distribution π on the states of the chain satisfying the condition $\pi = \pi P$ is termed an equilibrium (or stationary) distribution of the chain. The jth component

$$\pi_j \quad = \quad \sum_i \pi_i p_{ij} \tag{7.1}$$

of the equation $\pi = \pi P$ suggests a balance between the probabilistic flows into and out of state j. Indeed, if the left-hand side of equation (7.1) represents the probability of being in state j at the current epoch, then the right-hand side represents the probability of being in state j at the next epoch. At equilibrium, these two probabilities must match. For finite-state chains, equilibrium distributions always exist [59, 80]. The real issue is uniqueness.

Probabilists have attacked the uniqueness problem by defining appropriate ergodic conditions. For finite-state Markov chains, two ergodic assumptions are invoked. The first is aperiodicity; this means that the greatest

common divisor of the set $\{n \geq 1 : p_{ii}^{(n)} > 0\}$ is 1 for every state i. Aperiodicity trivially holds when $p_{ii} > 0$ for all i. The second ergodic assumption is irreducibility; this means that for every pair of states (i, j), there exists a positive integer n_{ij} such that $p_{ij}^{(n_{ij})} > 0$. In other words, every state is reachable from every other state. Said yet another way, all states communicate. For a finite-state irreducible chain, Problem 4 states that the integer n_{ij} can be chosen independently of the particular pair (i, j) if and only if the chain is also aperiodic. Thus, we can merge the two ergodic assumptions into the single assumption that some power P^n has all entries positive. Under this single ergodic condition, we show in Proposition 7.4.1 that a unique equilibrium distribution π exists and that $\lim_{n\to\infty} p_{ij}^{(n)} = \pi_j$. Because all states communicate, the entries of π are necessarily positive.

Equally important is the ergodic theorem [59, 80]. This theorem permits one to run a chain and approximate theoretical means by sample means. More precisely, let $f(z)$ be some real-valued function defined on the states of an ergodic chain. Then given that Z_i is the state of the chain at epoch i and π is the equilibrium distribution, we have

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f(Z_i) \;=\; \mathrm{E}_\pi[f(Z)] \;=\; \sum_z \pi_z f(z).$$

This result generalizes the law of large numbers for independent sampling and has important applications in Markov chain Monte Carlo methods as discussed later in this chapter. The ergodic theorem generalizes to periodic irreducible chains even though $\lim_{n\to\infty} P^n$ no longer exists. Problem 9 also indicates that uniqueness of the equilibrium distribution has nothing to do with aperiodicity.

In many Markov chain models, the equilibrium distribution satisfies the stronger condition

$$\pi_j p_{ji} \;=\; \pi_i p_{ij} \tag{7.2}$$

for all pairs (i, j). If this is the case, then the probability distribution π is said to satisfy detailed balance, and the Markov chain, provided it is irreducible, is said to be reversible. Summing equation (7.2) over i yields the equilibrium condition (7.1). Thus, detailed balance implies balance. Irreducibility is imposed as part of reversibility to guarantee that π is unique and has positive entries. Given the latter condition, detailed balance implies that $p_{ij} > 0$ if and only if $p_{ji} > 0$.

If $i_1, \ldots, i_m$ is any sequence of states in a reversible chain, then detailed balance also entails

$$\begin{aligned} \pi_{i_1} p_{i_1 i_2} &= \pi_{i_2} p_{i_2 i_1} \\ \pi_{i_2} p_{i_2 i_3} &= \pi_{i_3} p_{i_3 i_2} \\ &\vdots \end{aligned}$$

$$\begin{aligned}\pi_{i_{m-1}}p_{i_{m-1}i_m} &= \pi_{i_m}p_{i_mi_{m-1}}\\ \pi_{i_m}p_{i_mi_1} &= \pi_{i_1}p_{i_1i_m}.\end{aligned}$$

Multiplying these equations together and canceling the common positive factor $\pi_{i_1}\cdots\pi_{i_m}$ from both sides of the resulting equality give Kolmogorov's circulation criterion [111]

$$p_{i_1i_2}p_{i_2i_3}\cdots p_{i_{m-1}i_m}p_{i_mi_1} = p_{i_1i_m}p_{i_mi_{m-1}}\cdots p_{i_3i_2}p_{i_2i_1}. \tag{7.3}$$

Conversely, suppose an irreducible Markov chain satisfies Kolmogorov's criterion. One can easily demonstrate that $p_{ij} > 0$ if and only if $p_{ji} > 0$. Indeed, if $p_{ij} > 0$, then take a path of positive probability from j back to i. This creates a circuit from i to i whose first step goes from i to j. Kolmogorov's criterion shows that the reverse circuit from i to i whose last step goes from j to i also has positive probability. We can also prove that the chain is reversible by explicitly constructing the equilibrium distribution and showing that it satisfies detailed balance. The idea behind the construction is to choose some arbitrary reference state i and to pretend that π_i is given. If j is another state, let $i \to i_1 \to \cdots \to i_m \to j$ be any path leading from i to j. Then the formula

$$\pi_j = \pi_i\frac{p_{ii_1}p_{i_1i_2}\cdots p_{i_mj}}{p_{ji_m}p_{i_mi_{m-1}}\cdots p_{i_1i}} \tag{7.4}$$

defines π_j. A straightforward application of Kolmogorov's criterion (7.3) shows that definition (7.4) does not depend on the particular path chosen from i to j. To validate detailed balance, suppose that k is adjacent to j. Then $i \to i_1 \to \cdots \to i_m \to j \to k$ furnishes a path from i to k through j. It follows from (7.4) that

$$\pi_k = \pi_i\frac{p_{ii_1}p_{i_1i_2}\cdots p_{i_mj}p_{jk}}{p_{ji_m}p_{i_mi_{m-1}}\cdots p_{i_1i}p_{kj}} = \pi_j\frac{p_{jk}}{p_{kj}},$$

which is obviously equivalent to detailed balance. In general, the value of π_i is determined by the requirement that $\sum_j \pi_j = 1$. For a chain with a finite number of states, we can guarantee this condition by replacing π by $\tilde{\pi}$ with components

$$\tilde{\pi}_j = \frac{\pi_j}{\sum_k \pi_k}.$$

In practice, explicit calculation of the sum $\sum_k \pi_k$ may be nontrivial. For a chain with an infinite number of states, in contrast, it may be impossible to renormalize the π_j defined by equation (7.4) so that $\sum_j \pi_j = 1$. This situation occurs in Example 7.3.1 in the next section.

7.3 Examples

Here are a few examples of discrete-time chains classified according to the concepts just introduced. If possible, the unique equilibrium distribution is identified. For some irreducible chains, note that Kolmogorov's circulation criterion is trivial to verify. If we put an edge between two states i and j whenever $p_{ij} > 0$, then this construction induces a graph. If the graph reduces to a tree, then it can have no cycles, and Kolmogorov's circulation criterion is automatically satisfied.

Example 7.3.1 *Random Walk on a Graph*

Consider a connected graph with node set N and edge set E. The number of edges $d(v)$ incident on a given node v is called the degree of v. Owing to the connectedness assumption, $d(v) > 0$ for all $v \in N$. Now define the transition probability matrix $P = (p_{uv})$ by

$$p_{uv} = \begin{cases} \frac{1}{d(u)} & \text{for } \{u, v\} \in E \\ 0 & \text{for } \{u, v\} \notin E. \end{cases}$$

This Markov chain is irreducible because of the connectedness assumption. It is also aperiodic unless the graph is bipartite. (A graph is said to be bipartite if we can partition its node set into two disjoint subsets F and M, say females and males, such that each edge has one node in F and the other node in M.) If E has m edges, then the equilibrium distribution π of the chain has components $\pi_v = \frac{d(v)}{2m}$. It is trivial to show that this choice of π satisfies detailed balance.

One hardly needs this level of symmetry to achieve detailed balance. For instance, consider a random walk on the nonnegative integers with neighboring integers connected by an edge. For $i > 0$ let

$$p_{ij} = \begin{cases} q_i, & j = i - 1 \\ r_i, & j = i \\ p_i, & j = i + 1 \\ 0, & \text{otherwise.} \end{cases}$$

At the special state 0, set $p_{00} = r_0$ and $p_{01} = p_0$ for $p_0 + r_0 = 1$. With state 0 as a reference state, Kolmogorov's formula (7.4) becomes

$$\pi_i = \pi_0 \prod_{j=1}^{i} \frac{p_{j-1}}{q_j}.$$

This definition of π_i satisfies detailed balance because the graph of the chain is a tree. However, because the state space is infinite, we must impose the additional constraint

$$\sum_{i=0}^{\infty} \prod_{j=1}^{i} \frac{p_{j-1}}{q_j} < \infty$$

to achieve a legitimate equilibrium distribution. When this condition holds, we define

$$\pi_i = \frac{\prod_{j=1}^{i} \frac{p_{j-1}}{q_j}}{\sum_{k=0}^{\infty} \prod_{j=1}^{k} \frac{p_{j-1}}{q_j}} \tag{7.5}$$

and eliminate the unknown π_0. For instance, if all $p_j = p$ and all $q_j = q$, then $p < q$ is both necessary and sufficient for the existence of an equilibrium distribution. When $p < q$, formula (7.5) implies $\pi_i = \frac{(q-p)p^i}{q^{i+1}}$. ■

Example 7.3.2 *Wright-Fisher Model of Genetic Drift*

Consider a population of m organisms from some animal or plant species. Each member of this population carries two genes at some genetic locus, and these genes take two forms (or alleles) labeled a_1 and a_2. At each generation, the population reproduces itself by sampling $2m$ genes with replacement from the current pool of $2m$ genes. If Z_n denotes the number of a_1 alleles at generation n, then it is clear that the Z_n constitute a Markov chain with binomial transition probability matrix

$$p_{jk} = \binom{2m}{k} \left(\frac{j}{2m}\right)^k \left(1 - \frac{j}{2m}\right)^{2m-k}.$$

This chain is reducible because once one of the states 0 or $2m$ is reached, the corresponding allele is fixed in the population, and no further variation is possible. An infinity of equilibrium distributions π exist. Each one is characterized by $\pi_0 = \alpha$ and $\pi_{2m} = 1 - \alpha$ for some $\alpha \in [0, 1]$. ■

Example 7.3.3 *Ehrenfest's Model of Diffusion*

Consider a box with m gas molecules. Suppose the box is divided in half by a rigid partition with a very small hole. Molecules drift aimlessly around each half until one molecule encounters the hole and passes through. Let Z_n be the number of molecules in the left half of the box at epoch n. If epochs are timed to coincide with molecular passages, then the transition matrix of the chain is

$$p_{jk} = \begin{cases} 1 - \frac{j}{m} & \text{for } k = j + 1 \\ \frac{j}{m} & \text{for } k = j - 1 \\ 0 & \text{otherwise.} \end{cases}$$

This chain is a random walk with finite state space. It is periodic with period 2, irreducible, and reversible with equilibrium distribution

$$\pi_j = \binom{m}{j} \left(\frac{1}{2}\right)^m.$$

The binomial form of π_j follows from equation (7.5). ■

Example 7.3.4 *Discrete Renewal Process*

Many treatments of Markov chain theory depend on a prior development of renewal theory. Here we reverse the logical flow and consider a discrete renewal process as a Markov chain. A renewal process models repeated visits to a special state [59, 80]. After it enters the special state, the process departs and eventually returns for the first time after $j > 0$ steps with probability f_j. The return times following different visits are independent. In modeling this behavior by a Markov chain, we let Z_n denote the number of additional epochs left after epoch n until a return to the special state occurs. The renewal mechanism generates the transition matrix with entries

$$p_{ij} = \begin{cases} f_{j+1}, & i = 0 \\ 1, & i > 0 \text{ and } j = i - 1 \\ 0, & i > 0 \text{ and } j \neq i - 1. \end{cases}$$

In order for $\sum_j p_{0j} = 1$, we must have $f_0 = 0$. If $f_j = 0$ for $j > m$, then the chain has m states; otherwise, it has an infinite number of states. Because the chain always ratchets downward from $i > 0$ to $i-1$, it is both irreducible and irreversible. State 0, and therefore the whole chain, is aperiodic if and only if the set $\{j : f_j > 0\}$ has greatest common divisor 1. This number theoretic fact is covered in Appendix A.1.

One of the primary concerns in renewal theory is predicting what fraction of epochs are spent in the special state. This problem is solved by finding the equilibrium probability π_0 of state 0 in the associated Markov chain. Assuming the mean $\mu = \sum_i i f_i$ is finite, we can easily calculate the equilibrium distribution. The balance conditions defining equilibrium are

$$\pi_j = \pi_{j+1} + \pi_0 f_{j+1}.$$

One can demonstrate by induction that the unique solution to this system of equations is given by

$$\pi_j = \pi_0 \left(1 - \sum_{i=1}^{j} f_i\right) = \pi_0 \sum_{i=j+1}^{\infty} f_i$$

subject to

$$1 = \sum_{j=0}^{\infty} \pi_j = \pi_0 \sum_{j=0}^{\infty} \sum_{i=j+1}^{\infty} f_i = \pi_0 \mu.$$

Here we assume an infinite number of states and invoke Example 2.5.1. It follows that $\pi_0 = \mu^{-1}$ and $\pi_j = \mu^{-1} \sum_{i=j+1}^{\infty} f_i$.

Since the return visits to a specific state of a Markov chain constitute a renewal process, the formula $\pi_0 = \mu^{-1}$ provides an alternative way of defining the equilibrium distribution. For a symmetric random walk on

the integers, the mean first passage time μ from any state back to itself is infinite. Problems 36 and 37 ask the reader to check this fact as well as the fact that the return probabilities f_n sum to 1. These observations are obviously consistent with the failure of Kolmogorov's formula (7.4) to deliver an equilibrium distribution with finite mass. Markov chains with finitely many states cannot exhibit such null recurrent behavior. ■

Example 7.3.5 *Card Shuffling and Random Permutations*

Imagine a deck of cards labeled $1, \ldots, m$. The cards taken from top to bottom provide a permutation σ of these m labels. The usual method of shuffling cards, the so-called riffle shuffle, is difficult to analyze probabilistically. A far simpler shuffle is the top-in shuffle [3, 49]. In this shuffle, one takes the top card on the deck and moves it to a random position in the deck. Of course, if the randomly chosen position is the top position, then the deck suffers no change. Repeated applications of the top-in shuffle constitute a Markov chain. This chain is aperiodic and irreducible. Aperiodicity is obvious because the deck can remain constant for an arbitrary number of shuffles. Irreducibility is slightly more subtle. Suppose we follow the card originally at the bottom of the deck. Cards inserted below it occur in completely random order. Once the original bottom card reaches the top of the deck and is moved, then the whole deck is randomly rearranged. This argument shows that all permutations are ultimately equally likely and can be reached from any starting permutation. We extend this analysis in Example 7.4.3. Finally, the chain is irreversible. For example, if a deck of seven cards is currently in the order $\sigma = (4, 7, 5, 2, 3, 1, 6)$, equating left to top and right to bottom, then inserting the top card 4 in position 3 produces $\eta = (7, 5, 4, 2, 3, 1, 6)$. Clearly, it is impossible to return from η to σ by moving the new top card 7 to another position. Under reversibility, each individual step of a Markov chain can be reversed. ■

7.4 Coupling

In this section we undertake an investigation of the convergence of an ergodic Markov chain to its equilibrium. Our method of attack exploits a powerful proof technique known as coupling. By definition, two random variables or stochastic processes are coupled if they reside on the same probability space [134, 136]. As a warm-up, we illustrate coupling arguments by two examples having little to do with Markov chains.

Example 7.4.1 *Correlated Random Variables*

Suppose X is a random variable and the functions $f(x)$ and $g(x)$ are both increasing or both decreasing. If the random variables $f(X)$ and $g(X)$ have finite second moments, then it is reasonable to conjecture that

$\text{Cov}[f(X), g(X)] \geq 0$. To prove this fact by coupling, consider a second random variable Y independent of X but sharing the same distribution. If $f(x)$ and $g(x)$ are both increasing or both decreasing, then the product $[f(X) - f(Y)][g(X) - g(Y)] \geq 0$. Hence,

$$\begin{aligned} 0 &\leq \text{E}\{[f(X) - f(Y)][g(X) - g(Y)]\} \\ &= \text{E}[f(X)g(X)] + \text{E}[f(Y)g(Y)] - \text{E}[f(X)]\,\text{E}[g(Y)] - \text{E}[f(Y)]\,\text{E}[g(X)] \\ &= 2\,\text{Cov}[f(X), g(X)]. \end{aligned}$$

When one of the two functions is increasing and the other is decreasing, the same proof with obvious modifications shows that $\text{Cov}[f(X), g(X)] \leq 0$. ■

Example 7.4.2 *Monotonicity in Bernstein's Approximation*

In Example 3.5.1, we considered Bernstein's proof of the Weierstrass approximation theorem. When the continuous function $f(x)$ being approximated is increasing, it is plausible that the approximating polynomial

$$\text{E}\left[f\left(\frac{S_n}{n}\right)\right] = \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k}x^k(1-x)^{n-k}.$$

is increasing as well [136]. To prove this assertion by coupling, imagine scattering n points randomly on the unit interval. If $x \leq y$ and we interpret S_n as the number of points less than or equal to x and T_n as the number of points less than or equal to y, then these two binomially distributed random variables satisfy $S_n \leq T_n$. The desired inequality

$$\text{E}\left[f\left(\frac{S_n}{n}\right)\right] \leq \text{E}\left[f\left(\frac{T_n}{n}\right)\right]$$

now follows directly from the assumption that f is increasing. ■

Our coupling proof of the convergence of an ergodic Markov chain depends on quantifying the distance between the distributions π_X and π_Y of two integer-valued random variables X and Y. One candidate distance is the total variation norm

$$\begin{aligned} \|\pi_X - \pi_Y\|_{\text{TV}} &= \sup_{A \subset \mathcal{Z}} |\Pr(X \in A) - \Pr(Y \in A)| \\ &= \frac{1}{2}\sum_k |\Pr(X = k) - \Pr(Y = k)|, \end{aligned} \tag{7.6}$$

where A ranges over all subsets of the integers $\mathcal{Z}$ [49]. Problem 28 asks the reader to check that these two definitions of the total variation norm are equivalent. The coupling bound

$$\|\pi_X - \pi_Y\|_{\text{TV}} = \sup_{A \subset \mathcal{Z}} |\Pr(X \in A) - \Pr(Y \in A)|$$

$$\begin{aligned}
&= \sup_{A\subset\mathcal{Z}} |\Pr(X\in A, X=Y)+\Pr(X\in A, X\neq Y) \\
&\quad -\Pr(Y\in A, X=Y)-\Pr(Y\in A, X\neq Y)| \quad (7.7)\\
&= \sup_{A\subset\mathcal{Z}} |\Pr(X\in A, X\neq Y)-\Pr(Y\in A, X\neq Y)| \\
&\leq \sup_{A\subset\mathcal{Z}} \mathrm{E}(1_{\{X\neq Y\}}|1_A(X)-1_A(Y)|) \\
&\leq \Pr(X\neq Y)
\end{aligned}$$

has many important applications.

In our convergence proof, we actually consider two random sequences X_n and Y_n and a random stopping time T such that $X_n = Y_n$ for all $n \geq T$. The bound

$$\Pr(X_n \neq Y_n) \quad \leq \quad \Pr(T > n)$$

suggests that we study the tail probabilities $\Pr(T > n)$. By definition, a stopping time such as T relies only on the past and present and does not anticipate the future. More formally, if $\mathcal{F}_n$ is the σ-algebra of events determined by the X_i and Y_i with $i \leq n$, then $\{T \leq n\} \in \mathcal{F}_n$. In the setting of Proposition 7.4.1,

$$\Pr(T \leq n+r \mid \mathcal{F}_n) \quad \geq \quad \epsilon \tag{7.8}$$

for some $\epsilon > 0$ and $r \geq 1$ and all n. This implies the further inequality

$$\begin{aligned}
\Pr(T > n+r) &= \mathrm{E}(1_{\{T>n+r\}}) \\
&= \mathrm{E}(1_{\{T>n\}}1_{\{T>n+r\}}) \\
&= \mathrm{E}[1_{\{T>n\}}\,\mathrm{E}(1_{\{T>n+r\}} \mid \mathcal{F}_n)] \\
&\leq \mathrm{E}[1_{\{T>n\}}(1-\epsilon)] \\
&= (1-\epsilon)\Pr(T>n),
\end{aligned}$$

which can be iterated to produce

$$\Pr(T > kr) \quad \leq \quad (1-\epsilon)^k. \tag{7.9}$$

In the last step of the iteration, we must take $\mathcal{F}_0$ to be the trivial σ-algebra consisting of the null event and the whole sample space. From inequality (7.9) it is immediately evident that $\Pr(T < \infty) = 1$.

With these preliminaries out of the way, we now turn to proving convergence based on a standard coupling argument [134, 169].

Proposition 7.4.1 *Every finite-state ergodic Markov chain has a unique equilibrium distribution π. Furthermore, the n-step transition probabilities $p_{ij}^{(n)}$ satisfy $\lim_{n\to\infty} p_{ij}^{(n)} = \pi_j$.*

Proof: Without loss of generality, we identify the states of the chain with the integers $\{1, \ldots, m\}$. From the inequality

$$\begin{aligned} p_{ij}^{(n)} &= \sum_k p_{ik} p_{kj}^{(n-1)} \\ &\leq \max_l p_{lj}^{(n-1)} \sum_k p_{ik} \\ &= \max_l p_{lj}^{(n-1)} \end{aligned}$$

involving the n-step transition probabilities, we immediately deduce that $\max_i p_{ij}^{(n)}$ is decreasing in n. Likewise, $\min_i p_{ij}^{(n)}$ is increasing in n. If

$$\lim_{n \to \infty} |p_{uj}^{(n)} - p_{vj}^{(n)}| = 0$$

for all initial states u and v, then the gap between $\lim_{n\to\infty} \min_i p_{ij}^{(n)}$ and $\lim_{n\to\infty} \max_i p_{ij}^{(n)}$ is 0. This forces the existence of $\lim_{n\to\infty} p_{ij}^{(n)} = \pi_j$, which we identify as the equilibrium distribution of the chain.

We now construct two coupled chains X_n and Y_n on $\{1, \ldots, m\}$ that individually move according to the transition matrix $P = (p_{ij})$. The X chain starts at u and the Y chain at $v \neq u$. These two chains move independently until the first epoch $T = n$ at which $X_n = Y_n$. Thereafter, they move together. The pair of coupled chains has joint transition matrix

$$p_{(ij),(kl)} = \begin{cases} p_{ik} p_{jl} & \text{if } i \neq j \\ p_{ik} & \text{if } i = j \text{ and } k = l \\ 0 & \text{if } i = j \text{ and } k \neq l \,. \end{cases}$$

By definition it is clear that the probability that the coupled chains occupy the same state at epoch r is at least as great as the probability that two completely independent chains occupy the same state at epoch r. Invoking the ergodic assumption and choosing r so that some power P^r has all of its entries bounded below by a positive constant ϵ, it follows that

$$\Pr(T \leq r \mid X_0 = u, Y_0 = v) \geq \sum_k p_{uk}^{(r)} p_{vk}^{(r)} \geq \epsilon \sum_k p_{vk}^{(r)} = \epsilon.$$

Exactly the same reasoning demonstrates that at every rth epoch the two chains have a chance of colliding of at least ϵ, regardless of their starting positions r epochs previous. In other words, inequality (7.8) holds.

Because $\Pr(T > n)$ is decreasing in n, we now harvest the bound

$$\Pr(T > n) \leq (1 - \epsilon)^{\lfloor \frac{n}{r} \rfloor}$$

from inequality (7.9). Combining this bound with the coupling inequality (7.7) yields

$$\frac{1}{2} \sum_j |p_{uj}^{(n)} - p_{vj}^{(n)}| = \|\pi_{X_n} - \pi_{Y_n}\|_{\text{TV}}$$

$$\begin{aligned} &\leq \Pr(X_n \neq Y_n) \\ &\leq \Pr(T > n) \\ &\leq (1-\epsilon)^{\lfloor \frac{n}{r} \rfloor}. \end{aligned} \tag{7.10}$$

In view of the fact that u and v are arbitrary, this concludes the proof that the $\lim_{n\to\infty} p_{ij}^{(n)} = \pi_j$ exists. ■

In the next example, we concoct a different kind of stopping time $T < \infty$ connected with a Markov chain X_n. If at epoch T the chain achieves its equilibrium distribution π and X_T is independent of T, then T is said to be a strong stationary time. When a strong stationary time exists, it gives considerable insight into the rate of convergence of the underlying chain [3, 49]. In view of the fact that X_n is at equilibrium when $T \leq n$, we readily deduce the total variation bound

$$\|\pi_{X_n} - \pi\|_{\text{TV}} \leq \Pr(T > n). \tag{7.11}$$

Example 7.4.3 *A Strong Stationary Time for Top-in Shuffling*

In top-in card shuffling, let $T-1$ be the epoch at which the original bottom card reaches the top of the deck. At epoch T the bottom card is reinserted, and the deck achieves equilibrium. For our purposes it is fruitful to view T as the sum $T = S_1 + S_2 + \cdots + S_m$ of independent geometrically distributed random variables. If $T_0 = 0$ and T_i is the first epoch at which i total cards occur under the bottom card, then the waiting time $S_i = T_i - T_{i-1}$ is geometrically distributed with success probability $\frac{i}{m}$. Because $\mathrm{E}(S_i) = \frac{m}{i}$, the strong stationary time T has mean

$$\mathrm{E}(T) = \sum_{i=1}^{m} \frac{m}{i} \approx m \ln m.$$

To exploit the bound (7.11), we consider a random variable T^* having the same distribution as T but generated by a different mechanism. Suppose we randomly drop balls into m equally likely boxes. If we let T^* be the first trial at which no box is empty, then it is clear that we can decompose $T^* = S_m^* + S_{m-1}^* + \cdots + S_1^*$, where S_i^* is the number of trials necessary to go from i empty boxes to $i-1$ empty boxes. Once again the S_i^* are independent and geometrically distributed. This perspective makes it simple to bound $\Pr(T > n)$. Indeed, if A_i is the event that box i is empty after n trials, then

$$\begin{aligned} \Pr(T > n) &= \Pr(T^* > n) \\ &\leq \sum_{i=1}^{m} \Pr(A_i) \\ &= m\left(1 - \frac{1}{m}\right)^n \\ &\leq m e^{-\frac{n}{m}}. \end{aligned} \tag{7.12}$$

Here we invoke the inequality $\ln(1-x) \leq -x$ for $x \in [0,1)$.

Returning to the top-in shuffle problem, we now combine inequality (7.12) with inequality (7.11). If $n = (1+\epsilon)m \ln m$, we deduce that

$$\|\pi_{X_n} - \pi\|_{\text{TV}} \leq me^{-\frac{(1+\epsilon)m\ln m}{m}} = \frac{1}{m^\epsilon},$$

where π is the uniform distribution on permutations of $\{1, \ldots, m\}$. This shows that if we wait much beyond the mean of T, then top-in shuffling will completely randomize the deck. Hence, the mean $\mathrm{E}(T) \approx m \ln m$ serves as a fairly sharp cutoff for equilibrium. Some statistical applications of these ideas appear in reference [131]. ■

7.5 Convergence Rates for Reversible Chains

Although Proposition 7.4.1 proves convergence to equilibrium, it does not provide the best bounds on the rate of convergence. Example 7.4.3 is instructive because it constructs a better and more natural bound. Unfortunately, it is often impossible to identify a strong stationary time. The best estimates of the rate of convergence rely on understanding the eigenstructure of the transition probability matrix P [49, 169]. We now discuss this approach for a reversible ergodic chain with equilibrium distribution π. The inner products

$$\langle u, v\rangle_{1/\pi} = \sum_i \frac{1}{\pi_i} u_i v_i, \qquad \langle u, v\rangle_\pi = \sum_i \pi_i u_i v_i$$

feature prominently in our discussion.

For the chain in question, detailed balance translates into the condition

$$\sqrt{\pi_i} p_{ij} \frac{1}{\sqrt{\pi_j}} = \sqrt{\pi_j} p_{ji} \frac{1}{\sqrt{\pi_i}}. \tag{7.13}$$

If D is the diagonal matrix with ith diagonal entry $\sqrt{\pi_i}$, then the validity of equation (7.13) for all pairs (i, j) is equivalent to the symmetry of the matrix $Q = DPD^{-1}$. Let $Q = U\Lambda U^t$ be its spectral decomposition, where U is orthogonal, and Λ is diagonal with ith diagonal entry λ_i. One can rewrite the spectral decomposition as the sum of outer products

$$Q = \sum_i \lambda_i u^i (u^i)^t$$

using the columns u^i of U. The formulas $(u^i)^t u^j = 1_{\{i=j\}}$ and

$$Q^k = \sum_i \lambda_i^k u^i (u^i)^t$$

follow immediately. The formula for Q^k in turn implies

$$P^k = \sum_i \lambda_i^k D^{-1} u^i (u^i)^t D = \sum_i \lambda_i^k w^i v^i, \tag{7.14}$$

where $v^i = (u^i)^t D$ and $w^i = D^{-1} u^i$.

Rearranging the identity $DPD^{-1} = Q = U\Lambda U^t$ yields $U^t DP = \Lambda U^t D$. Hence, the rows v^i of $V = U^t D$ are row eigenvectors of P. These vectors satisfy the orthogonality relations

$$\langle v^i, v^j \rangle_{1/\pi} = v^i D^{-2} (v^j)^t = (u^i)^t u^j = 1_{\{i=j\}}$$

and therefore form a basis of the inner product space $\ell^2_{1/\pi}$. The identity $PD^{-1}U = D^{-1}U\Lambda$ shows that the columns w^j of $W = D^{-1}U$ are column eigenvectors of P. These dual vectors satisfy the orthogonality relations

$$\langle w^i, w^j \rangle_\pi = (w^i)^t D^2 w^j = (u^i)^t u^j = 1_{\{i=j\}}$$

and therefore form a basis of the inner product space ℓ^2_π. Finally, we have the biorthogonality relations

$$v^i w^j = 1_{\{i=j\}}$$

under the ordinary inner product. The trivial rescalings $w^i = D^{-2}(v^i)^t$ and $(v^i)^t = D^2 w^i$ allow one to pass back and forth between row eigenvectors and column eigenvectors.

The distance from equilibrium in the $\ell^2_{1/\pi}$ norm bounds the total variation distance from equilibrium in the sense that

$$\|\mu - \pi\|_{\text{TV}} \leq \frac{1}{2} \|\mu - \pi\|_{1/\pi}. \tag{7.15}$$

Problem 34 asks for a proof of this fact. With the understanding that $\lambda_1 = 1$, $v^1 = \pi$, and $w^1 = \mathbf{1}$, the next proposition provides an even more basic bound.

Proposition 7.5.1 *An initial distribution μ for a reversible ergodic chain with m states satisfies*

$$\|\mu P^k - \pi\|^2_{1/\pi} = \sum_{i=2}^m \lambda_i^{2k} [(\mu - \pi) w^i]^2 \tag{7.16}$$

$$\leq \rho^{2k} \|\mu - \pi\|^2_{1/\pi}, \tag{7.17}$$

where $\rho < 1$ is the absolute value of the second-largest eigenvalue in magnitude of the transition probability matrix P.

Proof: Proposition A.2.3 of Appendix A.2 shows that $\rho < 1$. In view of the identity $\pi P = \pi$, the expansion (7.14) gives

$$\begin{aligned}
\|\mu P^k - \pi\|_{1/\pi}^2 &= \|(\mu - \pi)P^k\|_{1/\pi}^2 \\
&= (\mu - \pi)\sum_{i=1}^m \lambda_i^k w^i v^i D^{-2} \sum_{j=1}^m \lambda_j^k (v^j)^t (w^j)^t (\mu - \pi)^t \\
&= (\mu - \pi)\sum_{i=1}^m \lambda_i^{2k} w^i (w^i)^t (\mu - \pi)^t \\
&= \sum_{i=1}^m \lambda_i^{2k} \left[(\mu - \pi)w^i\right]^2.
\end{aligned}$$

The two constraints $\sum_j \pi_j = \sum_j \mu_j = 1$ clearly imply $(\mu - \pi)w^1 = 0$. Equality (7.16) follows immediately. Because all remaining eigenvalues satisfy $|\lambda_j| \leq \rho$, one can show by similar reasoning that

$$\begin{aligned}
\sum_{j=2}^m \lambda_j^{2k} \left[(\mu - \pi)w^i\right]^2 &\leq \rho^{2k} \sum_{j=1}^m \left[(\mu - \pi)w^i\right]^2 \\
&= \rho^{2k} \|\mu - \pi\|_{1/\pi}^2.
\end{aligned}$$

This validates inequality (7.17). ■

7.6 Hitting Probabilities and Hitting Times

Consider a Markov chain X_k with state space $\{1, \ldots, n\}$ and transition matrix $P = (p_{ij})$. Suppose that we can divide the states into a transient set $B = \{1, \ldots, m\}$ and an absorbing set $A = \{m + 1, \ldots, n\}$ such that $p_{ij} = 0$ for every $i \in A$ and $j \in B$ and such that every $i \in B$ leads to at least one $j \in A$. Then every realization of the chain starting in B is eventually trapped in A. It is often of interest to find the probability h_{ij} that the chain started at $i \in B$ enters A at $j \in A$. The $m \times (n - m)$ matrix of hitting probabilities $H = (h_{ij})$ can be found by solving the system of equations

$$h_{ij} = p_{ij} + \sum_{k=1}^m p_{ik} h_{kj}$$

derived by conditioning on the next state visited by the chain starting from state i. We can summarize this system as the matrix equation $H = R + QH$ by decomposing P into the block matrix

$$P = \begin{pmatrix} Q & R \\ \mathbf{0} & S \end{pmatrix},$$

where Q is $m \times m$, R is $m \times (n-m)$, and S is $(n-m) \times (n-m)$. If I is the $m \times m$ identity matrix, then the formal solution of our system of equations is $H = (I-Q)^{-1}R$. To prove that the indicated matrix inverse exists, we turn to a simple proposition.

Proposition 7.6.1 *Suppose that* $\lim_{k\to\infty} Q^k = \mathbf{0}$, *where* $\mathbf{0}$ *is the* $m \times m$ *zero matrix. Then* $(I-Q)^{-1}$ *exists and equals* $\lim_{l\to\infty} \sum_{k=0}^{l} Q^k$.

Proof: By assumption $\lim_{k\to\infty}(I - Q^k) = I$. Because the determinant function is continuous and $\det I = 1$, it follows that $\det(I - Q^k) \neq 0$ for k sufficiently large. Taking determinants in the identity

$$(I-Q)(I+Q+\cdots+Q^{k-1}) \quad = \quad I - Q^k$$

yields

$$\det(I-Q)\det(I+Q+\cdots+Q^{k-1}) \quad = \quad \det(I-Q^k).$$

Thus, $\det(I-Q) \neq 0$, and $I-Q$ is nonsingular. Finally, the power series expansion for $I-Q$ follows from taking limits in

$$I+Q+\cdots+Q^{k-1} \quad = \quad (I-Q)^{-1}(I-Q^k).$$

This completes the proof. ■

To apply Proposition 7.6.1, we need to interpret the entries of the matrix $Q^k = (q_{ij}^{(k)})$. A moment's reflection shows that

$$q_{ij}^{(k)} \quad = \quad \sum_{l=1}^{m} q_{il}^{(k-1)} p_{lj}$$

is just the probability that the chain passes from i to j in k steps. Note that the sum defining $q_{ij}^{(k)}$ stops at $l = m$ because once the chain leaves the transient states, it can never reenter them. This fact also makes it intuitively obvious that $\lim_{k\to\infty} q_{ij}^{(k)} = 0$. To verify this limit, it suffices to prove that the chain leaves the transient states after a finite number of steps. Suppose on the contrary that the chain wanders from transient state to transient state forever. In this case, the chain visits some transient state i an infinite number of times. However, i leads to an absorbing state $j > m$ along a path of positive probability. One of these visits to i must successfully take the path to j. This argument can be tightened by defining a first passage time T to the transient states and invoking inequalities (7.8) and (7.9).

In much the same way that we calculate hitting probabilities, we can calculate the mean number of epochs t_{ij} that the chain spends in transient state j prior to absorption starting from transient state i. These expectations satisfy the system of equations

$$t_{ij} \quad = \quad 1_{\{j=i\}} + \sum_{k=1}^{m} p_{ik} t_{kj},$$

which reads as $T = I + QT$ in matrix form. The solution $T = (I-Q)^{-1}$ can be used to write the mean hitting time vector t with ith entry $t_i = \sum_j t_{ij}$ as $t = (I-Q)^{-1}\mathbf{1}$, where $\mathbf{1}$ is the vector with all entries 1. Finally, if f_{ij} is the probability of ever reaching transient state j starting from transient state i, we can rearrange the identity $t_{ij} = f_{ij}t_{jj}$ for $i \neq j$ to yield the simple formula $f_{ij} = t_{ij}/t_{jj}$ for f_{ij}. The analogous identity $t_{ii} = 1 + f_{ii}t_{ii}$ gives $f_{ii} = (t_{ii} - 1)/t_{ii}$.

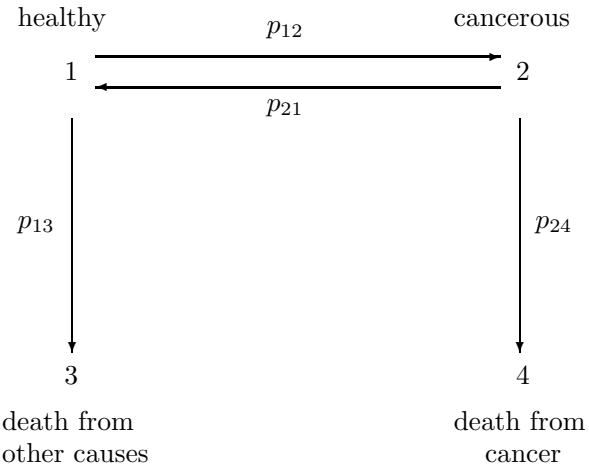


FIGURE 7.1. An Illness-Death Markov Chain

Example 7.6.1 *An Illness-Death Cancer Model*

Figure 7.1 depicts a naive Markov chain model for cancer morbidity and mortality [62]. The two transient states 1 (healthy) and 2 (cancerous) lead to the absorbing states 3 (death from other causes) and 4 (death from cancer). A brief calculation shows that

$$Q = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix}, \qquad R = \begin{pmatrix} p_{13} & 0 \\ 0 & p_{24} \end{pmatrix},$$

and

$$(I-Q)^{-1} = \frac{1}{(1-p_{11})(1-p_{22}) - p_{12}p_{21}} \begin{pmatrix} 1-p_{22} & p_{12} \\ p_{21} & 1-p_{11} \end{pmatrix}.$$

These are precisely the ingredients necessary to calculate the hitting probabilities $H = (I-Q)^{-1}R$ and mean hitting times $t = (I-Q)^{-1}\mathbf{1}$. ■

7.7 Markov Chain Monte Carlo

The Markov chain Monte Carlo (MCMC) revolution sweeping statistics is drastically changing how statisticians perform integration and summation. In particular, the Metropolis algorithm and Gibbs sampling make it straightforward to construct a Markov chain that samples from a complicated conditional distribution. Once a sample is available, then according to the ergodic theorem, any conditional expectation can be approximated by forming its corresponding sample average. The implications of this insight are profound for both classical and Bayesian statistics. As a bonus, trivial changes to the Metropolis algorithm yield simulated annealing, a general-purpose algorithm for solving difficult combinatorial optimization problems.

Our limited goal in this section is to introduce a few of the major MCMC themes. One issue of paramount importance is how rapidly the underlying chains reach equilibrium. This is the Achilles heel of the whole business and not just a mathematical nicety. Unfortunately, probing this delicate issue is scarcely possible in the confines of a brief overview. We analyze one example to give a feel for the power of coupling and spectral arguments. Readers interested in further pursuing MCMC methods and the related method of simulated annealing will enjoy the pioneering articles [69, 71, 85, 115, 142], the elementary surveys [35, 37], and the books [70, 73, 194].

7.7.1 The Hastings-Metropolis Algorithm

The Hastings-Metropolis algorithm is a device for constructing a Markov chain with a prescribed equilibrium distribution π on a given state space [85, 142]. Each step of the chain is broken into two stages, a proposal stage and an acceptance stage. If the chain is currently in state i, then in the proposal stage a new destination state j is proposed according to a probability density $q_{ij} = q(j \mid i)$. In the subsequent acceptance stage, a random number is drawn uniformly from $[0, 1]$ to determine whether the proposed step is actually taken. If this number is less than the Hastings-Metropolis acceptance probability

$$a_{ij} = \min\left\{\frac{\pi_j q_{ji}}{\pi_i q_{ij}}, 1\right\}, \tag{7.18}$$

then the proposed step is taken. Otherwise, the proposed step is declined, and the chain remains in place. Problem 41 indicates that equation (7.18) defines the most generous acceptance probability consistent with the given proposal mechanism.

Like most good ideas, the Hastings-Metropolis algorithm has undergone successive stages of abstraction and generalization. For instance, Metropolis et al. [142] considered only symmetric proposal densities with $q_{ij} = q_{ji}$. In

this case the acceptance probability reduces to

$$a_{ij} = \min\left\{\frac{\pi_j}{\pi_i}, 1\right\}. \tag{7.19}$$

In this simpler setting it is clear that any proposed destination j with $\pi_j > \pi_i$ is automatically accepted. In applying either formula (7.18) or formula (7.19), it is noteworthy that the π_i need only be known up to a multiplicative constant.

To prove that π is the equilibrium distribution of the chain constructed from the Hastings-Metropolis scheme (7.18), it suffices to check that detailed balance holds. If π puts positive weight on all points of the state space, then we must require the inequalities $q_{ij} > 0$ and $q_{ji} > 0$ to be simultaneously true or simultaneously false if detailed balance is to have any chance of holding. Now suppose without loss of generality that the fraction

$$\frac{\pi_j q_{ji}}{\pi_i q_{ij}} \leq 1$$

for some $j \neq i$. Then detailed balance follows immediately from

$$\begin{aligned} \pi_i q_{ij} a_{ij} &= \pi_i q_{ij} \frac{\pi_j q_{ji}}{\pi_i q_{ij}} \\ &= \pi_j q_{ji} \\ &= \pi_j q_{ji} a_{ji}. \end{aligned}$$

Besides checking that π is the equilibrium distribution, we should also be concerned about whether the Hastings-Metropolis chain is irreducible and aperiodic. Aperiodicity is the rule because the acceptance-rejection step allows the chain to remain in place. Problem 42 states a precise result and a counterexample. Irreducibility holds provided the entries of π are positive and the proposal matrix $Q = (q_{ij})$ is irreducible.

Example 7.7.1 *Random Walk on a Subset of the Integers*

Random walk sampling occurs when the proposal density $q_{ij} = q_{j-i}$ for some density q_k. This construction requires that the sample space be closed under subtraction. If $q_k = q_{-k}$, then the Metropolis acceptance probability (7.19) applies. ■

Example 7.7.2 *Independence Sampler*

If the proposal density satisfies $q_{ij} = q_j$, then candidate points are drawn independently of the current point. To achieve quick convergence of the chain, q_i should mimic π_i for most i. This intuition is justified by introducing the importance ratios $w_i = \pi_i/q_i$ and rewriting the acceptance probability (7.18) as

$$a_{ij} = \min\left\{\frac{w_j}{w_i}, 1\right\}. \tag{7.20}$$

It is now obvious that it is difficult to exit any state i with a large importance ratio w_i. ■

7.7.2 Gibbs Sampling

The Gibbs sampler is a special case of the Hastings-Metropolis algorithm for Cartesian product state spaces [69, 71, 73]. Suppose that each sample point $i = (i_1, \ldots, i_m)$ has m components. The Gibbs sampler updates one component of i at a time. If the component is chosen randomly and resampled conditional on the remaining components, then the acceptance probability is 1. To prove this assertion, let i_c be the uniformly chosen component, and denote the remaining components by $i_{-c} = (i_1, \ldots, i_{c-1}, i_{c+1}, \ldots, i_m)$. If j is a neighbor of i reachable by changing only component i_c, then $j_{-c} = i_{-c}$. For such a neighbor j, the proposal probability

$$q_{ij} \quad = \quad \frac{1}{m} \cdot \frac{\pi_j}{\sum_{\{k:k_{-c}=i_{-c}\}} \pi_k}$$

satisfies $\pi_i q_{ij} = \pi_j q_{ji}$, and the ratio appearing in the acceptance probability (7.18) is 1.

In contrast to random sampling of components, we can repeatedly cycle through the components in some fixed order, say $1, 2, \ldots, m$. If the transition matrix for changing component c while leaving other components unaltered is $P^{(c)}$, then the transition matrices for random sampling and sequential (or cyclic) sampling are $R = \frac{1}{m}\sum_c P^{(c)}$ and $S = P^{(1)} \cdots P^{(m)}$, respectively. Because each $P^{(c)}$ satisfies $\pi P^{(c)} = \pi$, we have $\pi R = \pi$ and $\pi S = \pi$ as well. Thus, π is the unique equilibrium distribution for R or S if either is irreducible. However as pointed out in Problem 43, R satisfies detailed balance while S ordinarily does not.

Example 7.7.3 *Ising Model*

Consider m elementary particles equally spaced around the boundary of the unit circle. Each particle c can be in one of two magnetic states—spin up with $i_c = 1$ or spin down with $i_c = -1$. The Gibbs distribution

$$\pi_i \quad \propto \quad e^{\beta \sum_d i_d i_{d+1}} \tag{7.21}$$

takes into account nearest-neighbor interactions in the sense that states like $(1, 1, 1, \ldots, 1, 1, 1)$ are favored and states like $(1, -1, 1, \ldots, 1, -1, 1)$ are shunned for $\beta > 0$. (Note that in equation (7.21) the index $m + 1$ of i_{m+1} is interpreted as 1.) There is no need to specify the normalizing constant (or partition function)

$$Z \quad = \quad \sum_i e^{\beta \sum_d i_d i_{d+1}}$$

to carry out Gibbs sampling. If we elect to resample component c, then the choices $j_c = -i_c$ and $j_c = i_c$ are made with respective probabilities

$$\begin{aligned}
\frac{e^{\beta(-i_{c-1}i_c - i_c i_{c+1})}}{e^{\beta(i_{c-1}i_c + i_c i_{c+1})} + e^{\beta(-i_{c-1}i_c - i_c i_{c+1})}} &= \frac{1}{e^{2\beta(i_{c-1}i_c + i_c i_{c+1})} + 1} \\
\frac{e^{\beta(i_{c-1}i_c + i_c i_{c+1})}}{e^{\beta(i_{c-1}i_c + i_c i_{c+1})} + e^{\beta(-i_{c-1}i_c - i_c i_{c+1})}} &= \frac{1}{1 + e^{-2\beta(i_{c-1}i_c + i_c i_{c+1})}}.
\end{aligned}$$

When the number of particles m is even, the odd-numbered particles are independent given the even-numbered particles, and vice versa. This fact suggests alternating between resampling all odd-numbered particles and resampling all even-numbered particles. Such multi-particle updates take longer to execute but create more radical rearrangements than single-particle updates. ■

7.7.3 Convergence of the Independence Sampler

For the independence sampler, it is possible to give a coupling bound on the rate of convergence to equilibrium [137]. Suppose that $X_0, X_1, \ldots$ represents the sequence of states visited by the independence sampler starting from $X_0 = x_0$. We couple this Markov chain to a second independence sampler $Y_0, Y_1, \ldots$ starting from the equilibrium distribution π. By definition, each Y_k has distribution π. The two chains are coupled by a common proposal stage and a common uniform deviate U sampled in deciding whether to accept the common proposed point. They differ in having different acceptance probabilities. If $X_n = Y_n$ for some n, then $X_k = Y_k$ for all $k \geq n$. Let T denote the random epoch when X_n first meets Y_n and the X chain attains equilibrium.

The importance ratios $w_j = \pi_j / q_j$ determine what proposed points are accepted. Without loss of generality, assume that the states of the chain are numbered $1, \ldots, m$ and that the importance ratios w_i are in decreasing order. If $X_n = x \neq y = Y_n$, then according to equation (7.18) the next proposed point is accepted by both chains with probability

$$\begin{aligned}
\sum_{j=1}^{m} q_j \min\left\{\frac{w_j}{w_x}, \frac{w_j}{w_y}, 1\right\} &= \sum_{j=1}^{m} \pi_j \min\left\{\frac{1}{w_x}, \frac{1}{w_y}, \frac{1}{w_j}\right\} \\
&\geq \frac{1}{w_1}.
\end{aligned}$$

In other words, at each trial the two chains meet with at least probability $1/w_1$. This translates into the tail probability bound

$$\Pr(T > n) \leq \left(1 - \frac{1}{w_1}\right)^n. \tag{7.22}$$

By the same type of reasoning that led to inequality (7.10), we deduce the further bound

$$\begin{aligned} \|\pi_{X_n} - \pi\|_{\text{TV}} &\leq \Pr(X_n \neq Y_n) \\ &= \Pr(T > n) \\ &\leq \left(1 - \frac{1}{w_1}\right)^n \end{aligned} \tag{7.23}$$

on the total variation distance of X_n from equilibrium.

It is interesting to compare this last bound with the bound entailed by Proposition 7.5.1. Based on our assumption that the importance ratios are decreasing, equation (7.20) shows that the transition probabilities of the independence sampler are

$$p_{ij} = \begin{cases} q_j & j < i \\ \pi_j / w_i & j > i \, . \end{cases}$$

In order for $\sum_j p_{ij} = 1$, we must set $p_{ii} = q_i + \lambda_i$, where

$$\lambda_i = \sum_{k=i}^{m} \left(q_k - \frac{\pi_k}{w_i}\right) = \sum_{k=i+1}^{m} \left(q_k - \frac{\pi_k}{w_i}\right).$$

With these formulas in mind, one can decompose the overall transition probability matrix as $P = U + \mathbf{1}q$, where $q = (q_1, \ldots, q_m)$ and U is the upper triangular matrix

$$U = \begin{pmatrix} \lambda_1 & \frac{q_2(w_2 - w_1)}{w_1} & \cdots & \cdots & \frac{q_{m-1}(w_{m-1} - w_1)}{w_1} & \frac{q_m(w_m - w_1)}{w_1} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_{m-1} & \frac{q_m(w_m - w_{m-1})}{w_{m-1}} \\ 0 & 0 & 0 & \cdots & 0 & \lambda_m \end{pmatrix}.$$

The eigenvalues of U are just its diagonal entries λ_1 through λ_m.

The reader can check that (a) $\lambda_1 = 1 - 1/w_1$, (b) the λ_i are decreasing, and (c) $\lambda_m = 0$. It turns out that P and U share most of their eigenvalues. They differ in the eigenvalue attached to the eigenvector $\mathbf{1}$ since $P\mathbf{1} = \mathbf{1}$ and $U\mathbf{1} = \mathbf{0}$. Suppose $Uv = \lambda_i v$ for some i between 1 and $m - 1$. Let us construct a column eigenvector of P with the eigenvalue λ_i. As a trial eigenvector we take $v + c\mathbf{1}$ and calculate

$$(U + \mathbf{1}q)(v + c\mathbf{1}) = \lambda_i v + qv\mathbf{1} + c\mathbf{1} = \lambda_i v + (qv + c)\mathbf{1}.$$

This is consistent with $v + c\mathbf{1}$ being an eigenvector provided we choose the constant c to satisfy $qv + c = \lambda_i c$. Because $\lambda_i \neq 1$, it is always possible to do so. The combination of Proposition 7.5.1 and inequality (7.15) gives a bound that decays at the same geometric rate $\lambda_1 = 1 - w_1^{-1}$ as the coupling bound (7.23). Thus, the coupling bound is about as good as one could hope for. Problems 44 and 45 ask the reader to flesh out our convergence arguments.

7.8 Simulated Annealing

In simulated annealing we are interested in finding the most probable state of a Markov chain [115, 163]. If this state is k, then $\pi_k > \pi_i$ for all $i \neq k$. To accentuate the weight given to state k, we can replace the equilibrium distribution π by a distribution putting probability

$$\pi_i^{(\tau)} = \frac{\pi_i^{1/\tau}}{\sum_j \pi_j^{1/\tau}}$$

on state i. Here τ is a positive parameter traditionally called temperature. With a symmetric proposal density, the distribution $\pi_i^{(\tau)}$ can be attained by running a chain with Metropolis acceptance probability

$$a_{ij} = \min\left\{\left(\frac{\pi_j}{\pi_i}\right)^{1/\tau}, 1\right\}. \tag{7.24}$$

In simulated annealing, the chain is run with τ gradually decreasing to 0 rather than with τ fixed. If τ starts out large, then in the early stages of simulated annealing, almost all proposed steps are accepted, and the chain broadly samples the state space. As τ declines, fewer unfavorable steps are taken, and the chain eventually settles on some nearly optimal state. With luck, this state is k or a state equivalent to k if several states are optimal. Simulated annealing is designed to mimic the gradual freezing of a substance into a crystalline state of perfect symmetry and hence minimum energy.

Example 7.8.1 *The Traveling Salesman Problem*

As discussed in Example 5.7.2, a salesman must visit m towns, starting and ending in his hometown. Given fixed distances d_{ij} between every pair of towns i and j, in what order should he visit the towns to minimize the length of his circuit? This problem belongs to the class of NP-complete problems; these have deterministic solutions that are conjectured to increase in complexity at an exponential rate in m.

In the simulated annealing approach to the traveling salesman problem, we assign to each permutation $\sigma = (\sigma_1, \ldots, \sigma_m)$ a cost $c_\sigma = \sum_{i=1}^m d_{\sigma_i, \sigma_{i+1}}$, where $\sigma_{m+1} = \sigma_1$. Defining $\pi_\sigma \propto e^{-c_\sigma}$ turns the problem of minimizing the cost into one of finding the most probable permutation σ. In the proposal stage of simulated annealing, we randomly select two indices $i < j$ and reverse the block of integers beginning at σ_i and ending at σ_j in the current permutation $(\sigma_1, \ldots, \sigma_m)$. Thus, if $(4, 7, 5, 2, 3, 1, 6)$ is the current permutation and indices 3 and 6 are selected, then the proposed permutation is $(4, 7, 1, 3, 2, 5, 6)$. A proposal is accepted with probability (7.24) depending on the temperature τ. In *Numerical Recipes'* [163] simulated annealing algorithm for the traveling salesman problem, τ is lowered in multiplicative

decrements of 10% after every $100m$ epochs or every $10m$ accepted steps, whichever comes first. ■

7.9 Problems

1. Take three numbers x_1, x_2, and x_3 and form the successive running averages $x_n = (x_{n-3} + x_{n-2} + x_{n-1})/3$ starting with x_4. Prove that
$$\lim_{n\to\infty} x_n = \frac{x_1 + 2x_2 + 3x_3}{6}.$$

2. A drunken knight is placed on an empty chess board and randomly moves according to the usual chess rule. Calculate the equilibrium distribution of the knight's position [152]. (Hints: Consider the squares to be nodes of a graph. Connect two squares by an edge if the knight can move from one to the other in one step. Show that the graph is connected and that 4 squares have degree 2, 8 squares have degree 3, 20 squares have degree 4, 16 squares have degree 6, and 16 squares have degree 8.)

3. Suppose you repeatedly throw a fair die and record the sum S_n of the exposed faces after n throws. Show that
$$\lim_{n\to\infty} \Pr(S_n \text{ is divisible by } 13) = \frac{1}{13}$$
by constructing an appropriate Markov chain [152].

4. Demonstrate that a finite-state Markov chain is ergodic (irreducible and aperiodic) if and only if some power P^n of the transition matrix P has all entries positive. (Hints: For sufficiency, show that if some power P^n has all entries positive, then P^{n+1} has all entries positive. For necessity, note that $p_{ij}^{(r+s+t)} \geq p_{ik}^{(r)} p_{kk}^{(s)} p_{kj}^{(t)}$, and use the number theoretic fact that the set $\{s : p_{kk}^{(s)} > 0\}$ contains all sufficiently large positive integers s if k is aperiodic. See Appendix A.1 for the requisite number theory.)

5. Consider the Cartesian product state space $A \times B$, where
$$A = \{0, 1, \ldots, a-1\}, \quad B = \{0, 1, \ldots, b-1\},$$
and a and b are positive integers. Define a Markov chain that moves from (x, y) to $(x+1 \bmod a, y)$ or $(x, y+1 \bmod b)$ with equal probability at each epoch. Show that the chain is irreducible. Also show that it is aperiodic if and only if the greatest common divisor of a and b is 1. (Hints: It helps to consider the special state $(0, 0)$. See Proposition A.1.4 of Appendix A.1.)

6. Prove that every state of an irreducible Markov chain has the same period.

7. Suppose an irreducible Markov chain has period d. Show that the states of the chain can be divided into d disjoint classes $C_0, \ldots, C_{d-1}$ such that $p_{ij} = 0$ unless $i \in C_k$ and $j \in C_l$ for $l = k + 1 \bmod d$. (Hint: Fix a state u and define $C_r = \{v : p_{uv}^{(nd+r)} > 0 \text{ for some } n \geq 0\}$.)

8. The transition matrix P of a finite Markov chain is said to be doubly stochastic if each of its column sums equals 1. Find an equilibrium distribution in this setting. Prove that symmetric transition matrices are doubly stochastic. For a nontrivial example of a doubly stochastic transition matrix, see Example 7.3.5.

9. Demonstrate that an irreducible Markov chain possesses at most one equilibrium distribution. This result applies regardless of whether the chain is finite or aperiodic. (Hints: Let $P = (p_{ij})$ be the transition matrix and π and μ be two different equilibrium distributions. Then there exist two states j and k with $\pi_j > \mu_j$ and $\pi_k < \mu_k$. For some state i choose m and n such that $p_{ji}^{(m)} > 0$ and $p_{ki}^{(n)} > 0$. If we define $Q = \frac{1}{2}P^m + \frac{1}{2}P^n$, then prove that $\pi = \pi Q$ and $\mu = \mu Q$. Furthermore, prove that strict inequality holds in the inequality

$$\begin{aligned} \|\pi - \mu\|_{\text{TV}} &= \frac{1}{2} \sum_i \Big| \sum_l (\pi_l - \mu_l) q_{li} \Big| \\ &\leq \frac{1}{2} \sum_l |\pi_l - \mu_l| \sum_i q_{li} \\ &= \|\pi - \mu\|_{\text{TV}}. \end{aligned}$$

   This contradiction gives the desired conclusion. Observe that the proof does not use the full force of irreducibility. The argument is valid for a chain with transient states provided they all can reach the designated state i.)

10. Show that Kolmogorov's criterion (7.3) implies that definition (7.4) does not depend on the particular path chosen from i to j.

11. In the Bernoulli-Laplace model, we imagine two boxes with m particles each. Among the $2m$ particles there are b black particles and w white particles, where $b + w = 2m$ and $b \leq w$. At each epoch one particle is randomly selected from each box, and the two particles are exchanged. Let Z_n be the number of black particles in the first box. Is the corresponding chain irreducible, aperiodic, and/or reversible? Show that its equilibrium distribution is hypergeometric.

12. In Example 7.3.1, show that the chain is aperiodic if and only if the underlying graph is not bipartite.

13. A random walk on a connected graph has equilibrium distribution $\pi_v = \frac{d(v)}{2m}$, where $d(v)$ is the degree of v and m is the number of edges. Let t_{uv} be the expected time the chain takes in traveling from node u to node v. If the graph is not bipartite, then the chain is aperiodic, and Example 7.3.4 shows that $t_{vv} = 1/\pi_v$. Write a recurrence relation connecting t_{vv} to the t_{uv} for nodes u in the neighborhood of v, and use the relation to demonstrate that $t_{uv} \leq 2m - d(v)$ for each such u.

14. Consider the $n!$ different permutations $\sigma = (\sigma_1, \ldots, \sigma_n)$ of the set $\{1, \ldots, n\}$ equipped with the uniform distribution $\pi_\sigma = 1/n!$ [49]. Declare a permutation ω to be a neighbor of σ if there exist two indices $i \neq j$ such that $\omega_i = \sigma_j$, $\omega_j = \sigma_i$, and $\omega_k = \sigma_k$ for $k \notin \{i, j\}$. How many neighbors does a permutation σ possess? Show how the set of permutations can be made into a reversible Markov chain using the construction of Example 7.3.1. Is the underlying graph bipartite? If we execute one step of the chain by randomly choosing two indices i and j and switching σ_i and σ_j, how can we slightly modify the chain so that it is aperiodic?

15. Consider a set of b light bulbs. At epoch n, a random subset of s light bulbs is selected. Those bulbs in the subset that are on are switched off, and those bulbs that are off are switched on. Let X_n equal the total number of on bulbs just after this random event.

    (a) Show that the stochastic process X_n is a Markov chain. What is the state space? (Hint: You may want to revise your answer after considering question (c).)

    (b) Demonstrate that the transition probability matrix has entries

    $$p_{jk} = \Pr(X_{n+1} = k \mid X_n = j) = \frac{\binom{j}{i}\binom{b-j}{s-i}}{\binom{b}{s}}.$$

    where $i = (s + j - k)/2$ must be an integer. Note that $p_{jk} > 0$ if and only if $p_{kj} > 0$.

    (c) Verify the following behavior. If s is an even integer and X_0 is even, then all subsequent X_n are even. If s is an even integer and X_0 is odd, then all subsequent X_n are odd. If s is an odd integer, then the X_n alternate between even and odd values. What is the period of the chain when s is odd? Recall that the period of state i is the greatest common divisor of the set $\{n \geq 1 : p_{ii}^{(n)} > 0\}$, where $p_{ii}^{(n)}$ is an n-step transition probability. If all states communicate, then every state has the same period.

    (d) If s is odd, then prove that all states communicate. If s is even, then prove that all even states communicate and that all odd

states communicate. (Hints: First, show that it is possible to pass in a finite number of steps from any state j to some state k with $k \leq s$. Second, show that it suffices to assume $b = s + 1$. Third, consider the paths

$$\begin{aligned} 0 \leftrightarrow s \leftrightarrow 2 \leftrightarrow s-2 \leftrightarrow 4 \leftrightarrow \cdots \leftrightarrow \lfloor \tfrac{s}{2} \rfloor \\ s+1 \leftrightarrow 1 \leftrightarrow s-1 \leftrightarrow 3 \leftrightarrow s-3 \leftrightarrow \cdots \leftrightarrow \lfloor \tfrac{s}{2} \rfloor + 1. \end{aligned}$$

Every state between 0 and $s + 1$ is visited by one of these two paths. When s is odd, the transition $\lfloor \frac{s}{2} \rfloor \leftrightarrow \lfloor \frac{s}{2} \rfloor + 1$ is possible. If this reasoning is too complicated, show how states communicate for a particular choice of s, say 4 or 5.)

(e) Verify that the unique stationary distribution π of the chain has entries

$$\pi_j = \frac{\binom{b}{j}}{2^b} \quad \text{or} \quad \pi_j = \frac{\binom{b}{j}}{2^{b-1}}.$$

(Hints: Check detailed balance. For the normalizing constant when s is even, suppose that X follows the equilibrium distribution π. If X is concentrated on the even integers, then it has generating function

$$\mathrm{E}(u^X) = \left(\frac{1}{2} + \frac{u}{2}\right)^b + \left(\frac{1}{2} - \frac{u}{2}\right)^b,$$

and if X is concentrated on the odd integers, then it has generating function

$$\mathrm{E}(u^X) = \left(\frac{1}{2} + \frac{u}{2}\right)^b - \left(\frac{1}{2} - \frac{u}{2}\right)^b.$$

Evaluate when $u = 1$.)

(f) Suppose that X follows the equilibrium distribution. Demonstrate that X has mean $\mathrm{E}(X) = \frac{b}{2}$. If s is even and X is concentrated on the even integers, then show that X has falling factorial moments

$$\mathrm{E}[(X)_k] = \begin{cases} \frac{(b)_k}{2^k} & 0 \leq k < b \\ \frac{b!}{2^b}[1 + (-1)^b] & k = b \\ 0 & k > b. \end{cases}$$

If s is even and X is concentrated on the odd integers, these factorial moments become

$$\mathrm{E}[(X)_k] = \begin{cases} \frac{(b)_k}{2^k} & 0 \leq k < b \\ \frac{b!}{2^b}[1 - (-1)^b] & k = b \\ 0 & k > b. \end{cases}$$

Let Y be binomially distributed with b trials and success probability $\frac{1}{2}$. It is interesting that $\mathrm{E}[(X)_k] = \mathrm{E}[(Y)_k]$ for $k < b$. This fact implies that X and Y have the same ordinary moments $\mathrm{E}(X^k) = \mathrm{E}(Y^k)$ for $k < b$. (Hint: See the hint to the last subproblem.)

16. In Example 7.4.1, suppose that $f(x)$ is strictly increasing and $g(x)$ is increasing. Show that $\mathrm{Cov}[f(X), g(X)] = 0$ occurs if and only if $\Pr[g(X) = c] = 1$ for some constant c. (Hint: For necessity, examine the proof of the example and show that $\mathrm{Cov}[f(X), g(X)] = 0$ entails $\Pr[g(X) = g(Y)] = 1$ and therefore $\mathrm{Var}[g(X) - g(Y)] = 0$.)

17. Consider a random graph with n nodes. Between every pair of nodes, independently introduce an edge with probability p. If $c(p)$ denotes the probability that the graph is connected, then it is intuitively clear that $c(p)$ is increasing in p. Give a coupling proof of this fact.

18. Consider a random walk on the integers $0, \ldots, m$ with transition probabilities

$$p_{ij} = \begin{cases} q_i & j = i - 1 \\ 1 - q_i & j = i + 1 \end{cases}$$

for $i = 1, \ldots, m-1$ and $p_{00} = p_{mm} = 1$. All other transition probabilities are 0. Eventually the walk gets trapped at 0 or m. Let f_i be the probability that the walk is absorbed at 0 starting from i. Show that f_i is an increasing function of the entries of $q = (q_1, \ldots, q_{m-1})$. (Hint: Let q and q^* satisfy $q_i \leq q_i^*$ for $i = 1, \ldots, m-1$. Construct coupled walks X_n and Y_n based on q and q^* such that $X_0 = Y_0 = i$ and such that at the first step $Y_1 \leq X_1$. This requires coordinating the first step of each chain. If $X_1 > Y_1$, then run the X_n chain until it reaches either m or Y_1. In the latter case, take another coordinated step of the two chains.)

19. Suppose that X follows the hypergeometric distribution

$$\Pr(X = i) = \frac{\binom{r}{i}\binom{n-r}{m-i}}{\binom{n}{m}}.$$

Let Y follow the same hypergeometric distribution except that $r+1$ replaces r. Give a coupling proof that $\Pr(X \geq k) \leq \Pr(Y \geq k)$ for all k. (Hint: Consider an urn with r red balls, 1 white ball, and $n-r-1$ black balls. If we draw m balls from the urn without replacement, then X is the number of red balls drawn, and Y is the number of red or white balls drawn.)

20. Let X be a binomially distributed random variable with n trials and success probability p. Show by a coupling argument that $\Pr(X \geq k)$ is increasing in n for fixed p and k and in p for fixed n and k.

21. Let Y be a Poisson random variable with mean λ. Demonstrate that $\Pr(Y \geq k)$ is increasing in λ for k fixed. (Hint: If $\lambda_1 < \lambda_2$, then construct coupled Poisson random variables Y_1 and Y_2 with means λ_1 and λ_2 such that $Y_1 \leq Y_2$.)

22. Let Y follow a negative binomial distribution that counts the number of failures until n successes. Demonstrate by a coupling argument that $\Pr(Y \geq k)$ is decreasing in the success probability p for k fixed.

23. Let X_1 follow a beta distribution with parameters α_1 and β_1 and X_2 follow a beta distribution with parameters α_2 and β_2. If $\alpha_1 \leq \alpha_2$ and $\alpha_1 + \beta_1 = \alpha_2 + \beta_2$, then demonstrate that $\Pr(X_1 \geq x) \leq \Pr(X_2 \geq x)$ for all $x \in [0, 1]$. How does this result carry over to the beta-binomial distribution? (Hint: Construct X_1 and X_2 from gamma distributed random variables.)

24. The random variable Y stochastically dominates the random variable X provided $\Pr(Y \leq u) \leq \Pr(X \leq u)$ for all real u. Using quantile coupling, we can construct on a common probability space probabilistic copies X_c of X and Y_c of Y such that $X_c \leq Y_c$ with probability 1. If X has distribution function $F(x)$ and Y has distribution function $G(y)$, then define $F^{[-1]}(u)$ and $G^{[-1]}(u)$ as instructed in Example 1.5.1 of Chapter 1. If U is uniformly distributed on $[0, 1]$, demonstrate that $X_c = F^{[-1]}(U)$ and $Y_c = G^{[-1]}(U)$ yield quantile couplings with the property $X_c \leq Y_c$. Problems 19 through 23 provide examples of stochastic domination.

25. Continuing Problem 24, suppose that X_1, X_2, Y_1, and Y_2 are random variables such that Y_1 dominates X_1, Y_2 dominates X_2, X_1 and X_2 are independent, and Y_1 and Y_2 are independent. Prove that $Y_1 + Y_2$ dominates $X_1 + X_2$.

26. Suppose that the random variable Y stochastically dominates the random variable X and that $f(u)$ is an increasing function of the real variable u. In view of Problem 24, prove that

$$\mathrm{E}[f(Y)] \quad \geq \quad \mathrm{E}[f(X)]$$

whenever both expectations exist. Conversely, if X and Y have this property, then show that Y stochastically dominates X.

27. Suppose the integer-valued random variable Y stochastically dominates the integer-valued random variable X. Prove the bound

$$\|\pi_X - \pi_Y\|_{\mathrm{TV}} \quad \leq \quad \mathrm{E}(Y) - \mathrm{E}(X)$$

by extending inequality (7.7). Explicitly evaluate this bound for the distributions featured in Problems 19 through 23. (Hint: According to Problem 24, one can assume that $Y \geq X$.)

28. Show that the two definitions of the total variation norm given in equation (7.6) coincide.

29. Let X have a Bernoulli distribution with success probability p and Y a Poisson distribution with mean p. Prove the total variation inequality

$$\|\pi_X - \pi_Y\|_{\text{TV}} \quad \leq \quad p^2 \tag{7.25}$$

involving the distributions π_X and π_Y of X and Y.

30. Suppose the integer-valued random variables U_1, U_2, V_1, and V_2 are such that U_1 and U_2 are independent and V_1 and V_2 are independent. Demonstrate that

$$\|\pi_{U_1+U_2} - \pi_{V_1+V_2}\|_{\text{TV}} \;\leq\; \|\pi_{U_1} - \pi_{V_1}\|_{\text{TV}} + \|\pi_{U_2} - \pi_{V_2}\|_{\text{TV}}. \tag{7.26}$$

31. A simple change of Ehrenfest's Markov chain in Example 7.3.3 renders it ergodic. At each step of the chain, flip a fair coin. If the coin lands heads, switch the chosen molecule to the other half of the box. If the coin lands tails, leave it where it is. Show that Ehrenfest's chain with holding is ergodic and converges to the binomial distribution π with m trials and success probability $\frac{1}{2}$. The rate of convergence to this equilibrium distribution can be understood by constructing a strong stationary time. As each molecule is encountered, check it off the list of molecules. Let T be the first time all m molecules are checked off. Argue that T is a strong stationary time. If $\pi^{(n)}$ is the state of the chain at epoch n, then also show that

$$\|\pi^{(n)} - \pi\|_{\text{TV}} \quad \leq \quad m\Big(1 - \frac{1}{m}\Big)^n$$

and therefore that

$$\|\pi^{(n)} - \pi\|_{\text{TV}} \quad \leq \quad e^{-c}$$

for $n = m \ln m + cm$ and $c > 0$.

32. Suppose in the Wright-Fisher model of Example 7.3.2 that each sampled a_1 allele has a chance of u of mutating to an a_2 allele and that each sampled a_2 allele has a chance of v of mutating to an a_1 allele, where the mutation rates u and v are taken from $(0, 1)$. If the number Z_n of a_1 alleles at generation n equals i, then show that Z_{n+1} is binomially distributed with success probability

$$p_i \quad = \quad \frac{i}{2m}(1-u) + \frac{2m-i}{2m}v.$$

Also prove:

(a) The p_i are increasing in i provided $u + v \leq 1$.

(b) The chain is ergodic.

(c) When $u + v = 1$, the chain is reversible with equilibrium distribution

$$\pi_j = \binom{2m}{j} v^j u^{2m-j}.$$

(d) When $u + v \neq 1$, the chain can be irreversible. For a counterexample, choose $m = 1$ and consider the path $0 \to 1 \to 2 \to 0$ and its reverse. Show that the circulation criterion can fail.

Although it is unclear what the equilibrium distribution is in general, Hua Zhou has constructed a coupling that bounds the rate of convergence of the chain to equilibrium. Assume that $u + v < 1$ and fix any two initial states $x < y$ in $\{0, 1, \ldots, 2m\}$. Let X_n be a realization of the chain starting from x and Y_n be a realization of the chain starting from y. We couple the chains by coordinated sampling of the $2m$ replacement genes at each generation. For the kth sampled gene in forming generation $n + 1$, let U_k be a uniform deviate from $(0, 1)$. If $U_k \leq p_{X_n}$, then declare the gene to be an a_1 allele in the X process. If $U_k \leq p_{Y_n}$, then declare the gene to be an a_1 allele in the Y process. Why does this imply that $X_{n+1} \leq Y_{n+1}$? Once $X_n = Y_n$ for some n, they stay coupled. In view of Problem 27, supply the reasons behind the following string of equalities and inequalities:

$$\begin{aligned} \|\pi_{X_n} - \pi_{Y_n}\|_{\text{TV}} &\leq \mathrm{E}(Y_n - X_n) \\ &= (1 - u - v)\,\mathrm{E}(Y_{n-1} - X_{n-1}) \\ &= (1 - u - v)^n (y - x) \\ &\leq 2m(1 - u - v)^n. \end{aligned}$$

Since x and y are arbitrary, this implies that the mixing time for the chain is on the order of $O\left(\frac{\ln(2m)}{u+v}\right)$ generations.

33. As another example of a strong uniform time, consider the inverse shuffling method of Reeds [49]. At every shuffle we imagine that each of c cards is assigned independently and uniformly to a top pile or a bottom pile. Hence, each pile has a binomial number of cards with mean $\frac{c}{2}$. The order of the two subpiles is kept consistent with the order of the parent pile, and in preparation for the next shuffle, the top pile is placed above the bottom pile. To keep track of the process, one can mark each card with a 0 (top pile) or 1 (bottom pile). Thus, shuffling induces an infinite binary sequence on each card that serves to track its fate. Let T denote the epoch when the first n digits for each card are unique. At T the cards reach a completely random

state where all c permutations are equally likely. Let π be the uniform distribution and π_{X_n} be the distribution of the cards after n shuffles. The probability $\Pr(T \leq n)$ is the same as the probability that c balls (digit strings) dropped independently and uniformly into 2^n boxes all wind up in different boxes. With this background in mind, deduce the bound

$$\|\pi_{X_n} - \pi\|_{\text{TV}} \quad \leq \quad 1 - \prod_{i=1}^{c-1} \left(1 - \frac{i}{2^n}\right).$$

Plot or tabulate the bound as a function of n for $c = 52$ cards. How many shuffles guarantee randomness with high probability?

34. Prove inequality (7.15) by applying the Cauchy-Schwarz inequality. Also verify that P satisfies the self-adjointness condition

$$\langle Pu, v\rangle_\pi \quad = \quad \langle u, Pv\rangle_\pi,$$

which yields a direct proof that P has only real eigenvalues.

35. Let $Z_0, Z_1, Z_2, \ldots$ be a realization of a finite-state ergodic chain. If we sample every kth epoch, then show (a) that the sampled chain $Z_0, Z_k, Z_{2k}, \ldots$ is ergodic, (b) that it possesses the same equilibrium distribution as the original chain, and (c) that it is reversible if the original chain is. Thus, based on the ergodic theorem, we can estimate theoretical means by sample averages using only every kth epoch of the original chain.

36. Consider the symmetric random walk S_n with

$$\Pr(S_{n+1} = S_n + 1) \quad = \quad \Pr(S_{n+1} = S_n - 1) \quad = \quad \frac{1}{2}.$$

Given $S_0 = i \neq 0$, let π_i be the probability that the random walk eventually hits 0. Show that

$$\begin{aligned} \pi_1 &= \frac{1}{2} + \frac{1}{2}\pi_2 \\ \pi_k &= \pi_1^k, \quad k > 0. \end{aligned}$$

Use these two equations to prove that all $\pi_i = 1$. (Hint: Symmetry.)

37. Continuing Problem 36, let μ_k be the expected waiting time for a first passage from k to 0. Show that $\mu_k = k\mu_1$ and that

$$\mu_k \quad = \quad 1 + \frac{1}{2}\mu_{k-1} + \frac{1}{2}\mu_{k+1}$$

for $k \geq 2$. Conclude from these identities that $\mu_k = \infty$ for all $k \geq 1$. Now reason that $\mu_0 = 1 + \mu_1$ and deduce that $\mu_0 = \infty$ as well.

38. Consider a random walk on the integers $\{0, 1, \ldots, n\}$. States 0 and n are absorbing in the sense that $p_{00} = p_{nn} = 1$. If i is a transient state, then the transition probabilities are $p_{i,i+1} = p$ and $p_{i,i-1} = q$, where $p + q = 1$. Verify that the hitting probabilities are

$$h_{in} = 1 - h_{i0} = \begin{cases} \frac{(\frac{q}{p})^i - 1}{(\frac{q}{p})^n - 1}, & p \neq q \\ \frac{i}{n}, & p = q \end{cases}$$

and the mean hitting times are

$$t_i = \begin{cases} \frac{n}{p-q} \frac{(\frac{q}{p})^i - 1}{(\frac{q}{p})^n - 1} - \frac{i}{p-q}, & p \neq q \\ i(n-i), & p = q. \end{cases}$$

(Hint: First argue that

$$t_i = 1 + \sum_{k=1}^{m} p_{ik} t_k$$

in the notation of Section 7.6.)

39. Arrange n points labeled $0, \ldots, n-1$ symmetrically on a circle, and imagine conducting a symmetric random walk with transition probabilities

$$p_{ij} = \begin{cases} \frac{1}{2} & j = i + 1 \bmod n \text{ or } j = i - 1 \bmod n \\ 0 & \text{otherwise}. \end{cases}$$

Thus, only transitions to nearest neighbors are allowed. Let e_k be the expected number of epochs until reaching point 0 starting at point k. Interpret e_0 as the expected number of epochs to return to 0. In finding the e_k, argue that it suffices to find $e_0, \ldots, e_m$, where $m = \lfloor \frac{n}{2} \rfloor$. Write a system of recurrence relations for the e_k, and show that the system has the solution

$$e_k = \begin{cases} n & k = 0 \\ k(n-k) & 1 \leq k \leq m. \end{cases}$$

Note that the last recurrence in the system differs depending on whether n is odd or even.

40. In the context of Section 7.6, one can consider leaving probabilities as well as hitting probabilities. Let l_{ij} be the probability of exiting the transient states from transient state j when the chain starts in transient state i. If $x_i = \sum_{k=m+1}^{n} p_{ik}$ is the exit probability from state i and $X = \text{diag}(x)$ is the $m \times m$ diagonal matrix with x_i as its ith diagonal entry, then show that $L = (I - Q)^{-1} X$, where $L = (l_{ij})$. Calculate L in the illness-death model of Section 7.6.

41. An acceptance function $a : (0, \infty) \mapsto [0, 1]$ satisfies the functional identity $a(x) = xa(1/x)$. Prove that the detailed balance condition

$$\pi_i q_{ij} a_{ij} \quad = \quad \pi_j q_{ji} a_{ji}$$

holds if the acceptance probability a_{ij} is defined by

$$a_{ij} \quad = \quad a\left(\frac{\pi_j q_{ji}}{\pi_i q_{ij}}\right)$$

in terms of an acceptance function $a(x)$. Check that the Barker function $a(x) = x/(1 + x)$ qualifies as an acceptance function and that any acceptance function is dominated by the Metropolis acceptance function in the sense that $a(x) \leq \min\{x, 1\}$ for all x.

42. The Metropolis acceptance mechanism (7.19) ordinarily implies aperiodicity of the underlying Markov chain. Show that if the proposal distribution is symmetric and if some state i has a neighboring state j such that $\pi_i > \pi_j$, then the period of state i is 1, and the chain, if irreducible, is aperiodic. For a counterexample, assign probability $\pi_i = 1/4$ to each vertex i of a square. If the two vertices adjacent to a given vertex i are each proposed with probability 1/2, then show that all proposed steps are accepted by the Metropolis criterion and that the chain is periodic with period 2.

43. Consider the Cartesian product space $\{0, 1\} \times \{0, 1\}$ equipped with the probability distribution

$$(\pi_{00}, \pi_{01}, \pi_{10}, \pi_{11}) \quad = \quad \left(\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{8}\right).$$

Demonstrate that sequential Gibbs sampling does not satisfy detailed balance by showing that $\pi_{00} s_{00,11} \neq \pi_{11} s_{11,00}$, where $s_{00,11}$ and $s_{11,00}$ are entries of the matrix S for first resampling component one and then resampling component two.

44. In our analysis of convergence of the independence sampler, we asserted that the eigenvalues $\lambda_1, \ldots, \lambda_m$ satisfied the properties: (a) $\lambda_1 = 1 - 1/w_1$, (b) the λ_i are decreasing, and (c) $\lambda_m = 0$. Verify these properties.

45. Find the row and column eigenvectors of the transition probability matrix P for the independence sampler. Show that they are orthogonal in the appropriate inner products.

46. It is known that every planar graph can be colored by four colors [32]. Design, program, and test a simulated annealing algorithm to find a four coloring of any planar graph. (Suggestions: Represent the

graph by a list of nodes and a list of edges. Assign to each node a color represented by a number between 1 and 4. The cost of a coloring is the number of edges with incident nodes of the same color. In the proposal stage of the simulated annealing solution, randomly choose a node, randomly reassign its color, and recalculate the cost. If successful, simulated annealing will find a coloring with the minimum cost of 0.)

47. A Sudoku puzzle is a 9×9 matrix, with some entries containing predefined digits. The goal is to completely fill in the matrix, using the digits 1 through 9, in such a way that each row, column, and symmetrically placed 3×3 submatrix displays each digit exactly once. In mathematical language, a completed Sudoku matrix is a Latin square subject to further constraints on the 3×3 submatrices. The initial partially filled in matrix is assumed to have a unique completion. Design, program, and test a simulated annealing algorithm to solve a Sudoku puzzle.

8
Continuous-Time Markov Chains

8.1 Introduction

This chapter introduces the subject of continuous-time Markov chains [23, 52, 59, 80, 106, 107, 118, 152]. In practice, continuous-time chains are more useful than discrete-time chains. For one thing, continuous-time chains permit variation in the waiting times for transitions between neighboring states. For another, they avoid the annoyances of periodic behavior. Balanced against these advantages is the disadvantage of a more complex theory involving linear differential equations. The primary distinction between the two types of chains is the substitution of transition intensities for transition probabilities. Once one grasps this difference, it is straightforward to formulate relevant continuous-time models. Implementing such models numerically and understanding them theoretically then require the matrix exponential function. Kendall's birth-death-immigration process, treated at the end of the chapter, involves an infinite number of states and transition intensities that depend on time.

8.2 Finite-Time Transition Probabilities

Just as with a discrete-time chain, the behavior of a continuous-time chain is described by an indexed family Z_t of random variables giving the state occupied by the chain at each time t. Now, however, the index t ranges over the nonnegative real numbers rather than the nonnegative integers. Of fun-

K. Lange, *Applied Probability,*
DOI 10.1007/978-1-4419-7165-4_8, © Springer Science+Business Media, LLC 2010

damental theoretical importance are the finite-time transition probabilities $p_{ij}(t) = \Pr(Z_{s+t} = j \mid Z_s = i)$ for all $s, t \geq 0$. We shall see momentarily how these probabilities can be found by solving a matrix differential equation.

The perspective of competing risks sharpens our intuitive understanding of how a continuous-time chain operates. Imagine that a particle executes a Markov chain by moving from state to state. If the particle is currently in state i, then each neighboring state independently beckons the particle to switch positions. The intensity of the temptation exerted by state j is the constant λ_{ij}. In the absence of competing temptations, the particle waits an exponential length of time T_{ij} with intensity λ_{ij} before moving to state j. Taking into account competing independent temptations, the particle moves at the moment $T_i = \min_{j \neq i} T_{ij}$, which is exponentially distributed with intensity $\lambda_i = \sum_{j \neq i} \lambda_{ij}$. Of course, exponentially distributed waiting times are inevitable in a Markovian model; otherwise, the intensity of leaving state i would depend on the history of waiting in i. Once the particle decides to leave i, it moves to j with probability $q_{ij} = \lambda_{ij}/\lambda_i$.

An important consequence of these assumptions is that the destination state D_i is chosen independently of the waiting time T_i. Indeed, conditioning on the value of T_{ik} gives

$$\begin{aligned}
\Pr(D_i = k, T_i \geq t) &= \Pr(T_{ik} \geq t, T_{ij} > T_{ik} \text{ for } j \neq k) \\
&= \int_t^\infty \lambda_{ik} e^{-\lambda_{ik} s} \Pr(T_{ij} > T_{ik} \text{ for } j \neq k \mid T_{ik} = s)\, ds \\
&= \int_t^\infty \lambda_{ik} e^{-\lambda_{ik} s} \prod_{j \notin \{i,k\}} e^{-\lambda_{ij} s} ds \\
&= \int_t^\infty \lambda_{ik} e^{-\lambda_i s} ds \\
&= q_{ik} e^{-\lambda_i t} \\
&= \Pr(D_i = k) \Pr(T_i \geq t).
\end{aligned}$$

Not only does this calculation establish the independence of D_i and T_i, but it also validates their claimed marginal distributions. If we ignore the times at which transitions occur, the sequence of transitions in a continuous-time chain determines a discrete-time chain with transition probability matrix $Q = (q_{ij})$.

At this juncture, it is helpful to pause and consider the nature of the finite-time transition matrix $P(t) = [p_{ij}(t)]$ for small times $t > 0$. Suppose the chain starts in state i at time 0. Because it will be in state i at time t if it never leaves during the interim, we have

$$p_{ii}(t) \geq e^{-\lambda_i t} = 1 - \lambda_i t + o(t).$$

The chain can reach a destination state $j \neq i$ if it makes a one-step transition to j sometime during $[0, t]$ and remains there for the duration of the

interval. Given that $\lambda_i q_{ij} = \lambda_{ij}$, this observation leads to the inequality

$$p_{ij}(t) \;\geq\; (1 - e^{-\lambda_i t}) q_{ij} e^{-\lambda_j t} \;=\; \lambda_{ij} t + o(t).$$

If there are only a finite number of states, the sum of these inequalities satisfies

$$1 \;=\; \sum_j p_{ij}(t) \;\geq\; 1 + o(t),$$

and therefore equality must hold in each of the participating inequalities to order $o(t)$. In view of the approximations embodied in the formula $p_{ij}(t) = \lambda_{ij} t + o(t)$, the transition intensities λ_{ij} are also termed infinitesimal transition probabilities.

8.3 Derivation of the Backward Equations

We now show that the finite-time transition probabilities $p_{ij}(t)$ satisfy a system of ordinary differential equations called the backward equations. The integral form of this system amounts to

$$p_{ij}(t) \;=\; 1_{\{j=i\}} e^{-\lambda_i t} + \int_0^t \lambda_i e^{-\lambda_i s} \sum_{k \neq i} q_{ik} p_{kj}(t-s)\, ds. \tag{8.1}$$

The first term on the right of equation (8.1) represents the probability that a particle initially in state i remains there throughout the period $[0, t]$. Of course, this is only possible when $j = i$. The integral contribution on the right of equation (8.1) involves conditioning on the time s of the first departure from state i. If state k is chosen as the destination for this departure, then the particle ends up in state j at time t with probability $p_{kj}(t-s)$.

Multiplying equation (8.1) by $e^{\lambda_i t}$ yields

$$\begin{aligned} e^{\lambda_i t} p_{ij}(t) &= 1_{\{j=i\}} + \int_0^t \lambda_i e^{\lambda_i (t-s)} \sum_{k \neq i} q_{ik} p_{kj}(t-s)\, ds \\ &= 1_{\{j=i\}} + \int_0^t \lambda_i e^{\lambda_i s} \sum_{k \neq i} q_{ik} p_{kj}(s)\, ds \end{aligned} \tag{8.2}$$

after an obvious change of variables. Because all of the finite-time transition probabilities satisfy $|p_{ik}(s)| \leq 1$ and all rows of the matrix Q satisfy $\sum_{k \neq i} q_{ik} = 1$, the integrand $\lambda_i e^{\lambda_i s} \sum_{k \neq i} q_{ik} p_{kj}(s)$ is bounded on every finite interval. It follows that both its integral and $p_{ij}(t)$ are continuous in t. Given continuity of the $p_{ij}(t)$, the integrand $\lambda_i e^{\lambda_i s} \sum_{k \neq i} q_{ik} p_{kj}(s)$ is

continuous, being the limit of a uniformly converging series of continuous functions. The fundamental theorem of calculus therefore implies that $p_{ij}(t)$ is differentiable. Taking derivatives in equation (8.2) produces

$$\lambda_i e^{\lambda_i t} p_{ij}(t) + e^{\lambda_i t} p'_{ij}(t) \quad = \quad \lambda_i e^{\lambda_i t} \sum_{k \neq i} q_{ik} p_{kj}(t).$$

After straightforward rearrangement using $\lambda_{ii} = -\lambda_i$, we arrive at the differential form

$$p'_{ij}(t) \quad = \quad \sum_k \lambda_{ik} p_{kj}(t) \tag{8.3}$$

of the backward equations.

For a chain with a finite number of states, the backward equation (8.3) can be summarized in matrix notation by introducing the two matrices $P(t) = [p_{ij}(t)]$ and $\Lambda = (\lambda_{ij})$. The backward equations in this notation become

$$\begin{aligned} P'(t) &= \Lambda P(t) \\ P(0) &= I, \end{aligned} \tag{8.4}$$

where I is the identity matrix. The solution of the initial value problem (8.4) is furnished by the matrix exponential [94, 118]

$$P(t) \quad = \quad e^{t\Lambda} \quad = \quad \sum_{k=0}^{\infty} \frac{1}{k!} (t\Lambda)^k. \tag{8.5}$$

One can check this fact formally by differentiating the series expansion (8.5) term by term. Probabilists call Λ the infinitesimal generator or infinitesimal transition matrix of the process. Because $\lambda_{ii} = -\sum_{j \neq i} \lambda_{ij}$, all row sums of Λ are 0, and the column vector $\mathbf{1}$ is an eigenvector of Λ with eigenvalue 0. The latter fact implies that the row sums of $P(t)$ are identically 1.

8.4 Equilibrium Distributions and Reversibility

A probability distribution $\pi = (\pi_i)$ on the states of a continuous-time Markov chain is a row vector whose components satisfy $\pi_i \geq 0$ for all i and $\sum_i \pi_i = 1$. If

$$\pi P(t) = \pi \tag{8.6}$$

holds for all $t \geq 0$, then π is said to be an equilibrium distribution for the chain. Written in components, the eigenvector equation (8.6) reduces

to $\sum_i \pi_i p_{ij}(t) = \pi_j$. For small t and a finite number of states, the series expansion (8.5) implies that equation (8.6) can be rewritten as

$$\pi(I + t\Lambda) + o(t) \quad = \quad \pi.$$

This approximate form makes it obvious that $\pi\Lambda = \mathbf{0}^t$ is a necessary condition for π to be an equilibrium distribution. Here $\mathbf{0}^t$ is a row vector of zeros. Multiplying equation (8.5) on the left by π shows that $\pi\Lambda = \mathbf{0}^t$ is also a sufficient condition for π to be an equilibrium distribution. In components this necessary and sufficient condition is equivalent to the balance equation

$$\sum_{j \neq i} \pi_j \lambda_{ji} \quad = \quad \pi_i \sum_{j \neq i} \lambda_{ij} \tag{8.7}$$

for all i.

In some chains it is easy to find an equilibrium distribution. For instance, when all column sums of the infinitesimal generator of a finite-state Markov chain vanish, the uniform distribution is stationary. If all of the states of a chain communicate, then the chain is said to be irreducible, and there is at most one equilibrium distribution π. If in addition the chain has a finite state space, π exists, and each row of $P(t)$ approaches π as t tends to ∞.

One particularly simple method of proving convergence to the equilibrium distribution is to construct a Liapunov function. A Liapunov function steadily declines along a trajectory of the chain until reaching its minimum at the equilibrium distribution. Let $q(t) = [q_j(t)]$ denote the distribution of the chain at time t. The relative information

$$H(t) \quad = \quad \sum_k q_k(t) \ln \frac{q_k(t)}{\pi_k} \quad = \quad \sum_k \pi_k \frac{q_k(t)}{\pi_k} \ln \frac{q_k(t)}{\pi_k}$$

furnishes one Liapunov function exploiting the strict convexity of the function $h(u) = u \ln u$ [111]. Proof that $H(t)$ is a Liapunov function hinges on the Chapman-Kolmogorov relation

$$q_k(t + d) \quad = \quad \sum_j q_j(t) p_{jk}(d)$$

for t and d nonnegative. This equation simply says that the process must pass through some intermediate state j at time t en route to state k at time $t + d$.

We now fix d and define $\alpha_{kj} = \pi_j p_{jk}(d)/\pi_k$. Because each α_{jk} is nonnegative and $\sum_j \alpha_{kj} = 1$, we deduce that

$$H(t + d) \quad = \quad \sum_k \pi_k h\left[\frac{q_k(t + d)}{\pi_k}\right]$$

$$\begin{aligned} &= \sum_k \pi_k h\left[\frac{\sum_j q_j(t)p_{jk}(d)}{\pi_k}\right] \\ &= \sum_k \pi_k h\left[\sum_j \alpha_{kj}\frac{q_j(t)}{\pi_j}\right] \\ &\leq \sum_j \sum_k \pi_k \alpha_{kj} h\left[\frac{q_j(t)}{\pi_j}\right] \qquad (8.8) \\ &= \sum_j \sum_k \pi_j p_{jk}(t) h\left[\frac{q_j(t)}{\pi_j}\right] \\ &= H(t), \end{aligned}$$

with strict inequality unless $q(t) = \pi$. Note here the implicit assumption that all entries of π are positive. This holds because all states communicate.

To prove that $\lim_{t\to\infty} q(t) = \pi$, we demonstrate that all cluster points of the trajectory $q(t)$ coincide with π. Certainly at least one cluster point exists when the number of states is finite because then $q(t)$ belongs to a compact (closed and bounded) set. Furthermore, $H(t)$ monotonically decreases to a finite limit c by the argument just presented. If $\lim_{n\to\infty} q(t_n) = \nu$, then the equality

$$q_k(t_n + d) = \sum_j q_j(t_n)p_{jk}(d)$$

implies that $\lim_{n\to\infty} q(t_n + d) = \omega$ exists as well. We now take limits in inequality (8.8) along the sequence t_n. In view of the continuity of $H(t)$ and its convergence to c, this gives

$$c = \sum_k \pi_k h\left(\frac{\omega_k}{\pi_k}\right) \leq \sum_k \pi_k h\left(\frac{\nu_k}{\pi_k}\right) = c. \qquad (8.9)$$

Rederivation of inequality (8.8) with ω_k substituted for $q_k(t + d)$ and ν_j substituted for $q_j(t)$ shows that strict inequality holds in (8.9) whenever $\nu \neq \pi$, contradicting the evident equality throughout. It follows that $\nu = \pi$ and that π is the limit of $q(t)$. For other proofs of convergence to the equilibrium distribution, see the references [118, 152].

Fortunately as pointed out in Problem 5, the annoying feature of periodicity present in the discrete-time theory disappears in the continuous-time theory. The definition and properties of reversible chains carry over directly from discrete time to continuous time provided we substitute transition intensities for transition probabilities [111]. For instance, the detailed balance condition becomes

$$\pi_i \lambda_{ij} = \pi_j \lambda_{ji} \qquad (8.10)$$

for all pairs $i \neq j$. Kolmogorov's circulation criterion for reversibility continues to hold. When it is true, the equilibrium distribution is constructed from the transition intensities exactly as in discrete time, substituting transition intensities for transition probabilities.

It helps to interpret equations (8.7) and (8.10) as probabilistic flows. Imagine a vast number of independent particles executing the same Markov chain. If we station ourselves at some state i, particles are constantly entering and leaving the state. Viewed from a distance, this particle swarm looks like a fluid flowing into and out of i. If we let the ensemble evolve, then it eventually reaches an equilibrium where the flows into and out of i match. Equation (8.7) is the quantification of this balance. With a reversible process, the flow from state i to state j must eventually match the flow from j to i; otherwise, the process reversed in time could be distinguished from the original process. Equation (8.10) is the quantification of detailed balance.

8.5 Examples

Here are a few examples of continuous-time Markov chains.

Example 8.5.1 *Oxygen Attachment to Hemoglobin*

A hemoglobin molecule has four possible sites to which oxygen (O_2) can attach. If the concentration s_o of O_2 is high compared to that of hemoglobin, then we can model the number of sites occupied on a single hemoglobin molecule as a continuous-time Markov chain [172]. Figure 8.1 depicts the model. In the figure, each arc is labeled by a transition intensity and each state by the circled number of O_2 molecules attached to the hemoglobin molecule. The forward rates $s_o k_{+j}$ incorporate the concentration of O_2. The higher the concentration of O_2, the more frequently successful collisions occur between O_2 and the hemoglobin attachment sites. Because this chain is reversible, we can calculate its equilibrium distribution starting from the reference state 0 as $\pi_i = \pi_0 s_o^i \prod_{j=1}^{i} k_{+j}/k_{-j}$. If each site operated independently, then we could postulate per site association and disassociation intensities of $s_o c_+$ and c_-, respectively. Under the independent-site hypothesis, $s_o k_{+j} = s_o(4 - j + 1)c_+$ and $k_{-j} = jc_-$. ■

Example 8.5.2 *Kimura's DNA Substitution Model*

Kimura has suggested a model for base-pair substitution in molecular evolution [113, 133]. Recall that DNA is a long double polymer constructed from the four bases (or nucleotides) adenine, guanine, cytosine, and thymine. These bases are abbreviated A, G, C, and T, respectively. Two of the bases are purines (A and G), and two are pyrimidines (C and T). The two strands of DNA form a double helix containing complementary hereditary information in the sense that A and T and C and G always pair across strands.

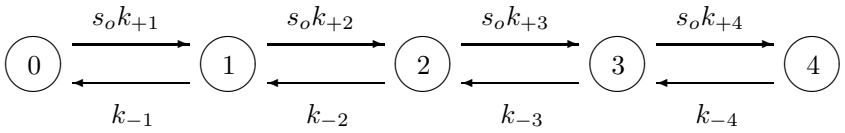


FIGURE 8.1. A Markov Chain Model for Oxygen Attachment to Hemoglobin

For example, if one strand contains the block –ACCGT– of bases, then the other strand contains the complementary block –TGGCA– of bases taken in reverse order. Thus, one is justified in following the evolutionary development of one strand and ignoring the other strand.

Kimura defines a continuous-time Markov chain with the four states A, G, C, and T that captures the evolutionary history of a species at a single position (or site) along a DNA strand. Mutations occur from time to time that change the base at the site. Let λ_{ij} be the intensity at which base i mutates to base j. Kimura radically simplifies these intensities. Let us write $i \simeq j$ if i and j are both purines or both pyrimidines and $i \not\simeq j$ if one is a purine and the other is a pyrimidine. Then Kimura assumes that

$$\lambda_{ij} = \begin{cases} \alpha & i \simeq j \\ \beta & i \not\simeq j. \end{cases}$$

These assumptions translate into the infinitesimal generator

$$\Lambda = \begin{array}{c} \\ \mathrm{A} \\ \mathrm{G} \\ \mathrm{C} \\ \mathrm{T} \end{array} \begin{array}{c} \begin{array}{cccc} \mathrm{A} & \mathrm{G} & \mathrm{C} & \mathrm{T} \end{array} \\ \begin{pmatrix} -(\alpha+2\beta) & \alpha & \beta & \beta \\ \alpha & -(\alpha+2\beta) & \beta & \beta \\ \beta & \beta & -(\alpha+2\beta) & \alpha \\ \beta & \beta & \alpha & -(\alpha+2\beta) \end{pmatrix} \end{array}.$$

Kimura's chain is reversible with the uniform distribution as its equilibrium distribution. ■

Example 8.5.3 *Incidence and Prevalence*

The distinction between the epidemiological terms "incidence" and "prevalence" can be illustrated by constructing a continuous-time Markov chain that records the numbers (M, N) of healthy and sick people traversing the diagram in Figure 8.2. The states H, S, and D correspond to a person being healthy, sick, or dead. New healthy people enter the process according to a Poisson process with rate β. Each healthy person has a rate α of converting to a sick person and a rate μ of dying. Sick people die at rate ν. In this model, prevalence is equated to $\mathrm{E}(N)$, and incidence is equated to the rate of conversion of healthy people to sick people.

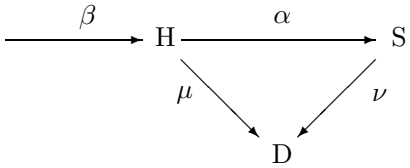


FIGURE 8.2. A Markov Chain Model for Incidence and Prevalence

The continuous-time Markov chain summarized by Figure 8.2 is irreversible. Indeed, the single-step transition from state $(m+1, n-1)$ to state (m, n) is possible while the reverse single-step transition is impossible. Nonetheless, we can write down the balance equations

$$\begin{aligned} \pi_{mn}[\beta + m(\alpha+\mu) + n\nu] &= \pi_{m-1,n}\beta + \pi_{m+1,n-1}(m+1)\alpha \\ &\quad + \pi_{m+1,n}(m+1)\mu + \pi_{m,n+1}(n+1)\nu \end{aligned} \tag{8.11}$$

for the equilibrium distribution by equating the probabilistic flow out of state (m, n) to the probabilistic flow into state (m, n). (Here we assume for the sake of simplicity that $m > 0$ and $n > 0$.) Although it is far from obvious how to solve this system of equations, it is clear on probabilistic grounds that the marginal distribution ω_m of the number of healthy people does not depend on the parameter ν. Furthermore, a little reflection suggests the balance equation

$$\omega_m m(\alpha+\mu) = \omega_{m-1}\beta.$$

From this it is trivial to deduce the Poisson distribution

$$\omega_m = \frac{\lambda_H^m}{m!}e^{-\lambda_H}, \qquad \lambda_H = \frac{\beta}{\alpha+\mu}.$$

At equilibrium, people in state H transfer to state S at rate (incidence)

$$\sum_{m=0}^{\infty} \frac{\lambda_H^m}{m!}e^{-\lambda_H} m\alpha = \alpha\lambda_H.$$

If we take the leap of faith that M and N are independent at equilibrium, then the equilibrium distribution ϕ_n of N should satisfy the balance equations

$$\phi_n n\nu = \phi_{n-1}\alpha\lambda_H$$

with Poisson solution

$$\phi_n = \frac{\lambda_S^n}{n!}e^{-\lambda_S}, \qquad \lambda_S = \frac{\alpha\beta}{\nu(\alpha+\mu)}.$$

The skeptical reader can now check that $\pi_{mn} = \omega_m \phi_n$ furnishes a solution to the balance equations (8.11). This model allows one to calculate at equilibrium the mean number of healthy people $\mathrm{E}(M) = \lambda_H$ and the mean number of sick people $\mathrm{E}(N) = \lambda_S$. The model also permits us to recover the classical equation

$$\text{prevalence} \;=\; \lambda_S \;=\; \alpha\lambda_H \frac{1}{\nu} \;=\; \text{incidence} \times \text{disease duration}$$

relating prevalence and incidence. This conservation equation from queuing theory generalizes to a broader epidemiological context [110]. Problem 18 provides an alternative method for finding the equilibrium distribution of M and N in the spirit of Example 6.8.1. ■

Example 8.5.4 *Circuit Theory*

Interesting but naive models of electrical circuits can be constructed using continuous-time Markov chains. Consider a model with $m+1$ nodes labeled $0, \ldots, m$. Nodes 0 and 1 correspond to the terminals of a battery. Node 0 has potential 0 and node 1 has potential 1. Each of the nodes can be occupied by electrons. For the sake of simplicity in discussing potential differences, we assume that an electron carries a positive rather than a negative charge. Suppose node j has a capacity of e_j electrons. The number of electrons present at node j is given by a random variable $n_j(t)$ at time t. At nodes 0 and 1 we assume that $n_0(t) = 0$ and $n_1(t) = e_1$. Provided we define appropriate infinitesimal transition probabilities, the random count vector $\mathbf{n}(t) = [n_0(t), n_1(t), \ldots, n_m(t)]$ constitutes a continuous-time Markov chain [111].

Transitions of this chain correspond to the transfer of electrons between pairs of nodes from $2, \ldots, m$, the absorption of an electron at node 0 from one of the nodes $2, \ldots, m$, and the introduction of an electron into one of the nodes $2, \ldots, m$ from node 1. Electrons absorbed by node 0 are immediately passed to the battery so that the count $n_0(t)$ remains at 0 for all time. Likewise, electrons leaving node 1 are immediately replaced by the battery so that $n_1(t)$ remains at e_1 for all time. Provided we refer to electron absorptions and introductions as transfer events, we can devise a useful model by giving the infinitesimal transfer rates between states. These will be phrased in terms of the conductance $c_{jk} = c_{kj}$ between two nodes j and k. The reciprocal of a conductance is a resistance. If we imagine the two nodes connected by a wire, then conductance indicates the mobility of the electrons through the wire. In the absence of a wire between the nodes, the conductance is 0. The transfer rate λ_{jk} between nodes j and k should incorporate the conductance c_{jk}, the possible saturation of each node by electrons, and the fact that electrons repel. The particular choice

$$\lambda_{jk} \;=\; c_{jk} \frac{n_j}{e_j} \frac{e_k - n_k}{e_k}$$

succinctly captures these requirements.

To monitor the mean number of electrons at node j, we derive a differential equation involving the transfer of electrons during a short time interval of duration s. Conditioning on the electron counts at time t, we find that

$$\begin{aligned}
& \mathrm{E}[n_j(t+s) - n_j(t)] \\
= \; & \mathrm{E}\{\mathrm{E}[n_j(t+s) - n_j(t) \mid \mathbf{n}(t)]\} \\
= \; & \sum_{k \neq j} \mathrm{E}\left\{ \mathrm{E}\left[c_{kj} \Big(\frac{e_j - n_j(t)}{e_j} \frac{n_k(t)}{e_k} - \frac{n_j(t)}{e_j} \frac{e_k - n_k(t)}{e_k} \Big) s \;\Big|\; \mathbf{n}(t) \right] \right\} + o(s) \\
= \; & \sum_{k \neq j} \mathrm{E}\left[c_{kj} \Big(\frac{n_k(t)}{e_k} - \frac{n_j(t)}{e_j} \Big) \right] s + o(s).
\end{aligned}$$

Forming the corresponding difference quotient and sending s to 0 give the differential equation

$$\frac{d}{dt} \mathrm{E}[n_j(t)] \;=\; \sum_{k \neq j} c_{kj}[p_k(t) - p_j(t)],$$

where $p_j(t) = \mathrm{E}[n_j(t)/e_j]$ and $p_k(t) = \mathrm{E}[n_k(t)/e_k]$ are the potentials at nodes j and k. The term $c_{kj}[p_k(t) - p_j(t)]$ represents the current flow from node j to node k and summarizes Ohm's law. At equilibrium, the mean number of electrons entering and leaving node j is 0. This translates into Kirchhoff's law

$$0 \;=\; \frac{d}{dt} \mathrm{E}[n_j(t)] \;=\; \sum_{k \neq j} c_{kj}(p_k - p_j),$$

where $p_k(t) = p_k$ and $p_j(t) = p_j$ are constant. An alternative Markov chain model for current flow is presented in reference [52]. ■

8.6 Calculation of Matrix Exponentials

From the definition of the matrix exponential e^A, it is easy to deduce that it is continuous in A and satisfies $e^{A+B} = e^A e^B$ whenever $AB = BA$. It is also straightforward to check the differentiability condition

$$\frac{d}{dt} e^{tA} \;=\; A e^{tA} \;=\; e^{tA} A.$$

Proofs of these facts depend on the introduction of vector and matrix norms. Of more practical importance is how one actually calculates e^{tA} [145]. In some cases it is possible to do so analytically. For instance, if u and v are column vectors with the same number of components, then

$$e^{suv^t} \;=\; \begin{cases} I + suv^t & \text{if } v^t u = 0 \\ I + \frac{e^{sv^t u} - 1}{v^t u} uv^t & \text{otherwise.} \end{cases}$$

This follows from the formula $(uv^t)^i = (v^t u)^{i-1} uv^t$. The special case where $u = (-\alpha, \beta)^t$ and $v = (1, -1)^t$ permits explicit calculation of the finite-time transition matrix

$$P(s) = \exp\left[s\begin{pmatrix} -\alpha & \alpha \\ \beta & -\beta \end{pmatrix}\right]$$

for a two-state Markov chain.

If A is a diagonalizable $n \times n$ matrix, then we can write $A = TDT^{-1}$ for D a diagonal matrix with ith diagonal entry ρ_i. Here ρ_i is an eigenvalue of A with eigenvector equal to the ith column of T. Because $A = TDT^{-1}$, we find that $A^2 = TDT^{-1}TDT^{-1} = TD^2T^{-1}$ and in general $A^i = TD^iT^{-1}$. Hence,

$$\begin{aligned} e^{tA} &= \sum_{i=0}^{\infty} \frac{1}{i!}(tA)^i \\ &= \sum_{i=0}^{\infty} \frac{1}{i!} T(tD)^i T^{-1} \\ &= Te^{tD}T^{-1}, \end{aligned} \tag{8.12}$$

where

$$e^{tD} = \begin{pmatrix} e^{\rho_1 t} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{\rho_n t} \end{pmatrix}.$$

Equation (8.12) suggests that the behavior of e^{tA} is determined by its dominant eigenvalue. This is the eigenvalue with largest real part. Suppose for the sake of argument that a dominant eigenvalue exists and is real. If the dominant eigenvalue is negative, then e^{tA} will tend to the zero matrix as t tends to ∞. If the dominant eigenvalue is positive, then usually the entries of e^{tA} will diverge. Finally, if the dominant eigenvalue is 0, then e^{tA} may converge to a constant matrix. As demonstrated in Section 8.4, the infinitesimal generator of an irreducible chain falls in this third category. Appendix A.2 develops these ideas rigorously.

Even if we cannot calculate e^{tA} analytically, we can usually do so numerically [92]. For instance when $t > 0$ is small, we can approximate e^{tA} by the truncated series $\sum_{i=0}^{n}(tA)^i/i!$ for n small. For larger t such truncation can lead to serious errors. If the truncated expansion is sufficiently accurate for all $t \leq c$, then for arbitrary t one can exploit the property $e^{(s+t)A} = e^{sA}e^{tA}$ of the matrix exponential. Thus, if $t > c$, take the smallest positive integer k such that $2^{-k}t \leq c$ and approximate $e^{2^{-k}tA}$ by the truncated series. Applying the multiplicative property, we can compute e^{tA} by squaring $e^{2^{-k}tA}$, squaring the result $e^{2^{-k+1}tA}$, squaring the result of this, and so forth, a total of k times.

Problem 3 features a method of computing matrix exponentials specifically tailored to infinitesimal generators. This uniformization technique is easy to implement numerically. The uniformization formula of the problem also demonstrates the obvious fact that all entries of the finite-time transition matrix $P(t) = e^{t\Lambda}$ are nonnegative.

Example 8.6.1 *Application to Kimura's Model*

Because the infinitesimal generator Λ in Kimura's model is a symmetric matrix, it has only real eigenvalues. These are 0, -4β, $-2(\alpha+\beta)$, and $-2(\alpha+\beta)$. The reader can check that the four corresponding eigenvectors

$$\frac{1}{2}\begin{pmatrix}1\\1\\1\\1\end{pmatrix}, \quad \frac{1}{2}\begin{pmatrix}-1\\-1\\1\\1\end{pmatrix}, \quad \frac{1}{\sqrt{2}}\begin{pmatrix}0\\0\\-1\\1\end{pmatrix}, \quad \frac{1}{\sqrt{2}}\begin{pmatrix}-1\\1\\0\\0\end{pmatrix}$$

are orthogonal unit vectors. Therefore, the matrix T constructed by concatenating these vectors is orthogonal with inverse T^t. Equation (8.12) gives

$$\begin{aligned} e^{t\Lambda} &= \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & 0 & -\frac{1}{\sqrt{2}} \\ \frac{1}{2} & -\frac{1}{2} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{\sqrt{2}} & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & e^{-4\beta t} & 0 & 0 \\ 0 & 0 & e^{-2(\alpha+\beta)t} & 0 \\ 0 & 0 & 0 & e^{-2(\alpha+\beta)t} \end{pmatrix} \\ &\quad \times \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 \end{pmatrix}. \end{aligned}$$

From this representation, we can calculate typical entries such as

$$\begin{aligned} p_{AA}(t) &= \frac{1}{4} + \frac{1}{4}e^{-4\beta t} + \frac{1}{2}e^{-2(\alpha+\beta)t} \\ p_{AG}(t) &= \frac{1}{4} + \frac{1}{4}e^{-4\beta t} - \frac{1}{2}e^{-2(\alpha+\beta)t} \\ p_{AC}(t) &= \frac{1}{4} - \frac{1}{4}e^{-4\beta t}. \end{aligned}$$

In fact, all entries of the finite-time transition matrix take one of these three forms. ■

8.7 Kendall's Birth-Death-Immigration Process

In this section we tackle a continuous-time Markov chain important in biological and chemical applications. This chain has nonconstant transition intensities and an infinite number of states. As a preliminary to understanding Kendall's birth-death-immigration process, we derive nonrigorously the forward equations governing such a chain X_t. Let $p_{ij}(t)$ be the probability that the chain is in state j at time t given it was in state i at time 0. Our point of departure is the Chapman-Kolmogorov relation

$$\begin{aligned} p_{ij}(t+h) &= p_{ij}(t)\Pr(X_{t+h}=j \mid X_t=j) \\ &\quad + \sum_{k\neq j} p_{ik}(t)\Pr(X_{t+h}=j \mid X_t=k). \end{aligned} \tag{8.13}$$

We now assume that

$$\begin{aligned} \Pr(X_{t+h}=j \mid X_t=j) &= 1-\lambda_j(t)h+o(h) \\ \Pr(X_{t+h}=j \mid X_t=k) &= \lambda_{kj}(t)h+o(h) \end{aligned} \tag{8.14}$$

for continuous intensities $\lambda_{kj}(t)$, that only a finite number of the $\lambda_{jk}(t)$ differ from 0 for a given j, and that $\lambda_j(t)=\sum_{k\neq j}\lambda_{jk}(t)$. Inserting the approximations (8.14) into the time-inhomogeneous Chapman-Kolmogorov relation (8.13) and rearranging terms yields the difference quotient

$$\frac{p_{ij}(t+h)-p_{ij}(t)}{h} = -p_{ij}(t)\lambda_j(t)+\sum_{k\neq j}p_{ik}(t)\lambda_{kj}(t)+\frac{o(h)}{h}.$$

Sending h to 0 therefore gives the system of forward equations

$$\frac{d}{dt}p_{ij}(t) = -p_{ij}(t)\lambda_j(t)+\sum_{k\neq j}p_{ik}(t)\lambda_{kj}(t) \tag{8.15}$$

with the obvious initial conditions $p_{ij}(0)=1_{\{i=j\}}$.

In Kendall's birth-death-immigration process, X_t counts the number of particles at time t. Particles are of a single type and independently die and reproduce. The death rate per particle is $\mu(t)$, and the birth rate per particle is $\alpha(t)$. Reproduction occurs throughout the lifetime of a particle, not at its death. A Poisson process with intensity $\nu(t)$ feeds new particles into the process. Each new immigrant starts an independently evolving clan of particles. If there are initially i particles, then the forward equations can be summarized as

$$\frac{d}{dt}p_{i0}(t) = -\nu(t)p_{i0}(t)+\mu(t)p_{i1}(t)$$

and

$$\begin{aligned} \frac{d}{dt}p_{ij}(t) &= -[\nu(t)+j\alpha(t)+j\mu(t)]p_{ij}(t)+[\nu(t)+(j-1)\alpha(t)]p_{i,j-1}(t) \\ &\quad +(j+1)\mu(t)p_{i,j+1}(t) \end{aligned}$$

for $j > 0$. Note here the assumption that each birth involves a single daughter particle.

To better understand this infinite system of coupled ordinary differential equations, we define the generating function

$$G(s,t) = \sum_{j=0}^{\infty} p_{ij}(t)s^j$$

for $s \in [0,1]$. If we multiply the jth forward equation by s^j and sum on j, then we find that

$$\begin{aligned}
\frac{\partial}{\partial t}G(s,t) &= \sum_{j=0}^{\infty} \frac{d}{dt}p_{ij}(t)s^j \\
&= -\nu(t)\sum_{j=0}^{\infty} p_{ij}(t)s^j - [\alpha(t)+\mu(t)]s\sum_{j=1}^{\infty} jp_{ij}(t)s^{j-1} \\
&\quad + \nu(t)s\sum_{j=1}^{\infty} p_{i,j-1}(t)s^{j-1} + \alpha(t)s^2\sum_{j=1}^{\infty}(j-1)p_{i,j-1}(t)s^{j-2} \\
&\quad + \mu(t)\sum_{j=0}^{\infty}(j+1)p_{i,j+1}(t)s^j \\
&= -\nu(t)G(s,t) - [\alpha(t)+\mu(t)]s\frac{\partial}{\partial s}G(s,t) + \nu(t)sG(s,t) \\
&\quad + \alpha(t)s^2\frac{\partial}{\partial s}G(s,t) + \mu(t)\frac{\partial}{\partial s}G(s,t).
\end{aligned}$$

Collecting terms yields the partial differential equation

$$\frac{\partial}{\partial t}G(s,t) = [\alpha(t)s-\mu(t)](s-1)\frac{\partial}{\partial s}G(s,t) + \nu(t)(s-1)G(s,t) \quad (8.16)$$

with the initial condition $G(s,0) = s^i$.

Before solving equation (8.16), it is worth solving for the mean number of particles $m_i(t)$ at time t. In view of the fact that $m_i(t) = \frac{\partial}{\partial s}G(s,t)|_{s=1}$, we can differentiate equation (8.16) and derive the ordinary differential equation

$$\begin{aligned}
\frac{d}{dt}m_i(t) &= \frac{\partial^2}{\partial s \partial t}G(1,t) \\
&= [\alpha(t)-\mu(t)]\frac{\partial}{\partial s}G(1,t) + \nu(t)G(1,t) \\
&= [\alpha(t)-\mu(t)]m_i(t) + \nu(t)
\end{aligned}$$

with the initial condition $m_i(0) = i$. To solve this first-order ordinary differential equation, we multiply both sides by $e^{w(t)}$, where

$$w(t) = \int_0^t [\mu(\tau)-\alpha(\tau)]\,d\tau.$$

This action produces

$$\frac{d}{dt}\Big[m_i(t)e^{w(t)}\Big] \;=\; \nu(t)e^{w(t)}.$$

Integrating and rearranging then yields the solution

$$m_i(t) \;=\; ie^{-w(t)} + \int_0^t \nu(\tau)e^{w(\tau)-w(t)}\,d\tau.$$

The special case where $\alpha(t)$, $\mu(t)$, and $\nu(t)$ are constant simplifies to

$$m_i(t) \;=\; ie^{(\alpha-\mu)t} + \frac{\nu}{\alpha-\mu}\Big[e^{(\alpha-\mu)t} - 1\Big]. \tag{8.17}$$

When $\alpha < \mu$, the forces of birth and immigration eventually balance the force of death, and the process reaches equilibrium. The equilibrium distribution has mean $\lim_{t\to\infty} m_i(t) = \nu/(\mu-\alpha)$.

Example 8.7.1 *Inhomogeneous Poisson Process*

If we take $i = 0$, $\alpha(t) = 0$, and $\mu(t) = 0$, then Kendall's process coincides with an inhomogeneous Poisson process. The reader can check that the partial differential equation (8.16) reduces to

$$\frac{\partial}{\partial t}G(s,t) \;=\; \nu(t)(s-1)G(s,t)$$

with solution

$$G(s,t) \;=\; e^{-(1-s)\int_0^t \nu(\tau)d\tau}.$$

From $G(s,t)$ we reap the Poisson probabilities

$$p_{0j}(t) \;=\; \frac{\Big(\int_0^t \nu(\tau)\,d\tau\Big)^j}{j!}e^{-\int_0^t \nu(\tau)\,d\tau}.$$

■

Example 8.7.2 *Inhomogeneous Pure Death Process*

When $i > 0$, $\alpha(t) = 0$, and $\nu(t) = 0$, Kendall's process represents a pure death process. The partial differential equation (8.16) reduces to

$$\frac{\partial}{\partial t}G(s,t) \;=\; -\mu(t)(s-1)\frac{\partial}{\partial s}G(s,t)$$

with solution

$$G(s,t) \;=\; 1 - e^{-\int_0^t \mu(\tau)\,d\tau} + se^{-\int_0^t \mu(\tau)\,d\tau}$$

if $i = 1$. In other words, a single initial particle is still alive at time t with probability $\exp[-\int_0^t \mu(\tau)\,d\tau]$. Another interpretation of this process is possible if t is taken as the distance traveled by a particle through some attenuating medium with attenuation coefficient $\mu(t)$. Death corresponds to the particle being stopped. The solution for general $i > 0$ is given by the binomial generating function

$$G(s,t) = \left[1 - e^{-\int_0^t \mu(\tau)\,d\tau} + se^{-\int_0^t \mu(\tau)\,d\tau}\right]^i$$

because all i particles behave independently. ■

Example 8.7.3 *Inhomogeneous Pure Birth Process*

When $i > 0$, $\mu(t) = 0$, and $\nu(t) = 0$, Kendall's process represents a pure birth process. The partial differential equation (8.16) becomes

$$\frac{\partial}{\partial t}G(s,t) = \alpha(t)s(s-1)\frac{\partial}{\partial s}G(s,t)$$

with solution

$$G(s,t) = \left\{\frac{se^{-\int_0^t \alpha(\tau)\,d\tau}}{1 - s\left[1 - e^{-\int_0^t \alpha(\tau)\,d\tau}\right]}\right\}^i,$$

which is the generating function of a negative binomial distribution with success probability $\exp[-\int_0^t \alpha(\tau)\,d\tau]$ and required number of successes i. When $i = 1$,

$$p_{1j}(t) = e^{-\int_0^t \alpha(\tau)\,d\tau}\left[1 - e^{-\int_0^t \alpha(\tau)\,d\tau}\right]^{j-1}$$

is just the probability that a single ancestral particle generates $j-1$ descendant particles. If we reverse time in this process, then births appear to be deaths, and $p_{1j}(t)$ coincides with the probability that all $j-1$ descendant particles present at time t die before time 0 and the ancestral particle lives throughout the time interval. ■

8.8 Solution of Kendall's Equation

Remarkably enough, the dynamics of Kendall's birth-death-immigration process can be fully specified by solving equation (8.16). Let us begin by supposing that the immigration rate $\nu(t)$ is identically 0 and that $X_0 = 1$. In this situation, $G(s,t)$ is given by the formidable expression

$$G(s,t) = \frac{e^{w(t)} - (s-1)[\int_0^t \alpha(\tau)e^{w(\tau)}d\tau - 1]}{e^{w(t)} - (s-1)\int_0^t \alpha(\tau)e^{w(\tau)}d\tau} \tag{8.18}$$

$$= 1 + \frac{1}{\frac{e^{w(t)}}{s-1} - \int_0^t \alpha(\tau) e^{w(\tau)} d\tau},$$

where $w(t) = \int_0^t [\mu(\tau) - \alpha(\tau)]\, d\tau$. It follows from (8.18) that $G(s,0) = s$, consistent with starting with a single particle.

To find $G(s,t)$, we consider a curve $s(t)$ in (s,t) space parameterized by t and determined implicitly by the relation $G(s,t) = s_0$, where s_0 is some constant in [0,1]. Differentiating $G(s,t) = s_0$ with respect to t produces

$$\frac{\partial}{\partial t} G(s,t) + \frac{\partial}{\partial s} G(s,t) \frac{ds}{dt} = 0.$$

Comparing this equation to equation (8.16), we conclude that

$$\frac{ds}{dt} = (\alpha s - \mu)(1-s) = [\alpha - \mu + \alpha(s-1)](1-s), \tag{8.19}$$

where we have omitted the dependence of the various functions on t.

If we let $u = 1 - s$ and $r = \mu - \alpha$, then the ordinary differential equation (8.19) is equivalent to

$$\frac{du}{dt} = (r + \alpha u) u.$$

The further transformation $v = \ln u$ gives

$$\frac{dv}{dt} = r + \alpha e^v,$$

which in turn yields

$$\frac{d}{dt}(v - w) = \alpha e^v = \alpha e^{v-w} e^w$$

for $w(t) = \int_0^t r(\tau)\, d\tau$. Once we write this last equation as

$$\frac{d}{dt} e^{-(v-w)} = -\alpha e^w,$$

the solution

$$e^{-v(t)+w(t)} - e^{-v(0)+w(0)} = -\int_0^t \alpha(\tau) e^{w(\tau)} d\tau$$

is obvious. When we impose the initial condition $s(0) = s_0$ and recall the definitions of v and u, it follows that

$$\frac{e^{w(t)}}{1-s} - \frac{1}{1-s_0} = -\int_0^t \alpha(\tau) e^{w(\tau)} d\tau.$$

A final rearrangement now gives

$$s_0 = 1 + \frac{1}{\frac{e^{w(t)}}{s-1} - \int_0^t \alpha(\tau) e^{w(\tau)} d\tau},$$

which validates the representation (8.18) of $G(s,t) = s_0$.

Before adding the complication of immigration, let us point out that the generating function (8.18) collapses to the appropriate expression for $G(s,t)$ when the process is a pure birth process or a pure death process. It is also noteworthy that $G(s,t)^i$ is the generating function of X_t when $X_0 = i$ rather than $X_0 = 1$. This is just a manifestation of the fact that particles behave independently in the model.

The key to including immigration is the observation that when an immigrant particle arrives, it generates a clan of particles similar to the clan of particles issuing from a particle initially present. However, the clan originating from the immigrant has less time to develop. This suggests that we consider the behavior of Kendall's birth-death process starting with a single particle at some later time $u > 0$. Denote the generating function associated with this delayed process by $G(s,t,u)$ for $t \geq u$. Our discussion above indicates that $G(s,t,u)$ is given by formula (8.18), provided we replace 0 by u in the lower limits of integration in the formula and in the definition of the function $w(t)$.

If we now assume $X_0 = 0$, then the generating function $H(s,t)$ of X_t in the presence of immigration is given by

$$H(s,t) = \exp\left\{ \int_0^t [G(s,t,u) - 1]\nu(u)\, du \right\}. \tag{8.20}$$

We will prove this by constructing an appropriate marked Poisson process and appealing to Campbell's formula (6.18). Recall that immigrant particles arrive according to a Poisson process with intensity $\nu(t)$. Now imagine marking a new immigrant particle at time u by the size of the clan y it generates at the subsequent time t. The random marked points (U, Y) constitute a marked Poisson process Π with intensity $p(y \mid u)\nu(u)$, where $p(y \mid u)$ is the conditional discrete density of Y given the immigration time u. Thus, formula (6.18) implies that $X_t = \sum_{(u,y)\in\Pi} y$ has generating function

$$H(s,t) = \exp\left\{ \int \sum_y (s^y - 1) p(y \mid u)\nu(u)\, du \right\}.$$

In view of the identities $\sum_y p(y \mid u) = 1$ and $\sum_y s^y p(y \mid u) = G(s,t,u)$, this proves equation (8.20). Because particles behave independently, if initially $X_0 = i$ instead of $X_0 = 0$, then the generating function of X_t is

$$\mathrm{E}(s^{X_t}) = G(s,t,0)^i \exp\left\{ \int_0^t [G(s,t,u) - 1]\nu(u)\, du \right\}. \tag{8.21}$$

In the special case where birth, death, and immigration rates are constant, the above expression fortunately simplifies. Thus, the two explicit formulas

$$\begin{aligned}\int_u^t (\mu-\alpha)\,d\tau &= (\mu-\alpha)(t-u)\\ \int_u^t \alpha e^{\int_u^\tau (\mu-\alpha)d\eta} d\tau &= \frac{\alpha}{\mu-\alpha}[e^{(\mu-\alpha)(t-u)}-1]\end{aligned}$$

lead to

$$\begin{aligned}G(s,t,u) &= 1+\frac{1}{\frac{e^{(\mu-\alpha)(t-u)}}{s-1}-\frac{\alpha}{\mu-\alpha}[e^{(\mu-\alpha)(t-u)}-1]}\\ &= 1+\frac{(s-1)(\mu-\alpha)e^{(\alpha-\mu)(t-u)}}{\mu-\alpha s+\alpha(s-1)e^{(\alpha-\mu)(t-u)}},\end{aligned}$$

which in turn implies

$$\begin{aligned}\exp\left\{\int_0^t [G(s,t,u)-1]\nu\,du\right\} &= e^{\frac{\nu}{\alpha}\int_0^t \frac{d}{du}\ln(\mu-\alpha s+\alpha(s-1)e^{(\alpha-\mu)(t-u)})\,du}\\ &= \left(\frac{\mu-\alpha}{\mu-\alpha s+\alpha(s-1)e^{(\alpha-\mu)t}}\right)^{\frac{\nu}{\alpha}}.\end{aligned}$$

From these pieces the full generating function (8.21) can be assembled.

8.9 Problems

1. Let U, V, and W be independent exponentially distributed random variables with intensities λ, μ, and ν, respectively. Consider the random variables $X = \min\{U, W\}$ and $Y = \min\{V, W\}$. Demonstrate that X and Y are exponentially distributed. What are their means? The bivariate distribution of (X, Y) is interesting. Prove that its right-tail probability satisfies

$$\Pr(X \geq x, Y \geq y) = e^{-\lambda x-\mu y-\nu\max\{x,y\}}.$$

By invoking the notion of competing risks, show that

$$\begin{aligned}\Pr(X<Y) &= \frac{\lambda}{\lambda+\mu+\nu}\\ \Pr(Y<X) &= \frac{\mu}{\lambda+\mu+\nu}\\ \Pr(X=Y) &= \frac{\nu}{\lambda+\mu+\nu}.\end{aligned}$$

The bivariate distribution of (X, Y) possesses a density $f(x, y)$ off the line $y = x$. Show that

$$f(x,y) = \begin{cases} \lambda(\mu+\nu)e^{-\lambda x-(\mu+\nu)y} & x < y \\ \mu(\lambda+\nu)e^{-(\lambda+\nu)x-\mu y} & x > y. \end{cases}$$

Finally, demonstrate that

$$\text{Cov}(X,Y) = \frac{\nu}{(\lambda+\nu)(\mu+\nu)(\lambda+\mu+\nu)}.$$

(Hint: For the covariance, it helps to condition on W and use the identity $\text{E}(X \mid W) = (1 - e^{-\lambda W})/\lambda$.)

2. Let $\Lambda = (\lambda_{ij})$ be an $m \times m$ matrix and $\pi = (\pi_i)$ be a $1 \times m$ row vector. Show that the equality $\pi_i\lambda_{ij} = \pi_j\lambda_{ji}$ is true for all pairs (i, j) if and only if $\text{diag}(\pi)\Lambda = \Lambda^t\text{diag}(\pi)$, where $\text{diag}(\pi)$ is a diagonal matrix with ith diagonal entry π_i. Now suppose Λ is an infinitesimal generator with equilibrium distribution π. If $P(t) = e^{t\Lambda}$ is its finite-time transition matrix, then show that detailed balance $\pi_i\lambda_{ij} = \pi_j\lambda_{ji}$ for all pairs (i, j) is equivalent to finite-time detailed balance $\pi_i p_{ij}(t) = \pi_j p_{ji}(t)$ for all pairs (i, j) and times $t \geq 0$.

3. Suppose that Λ is the infinitesimal generator of a continuous-time finite-state Markov chain, and let $\mu \geq \max_i \lambda_i$. If $R = I + \mu^{-1}\Lambda$, then prove that R has nonnegative entries and that

$$S(t) = \sum_{i=0}^{\infty} e^{-\mu t}\frac{(\mu t)^i}{i!}R^i$$

coincides with $P(t)$. Conclude from this formula that all entries of $P(t)$ are nonnegative. (Hint: Verify that $S(t)$ satisfies the same defining differential equation and the same initial condition as $P(t)$.)

4. Consider a continuous-time Markov chain with infinitesimal generator Λ and equilibrium distribution π. If the chain is at equilibrium at time 0, then show that it experiences $t\sum_i \pi_i\lambda_i$ transitions on average during the time interval $[0, t]$, where $\lambda_i = \sum_{j\neq i} \lambda_{ij}$ and λ_{ij} denotes a typical off-diagonal entry of Λ.

5. Let $P(t) = [p_{ij}(t)]$ be the finite-time transition matrix of a finite-state irreducible Markov chain. Show that $p_{ij}(t) > 0$ for all i, j, and $t > 0$. Thus, no state displays periodic behavior. (Hint: Use Problem 3.)

6. Let X_t be a finite-state reversible Markov chain with equilibrium distribution π and infinitesimal generator Λ. Suppose $\{w^i\}_i$ is an orthonormal basis of column eigenvectors of Λ in ℓ^2_π and $\{v^i\}_i$ is the corresponding orthonormal basis of row eigenvectors in $\ell^2_{1/\pi}$. Arrange

the eigenvalues $\gamma_1 = 0, \gamma_2, \ldots$ of Λ so that their real parts decline. Demonstrate that

$$e^{t\Lambda} = \sum_i e^{\gamma_i t} w^i v^i = \sum_i e^{\gamma_i t} w^i (w^i)^t \operatorname{diag}(\pi)$$

and that

$$\begin{aligned} \|\mu e^{t\Lambda} - \pi\|_{1/\pi}^2 &= \sum_{j \geq 2} e^{2\gamma_j t} \langle \mu - \pi, v^j \rangle_{1/\pi}^2 \\ &= \sum_{j \geq 2} e^{2\gamma_j t} [(\mu - \pi) w^j]^2 \\ &\leq e^{2\gamma_2 t} \|\mu - \pi\|_{1/\pi}^2 \end{aligned}$$

for any initial distribution μ. (Hints: Verify that the matrix exponential satisfies its defining differential equation, and see Section 7.5.)

7. A village with $n+1$ people suffers an epidemic. Let X_t be the number of sick people at time t, and suppose that $X_0 = 1$. If we model X_t as a continuous-time Markov chain, then a plausible model is to take the infinitesimal transition probability $\lambda_{i,i+1} = \lambda i(n+1-i)$ to be proportional to the number of encounters between sick and well people. All other $\lambda_{ij} = 0$. Now let T be the time at which the last member of the village succumbs to the disease. Since the waiting time to move from state i to state $i+1$ is exponential with intensity $\lambda_{i,i+1}$, show that $\mathrm{E}(T) \approx 2(\ln n + \gamma)/[\lambda(n+1)]$, where $\gamma \approx .5772$ is Euler's constant. It is interesting that $\mathrm{E}(T)$ decreases with n for large n.

8. Show that $e^{A+B} = e^A e^B = e^B e^A$ when $AB = BA$. (Hint: Prove that all three functions $e^{t(A+B)}$, $e^{tA}e^{tB}$, and $e^{tB}e^{tA}$ satisfy the ordinary differential equation $P'(t) = (A+B)P(t)$ with initial condition $P(0) = I$.)

9. Consider a square matrix M. Demonstrate that (a) e^{-M} is the inverse of e^M, (b) e^M is positive definite when M is symmetric, and (c) e^M is orthogonal when M is skew symmetric in the sense that $M^t = -M$. (Hint: Apply Problem 8.)

10. Let A and B be the 2×2 real matrices

$$A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}, \qquad B = \begin{pmatrix} \lambda & 0 \\ 1 & \lambda \end{pmatrix}.$$

Show that

$$e^A = e^a \begin{pmatrix} \cos b & -\sin b \\ \sin b & \cos b \end{pmatrix}, \qquad e^B = e^\lambda \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

(Hints: Note that 2×2 matrices of the form $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ are isomorphic to the complex numbers under the correspondence $\begin{pmatrix} a & -b \\ b & a \end{pmatrix} \leftrightarrow a + bi$. For the second case write $B = \lambda I + C$.)

11. Define matrices

$$A = \begin{pmatrix} a & 0 \\ 1 & a \end{pmatrix}, \qquad B = \begin{pmatrix} b & 1 \\ 0 & b \end{pmatrix}.$$

Show that $AB \neq BA$ and that

$$\begin{aligned} e^A e^B &= e^{a+b} \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \\ e^{A+B} &= e^{a+b} \left[\cosh(1) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \sinh(1) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right]. \end{aligned}$$

Hence, $e^A e^B \neq e^{A+B}$. (Hint: Use Problem 10 to calculate e^A and e^B. For e^{A+B} write $A + B = (a + b)I + R$ with R satisfying $R^2 = I$.)

12. Prove that $\det(e^A) = e^{\mathrm{tr}(A)}$, where tr is the trace function. (Hint: Since the diagonalizable matrices are dense in the set of matrices [94], by continuity you may assume that A is diagonalizable.)

13. Verify the Duhamel-Dyson identity

$$e^{t(A+B)} = e^{tA} + \int_0^t e^{(t-s)(A+B)} B e^{sA} ds$$

for matrix exponentials. (Hint: Both sides satisfy the same differential equation.)

14. Consider a random walk on the set $\{0, 1, \ldots, n\}$ with transition intensities

$$\lambda_{ij} = \begin{cases} \alpha_i & j = i + 1 \\ \beta_i & j = i - 1 \\ 0 & \text{otherwise} \end{cases}$$

for $1 \leq i \leq n - 1$. Let h_k be the probability that the process hits the absorbing state n before the absorbing state 0 starting from state k. Demonstrate that

$$h_k = \frac{\sum_{i=1}^{k-1} \prod_{j=1}^{i} \frac{\beta_j}{\alpha_j}}{\sum_{i=1}^{n-1} \prod_{j=1}^{i} \frac{\beta_j}{\alpha_j}}.$$

Simplify this expression when all $\alpha_i = \alpha$ and all $\beta_i = \beta$. (Hints: Write a difference equation for h_k by conditioning on the outcome of the first

jump. Rewrite the equation in terms of the differences $d_k = h_{k+1} - h_k$ and solve in terms of $d_0 = h_1$. This gives $h_k = d_{k-1} + \cdots + d_0$ up to the unknown h_1. The value of h_1 is determined by the initial condition $h_n = 1$.)

15. In the random walk of Problem 14, suppose that escape is possible from state 0 to state 1 with transition intensity $\alpha_0 > 0$. No other transitions out of state 0 are permitted. Let t_l be the expected time until absorption by state n starting in state l. Show that

$$t_l = \sum_{k=l}^{n-1} \sum_{j=0}^{k} \frac{1}{\alpha_k} \prod_{i=j+1}^{k} \frac{\beta_i}{\alpha_{i-1}}.$$

(Hints: Write a difference equation for t_l by conditioning on the outcome of the first jump. Rewrite the equation in terms of the differences $d_l = t_l - t_{l+1}$ and solve. Finally, invoke the telescoping series $h_l = d_l + \cdots + d_{n-1}$.)

16. In our discussion of mean hitting times in Section 7.6, we derived the formula $t = (I - Q)^{-1}\mathbf{1}$ for the vector of mean times spent in the transient states $\{1, \ldots, m\}$ en route to the absorbing states $\{m+1, \ldots, n\}$. If we pass to continuous time and replace transition probabilities p_{ij} by transition intensities λ_{ij}, then show that the mean time t_i spent in the transient states beginning at state i satisfies the equation

$$t_i = \frac{1}{\lambda_i} + \sum_{\{j: j \neq i, 1 \leq j \leq m\}} \frac{\lambda_{ij}}{\lambda_i} t_j.$$

Prove that the solution to this system can be expressed as

$$t = (I - Q)^{-1}\omega = -\Upsilon^{-1}\mathbf{1},$$

where ω is the $m \times 1$ column vector with ith entry λ_i^{-1} and Υ is the upper left $m \times m$ block of the infinitesimal generator $\Lambda = (\lambda_{ij})$.

17. In Kimura's model, suppose that two new species bifurcate at time 0 from an ancestral species and evolve independently thereafter. Show that the probability that the two species possess the same base at a given site at time t is

$$\frac{1}{4} + \frac{1}{4}e^{-8\beta t} + \frac{1}{2}e^{-4(\alpha+\beta)t}.$$

(Hint: By symmetry this formula holds regardless of what base was present at the site in the ancestral species.)

18. Recall from Example 7.3.3 that Ehrenfest's model of diffusion involves a box with n gas molecules. The box is divided in half by a rigid partition with a very small hole. Molecules drift aimlessly around and occasionally pass through the hole. Here we consider the continuous version of the process. During a short time interval h, a given molecule changes sides with probability $\lambda h + o(h)$. Show that a single molecule at time $t > 0$ is on the same side of the box as it started at time 0 with probability $\frac{1}{2}(1 + e^{-2\lambda t})$. Now consider the continuous-time Markov chain for the number of molecules in the left half of the box. Given that the n molecules behave independently, prove that finite-time transition probability $p_{ij}(t)$ amounts to

$$\begin{aligned} p_{ij}(t) &= \left(\frac{1}{2}\right)^n \sum_{k=\max\{0,i+j-n\}}^{\min\{i,j\}} \binom{i}{k}\binom{n-i}{j-k}\left(1+e^{-2\lambda t}\right)^{n-i-j+2k} \\ &\quad \times \left(1-e^{-2\lambda t}\right)^{i+j-2k}. \end{aligned}$$

(Hint: The summation index k is the number of molecules initially in the left half that end up in the left half at time t.)

19. Continuing Problem 18, it is possible to find a strong stationary time. For a single particle imagine a Poisson process with intensity 2λ. At each event of the process, flip a fair coin. If the coin lands heads, move the particle to the other half of the box. If the coin lands tails, leave the particle where it is. Why is this more elaborate process consistent with the original process? Why does the particle reach equilibrium at the moment T_i of the first event of the new process? Why do all particles reach equilibrium at the random time $T = \max_i T_i$? Let π^t be the distribution of the chain at time t, and let π be the equilibrium distribution. Deduce the total variation bound

$$\|\pi^t - \pi\|_{\text{TV}} \leq 1 - \left(1 - e^{-2\lambda t}\right)^n \approx 1 - e^{-ne^{-2\lambda t}}.$$

Why does this imply that equilibrium is reached shortly after the time $\frac{\ln n}{2\lambda}$?

20. A chemical solution initially contains $n/2$ molecules of each of the four types A, B, C, and D. Here n is a positive even integer. Each pair of A and B molecules collides at rate α to produce one C molecule and one D molecule. Likewise, each pair of C and D molecules collides at rate β to produce one A molecule and one B molecule. In this problem, we model the dynamics of these reactions as a continuous-time Markov chain X_t and seek the equilibrium distribution. The random variable X_t tracks the number of A molecules at time t [15].

(a) Argue that the infinitesimal transition rates of the chain amount to

$$\begin{aligned}\lambda_{i,i-1} &= i^2\alpha \\ \lambda_{i,i+1} &= (n-i)^2\beta.\end{aligned}$$

What about the other rates?

(b) Show that the chain is irreducible and reversible.

(c) Use Kolmogorov's formula and calculate the equilibrium distribution

$$\pi_k = \pi_0 \left(\frac{\beta}{\alpha}\right)^k \binom{n}{k}^2$$

for k between 0 and n.

(d) For the special case $\alpha = \beta$, demonstrate that

$$\pi_k = \frac{\binom{n}{k}^2}{\binom{2n}{n}}.$$

To do so first prove the identity

$$\sum_{k=0}^{n} \binom{n}{k}^2 = \binom{2n}{n}.$$

(e) To handle the case $\alpha \neq \beta$, we revert to the normal approximation to the binomial distribution. Argue that

$$\begin{aligned}\binom{n}{k} p^k q^{n-k} &= q^n \binom{n}{k} \left(\frac{p}{q}\right)^k \\ &\approx \frac{1}{\sqrt{2\pi npq}} e^{-\frac{(k-np)^2}{2npq}}\end{aligned}$$

for $p + q = 1$. Show that this implies

$$\binom{n}{k}^2 \left(\frac{p^2}{q^2}\right)^k \approx \frac{1}{2\pi npq^{2n+1}} e^{-\frac{(k-np)^2}{npq}}.$$

Now choose p so that $p^2/q^2 = \beta/\alpha$ and prove that the equilibrium distribution is approximately normally distributed with mean and variance

$$\begin{aligned}\mathrm{E}(X_\infty) &= \frac{n\sqrt{\frac{\beta}{\alpha}}}{1+\sqrt{\frac{\beta}{\alpha}}} \\ \mathrm{Var}(X_\infty) &= \frac{n\sqrt{\frac{\beta}{\alpha}}}{2\left(1+\sqrt{\frac{\beta}{\alpha}}\right)^2}.\end{aligned}$$

21. Let n indistinguishable particles independently execute the same continuous-time Markov chain with infinitesimal transition probabilities λ_{ij}. Define a new Markov chain called the composition chain for the particles by recording how many of the n total particles are in each of the s possible states. A state of the new chain is a sequence of nonnegative integers $(k_1, \ldots, k_s)$ such that $\sum_{i=1}^{s} k_i = n$. For instance, with $n = 3$ particles and $s = 2$ states, the composition chain has the four states $(3, 0)$, $(2, 1)$, $(1, 2)$, and $(0, 3)$. Find the infinitesimal transition probabilities of the composition chain. If the original chain is ergodic with equilibrium distribution $\pi = (\pi_1, \ldots, \pi_s)$, find the equilibrium distribution of the composition chain. Finally, show that the composition chain is reversible if the original chain is reversible.

22. Apply Problem 21 to the hemoglobin model in Example 8.5.1 with the understanding that the attachment sites operate independently with the same rates. What are the particles? How many states can each particle occupy? Identify the infinitesimal transition probabilities and the equilibrium distribution based on the results of Problem 21.

23. The equilibrium distribution of the numbers (M, N) of healthy and sick people in Example 8.5.3 can be found by constructing a marked Poisson process. The time X at which a random person enters state H is determined by a homogeneous Poisson process. Let Y be the time he spends in state H and Z be the time he spends in state S. If we mark each X by the pair (Y, Z), then we get a marked Poisson process on R^3. Here we suppose for the moment that all healthy people get sick before they eventually die of the given disease. Show that the random variables M and N at a given time, say $t = 0$, count the number of points in disjoint regions of R^3. Hence, these random variables are independent and Poisson distributed. Finally, prove that you can correct for our incorrect assumption by randomly thinning some of the points corresponding to N. What thinning probability should you use to get the correct mean $\mathrm{E}(N) = \frac{\alpha\beta}{\nu(\alpha+\mu)}$?

24. Cars arrive at an auto repair shop according to a Poisson process with intensity λ. There are m mechanics on duty, and each takes an independent exponential length of time with intensity μ to repair a car. If all mechanics are busy when a car arrives, it is turned away. Let X_t denote the number of mechanics busy at time t. Show that the equilibrium distribution π of X_t has components

$$\pi_k = \frac{\left(\frac{\lambda}{\mu}\right)^k \frac{1}{k!}}{\sum_{j=0}^{m} \left(\frac{\lambda}{\mu}\right)^j \frac{1}{j!}}.$$

25. Consider a pure birth process X_t with birth intensity λ_j when $X_t = j$. Let T_j denote the waiting time for a passage from state j to state $j+1$. The random variable $T = \sum_j T_j$ is the time required for the population to reach infinite size. If the series $\sum_j \lambda_j^{-1}$ converges, then demonstrate that $\Pr(T < \infty) = 1$. If the series $\sum_j \lambda_j^{-1}$ diverges, then demonstrate that $\Pr(T = \infty) = 1$. (Hints: Show that $\mathrm{E}(T) = \sum_j \lambda_j^{-1}$ and that the Laplace transform $\mathrm{E}(e^{-T}) = \prod_j (1+\lambda_j^{-1})^{-1}$. The infinite product converges to 0 if and only if the series diverges.)

26. On the lattice $\{0, 1, 2, \ldots, n\}$, particles are fed into site 0 according to a Poisson process with intensity λ. Once on the lattice a particle hops one step to the right with intensity β and evaporates with intensity μ. Show that a particle eventually reaches site n with probability $\beta^n/(\beta+\mu)^n$. Further demonstrate that, conditional on this event, the corresponding waiting time follows a gamma distribution $F_{n,\beta+\mu}(t)$ with shape parameter n and intensity $\beta+\mu$. Finally, let N_t denote the number of particles that reach site n by time $t \geq 0$. Prove that N_t is Poisson distributed with mean

$$\mathrm{E}(N_t) \;=\; \lambda\Big(\frac{\beta}{\beta+\mu}\Big)^n \int_0^t F_{n,\beta+\mu}(s)\, ds.$$

See the reference [182] for a list of biological applications and further theory. (Hint: If a particle arrives at site 0 at time X and ultimately reaches site n, then mark it by the corresponding waiting time Y. The pairs (X, Y) constitute a marked Poisson process.)

27. Prove that $G(s,t)$ defined by equation (8.18) satisfies the partial differential equation (8.16) with initial condition $G(s,0) = s$ and $\nu(t) = 0$.

28. In the homogeneous version of Kendall's process, show that

$$\begin{aligned}\mathrm{Var}(X_t) \;&=\; \frac{\nu}{(\alpha-\mu)^2}\Big[\alpha e^{(\alpha-\mu)t} - \mu\Big]\Big[e^{(\alpha-\mu)t} - 1\Big] \\ &\quad + \frac{i(\alpha+\mu)e^{(\alpha-\mu)t}}{\alpha-\mu}\Big[e^{(\alpha-\mu)t} - 1\Big]\end{aligned}$$

when $X_0 = i$.

29. Continuing Problem 28, demonstrate that

$$\mathrm{Cov}(X_{t_2}, X_{t_1}) \;=\; e^{(\alpha-\mu)(t_2-t_1)}\,\mathrm{Var}(X_{t_1})$$

for $0 \leq t_1 \leq t_2$. (Hints: First show that

$$\mathrm{Cov}(X_{t_2}, X_{t_1}) \;=\; \mathrm{Cov}[\mathrm{E}(X_{t_2} \mid X_{t_1}), X_{t_1}].$$

Then apply Problem 28.)

30. In the homogeneous version of Kendall's process, show that the generating function $G(s,t)$ of X_t satisfies

$$\lim_{t\to\infty} G(s,t) = \frac{\left(1-\frac{\alpha}{\mu}\right)^{\nu/\alpha}}{\left(1-\frac{\alpha s}{\mu}\right)^{\nu/\alpha}} \tag{8.22}$$

when $\alpha < \mu$.

31. Continuing Problem 30, prove that the equilibrium distribution π has jth component

$$\pi_j = \left(1-\frac{\alpha}{\mu}\right)^{\nu/\alpha}\left(-\frac{\alpha}{\mu}\right)^{j}\binom{-\frac{\nu}{\alpha}}{j}.$$

Do this by expanding the generating function on the right-hand side of equation (8.22) and also by applying Kolmogorov's method to Kendall's process. Note that the process is reversible.

32. Consider a time-homogeneous Kendall process with no immigration. Show that the generating function $G(s,t)$ of X_t satisfies the limit

$$\lim_{t\to\infty} \frac{G(s,t)-G(0,t)}{1-G(0,t)} = \frac{s(\mu-\alpha)}{\mu-\alpha s}$$

when $\alpha < \mu$. Prove that this limit entails the geometric probability

$$\lim_{t\to\infty} \Pr(X_t = k \mid X_t > 0) = \left(\frac{\alpha}{\mu}\right)^{k-1}\left(1-\frac{\alpha}{\mu}\right)$$

for each $k > 0$.

33. Consider the time averages

$$Y_t = \frac{1}{t}\int_0^t X_s\,ds$$

of a nonnegative stochastic process X_t with finite means and variances. Prove that

$$\begin{aligned}\mathrm{E}(Y_t) &= \frac{1}{t}\int_0^t \mathrm{E}(X_s)\,ds \\ \mathrm{Var}(Y_t) &= \frac{2}{t^2}\int_0^t\int_0^r \mathrm{Cov}(X_s,X_r)\,ds\,dr.\end{aligned}$$

In particular for a time-homogeneous Kendall process with $X_0 = 0$, show that

$$\mathrm{E}(Y_t) = \frac{\nu}{t(\alpha-\mu)^2}\left[e^{(\alpha-\mu)t}-1\right]-\frac{\nu}{\alpha-\mu}.$$

Under the same circumstances, calculate $\mathrm{Var}(Y_t)$ using Problems 28 and 29. (Hint: Apply Fubini's theorem to the first and second moments of Y_t.)

9
Branching Processes

9.1 Introduction

A branching process models the reproduction of particles such as human beings, cells, or neutrons. In the simplest branching processes, time is measured discretely in generations, and particles are of only one type. Each particle is viewed as living one generation; during this period it produces offspring contributing to the next generation. The key assumption that drives the theory is that particles reproduce independently according to the same probabilistic law. Interactions between particles are forbidden. Within this context one can ask and at least partially answer interesting questions concerning the random number X_n of particles at generation n. For instance, what are the mean $\mathrm{E}(X_n)$ and the variance $\mathrm{Var}(X_n)$? What is the extinction probability $\Pr(X_n = 0)$ on or before generation n, and what is the ultimate extinction probability $\lim_{n\to\infty} \Pr(X_n = 0)$?

Probabilists have studied many interesting elaborations of the simple branching process paradigm. For example, in some applications it is natural to include immigration of particles from outside the system and to investigate the stochastic balance between immigration and extinction. The fact that branching processes are Markov chains suggests the natural generalization to continuous time. Here each particle lives an exponentially distributed length of time. Reproduction comes as the particle dies. We have already met one such process in the guise of the time-homogeneous Kendall process. A final generalization is to processes with multiple particle types. In continuous time, each type has its own mean lifetime and own

K. Lange, *Applied Probability,*
DOI 10.1007/978-1-4419-7165-4_9, © Springer Science+Business Media, LLC 2010

reproductive pattern. Particles of one type can produce both offspring of their own type and offspring of other types.

The theory of branching processes has a certain spin that sets it apart from the general theory of Markov chains. Our focus in this chapter is on elementary results and applications to biological models. We stress computational topics such as the finite Fourier transform and the matrix exponential function rather than asymptotic results. Readers interested in pursuing the theory of branching processes in more detail should consult the references [13, 14, 51, 59, 80, 84, 103, 106, 170]. Statistical inference questions arising in branching processes are considered in reference [81].

9.2 Examples of Branching Processes

We commence our discussion of branching processes in discrete time by assuming that all particles are of the same type and that no immigration occurs. The reproductive behavior of the branching process is encapsulated in the progeny generating function $Q(s) = \sum_{k=0}^{\infty} q_k s^k$ for the number of progeny (equivalently, offspring or daughter particles) born to a single particle. If the initial number of particles $X_0 = 1$, then $Q(s)$ is the generating function of X_1. For the sake of brevity in this chapter, we refer to probability generating functions simply as generating functions. Before launching into a discussion of theory, it is useful to look at a few concrete models of branching processes.

Example 9.2.1 *Cell Division*

A cell eventually either dies with probability q_0 or divides with probability q_2. In a cell culture, cells can be made to reproduce synchronously at discrete generation times. Starting from a certain number of progenitor cells, the number of cells at successive generations forms a branching process. The progeny generating function of this process is $Q(s) = q_0 + q_2 s^2$. ■

Example 9.2.2 *Neutron Chain Reaction*

In a fission reactor, a free neutron starts a chain reaction by striking and splitting a nucleus. Typically, this generates a fixed number m of secondary neutrons. These secondary neutrons are either harmlessly absorbed or strike further nuclei and release tertiary neutrons, and so forth. The progeny generating function $Q(s) = q_0 + q_m s^m$ of the resulting branching process has mean $\mu = mq_m$. The chain reaction environment is said to be subcritical, critical, or supercritical depending on whether $\mu < 1$, $\mu = 1$, or $\mu > 1$. A nuclear reactor is carefully modulated by drawing off excess neutrons to maintain the critical state and avoid an explosion. ■

Example 9.2.3 *Survival of Family Names*

In most cultures, family names (surnames) are passed through the male line. The male descendants of a given man constitute a branching process to a good approximation. Bienaymé, Galton, and Watson introduced branching processes to study the phenomenon of extinction of family names. We will see later that extinction is certain for subcritical and critical processes. A supercritical process either goes extinct or eventually grows at a geometric rate. These remarks partially explain why some long-established countries have relatively few family names. ■

Example 9.2.4 *Epidemics*

The early stages of an epidemic are well modeled by a branching process. If the number of infected people is small, then they act approximately independently of each other. The coefficient q_k in the progeny generating function $Q(s) = \sum_{k=0}^{\infty} q_k s^k$ is the probability that an infected individual infects k further people before he or she dies or recovers from the infection. The extinction question is of paramount importance in assessing the efficacy of vaccines in preventing epidemics. ■

Example 9.2.5 *Survival of Mutant Genes*

If a dominant deleterious (harmful) mutation occurs at an autosomal genetic locus, then the person affected by the mutation starts a branching process of mutant people. (An autosomal locus is a gene location on a nonsex chromosome.) For instance, a clan of related people afflicted with the neurological disorder Huntington's disease can be viewed from this perspective [165]. Instead of sons, we follow the descendants, male and female, carrying the mutation. Because a deleterious gene is rare, we can safely assume that carriers mate only with normal people. On average, half of the children born to a carrier will be normal, and half will be carriers. The fate of each child is determined independently. Usually the reproductive fitness of carriers is sufficiently reduced so that the branching process is subcritical. ■

9.3 Elementary Theory

For the sake of simplicity, we now take $X_0 = 1$ unless noted to the contrary. To better understand the nature of a branching process, we divide the X_n descendants at generation n into X_1 clans. The kth clan consists of the descendants of the kth progeny of the founder of the process. Translating this verbal description into symbols gives the representation

$$X_n = \sum_{k=1}^{X_1} X_{nk}, \tag{9.1}$$

where X_{nk} is the number of particles at generation n in the kth clan. The assumptions defining a branching process imply that the X_{nk} are independent, probabilistic replicas of X_{n-1}. Our calculations in Example 2.4.4 consequently indicate that $\mathrm{E}(X_n) = \mathrm{E}(X_1)\,\mathrm{E}(X_{n-1})$. If we let μ be the mean number $E(X_1) = Q'(1)$ of progeny per particle, then this recurrence relation entails $\mathrm{E}(X_n) = \mu^n$.

Calculation of the variance $\mathrm{Var}(X_n)$ also yields to the analysis of Example 2.4.4. If σ^2 is the variance of X_1, then the representation (9.1) implies

$$\begin{aligned}\mathrm{Var}(X_n) &= \mu\,\mathrm{Var}(X_{n-1}) + \sigma^2\,\mathrm{E}(X_{n-1})^2 \\ &= \mu\,\mathrm{Var}(X_{n-1}) + \sigma^2\mu^{2(n-1)}. \end{aligned} \tag{9.2}$$

We claim that this recurrence relation has solution

$$\mathrm{Var}(X_n) = \begin{cases} n\sigma^2, & \mu = 1 \\ \frac{\mu^{n-1}(1-\mu^n)\sigma^2}{1-\mu}, & \mu \neq 1 \end{cases} \tag{9.3}$$

subject to the initial condition $\mathrm{Var}(X_0) = 0$. The stated formula obviously satisfies the recurrence (9.2) for $\mu = 1$. When $\mu \neq 1$ and $n = 0$, the formula also holds. If we assume by induction that it is true for $n-1$, then the calculation

$$\begin{aligned}\mathrm{Var}(X_n) &= \mu\frac{\mu^{n-2}(1-\mu^{n-1})\sigma^2}{1-\mu} + \sigma^2\mu^{2(n-1)} \\ &= \mu\sigma^2\frac{\mu^{n-2}(1-\mu^{n-1}) + \mu^{2n-3}(1-\mu)}{1-\mu} \\ &= \frac{\mu^{n-1}(1-\mu^n)\sigma^2}{1-\mu}\end{aligned}$$

combining equations (9.2) and (9.3) completes the proof for $\mu \neq 1$.

Finally, the representation (9.1) implies that X_n has generating function $Q_n(s) = Q(Q_{n-1}(s))$, which is clearly the n-fold functional composition of $Q(s)$ with itself. The next example furnishes one of the rare instances in which $Q_n(s)$ can be explicitly found.

Example 9.3.1 *Geometric Progeny Distribution*

In a sequence of Bernoulli trials with success probability p and failure probability $q = 1-p$, the number of failures Y until the first success has a geometric distribution with generating function

$$Q(s) = \sum_{k=0}^{\infty} pq^k s^k = \frac{p}{1-qs}.$$

The mean and variance of Y are

$$\mu = \frac{pq}{(1-qs)^2}\Big|_{s=1} = \frac{q}{p}$$

and

$$\sigma^2 = \frac{2pq^2}{(1-qs)^3}\Big|_{s=1} + \mu - \mu^2 = \frac{q}{p^2}.$$

We can verify inductively that

$$Q_n(s) = \begin{cases} \frac{n-(n-1)s}{n+1-ns}, & p = q \\ p\frac{q^n-p^n-(q^{n-1}-p^{n-1})qs}{q^{n+1}-p^{n+1}-(q^n-p^n)qs}, & p \neq q\,. \end{cases} \tag{9.4}$$

When $n = 1$, the second of these formulas holds because

$$\begin{aligned} \frac{p(q-p)}{q^2-p^2-(q-p)qs} &= \frac{p(q-p)}{(q+p)(q-p)-(q-p)qs} \\ &= \frac{p}{1-qs}. \end{aligned}$$

Assuming that the second formula in (9.4) is true for $n-1$, we reason that

$$\begin{aligned} Q_n(s) &= Q(Q_{n-1}(s)) \\ &= \frac{p}{1-qp\dfrac{q^{n-1}-p^{n-1}-(q^{n-2}-p^{n-2})qs}{q^n-p^n-(q^{n-1}-p^{n-1})qs}} \\ &= p\frac{q^n-p^n-(q^{n-1}-p^{n-1})qs}{q^n(1-p)-p^n(1-q)-q^{n-1}(1-p)qs+p^{n-1}(1-q)qs} \\ &= p\frac{q^n-p^n-(q^{n-1}-p^{n-1})qs}{q^{n+1}-p^{n+1}-(q^n-p^n)qs}. \end{aligned}$$

The inductive proof of the first formula in (9.4) is left to the reader. ■

9.4 Extinction

Starting with a single particle at generation 0, we now ask for the probability s_∞ that a branching process eventually goes extinct. To characterize s_∞, we condition on the number of progeny $X_1 = k$ born to the initial particle. If extinction is to occur, then each of the clans emanating from the k progeny must go extinct. By independence, this event occurs with probability s_∞^k. Thus, s_∞ satisfies the functional equation $s_\infty = \sum_{k=0}^{\infty} q_k s_\infty^k = Q(s_\infty)$, where $Q(s)$ is the progeny generating function.

One can find the extinction probability by functional iteration starting at $s = 0$. Let s_n be the probability that extinction occurs in the branching process at or before generation n. Then $s_0 = 0$, $s_1 = q_0 = Q(s_0)$, and, in general, $s_{n+1} = Q(s_n) = Q_{n+1}(0)$. This recurrence relation can be deduced by conditioning once again on the number of progeny in the first generation.

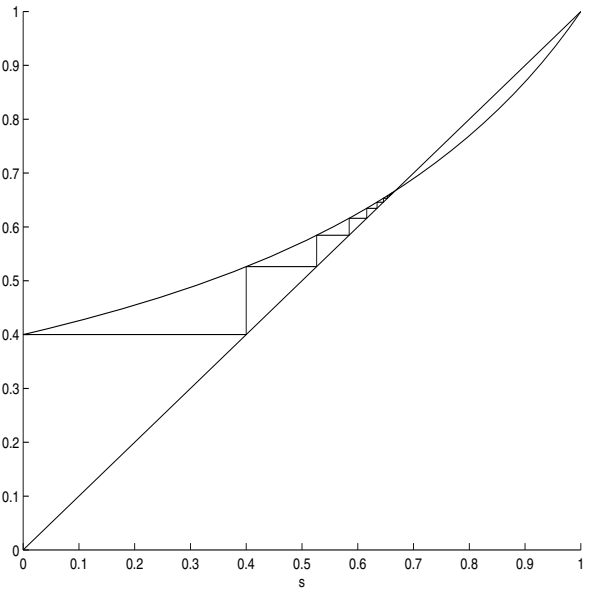


FIGURE 9.1. Extinction Iterates in a Supercritical Branching Process

If extinction is to occur at or before generation $n+1$, then it must occur in n additional generations or sooner for each clan emanating from a daughter particle of the founding particle.

On probabilistic grounds it is obvious that the sequence s_n increases monotonically to the extinction probability s_∞. To understand what is happening analytically, we need to know the number of roots of $s = Q(s)$ and which of these roots is s_∞. Because $Q''(s) = \sum_{k=2}^{\infty} k(k-1)q_k s^{k-2} \geq 0$, the curve $Q(s)$ is convex. It starts at $Q(0) = q_0 > 0$ above the diagonal line $t = s$ in Figure 9.1. (Note that if $q_0 = 0$, then the process can never go extinct.) On the interval $[0, 1]$, the curve $Q(s)$ and the line $t = s$ intersect in either one or two points. The point $s = 1$ is certainly an intersection point because $Q(1) = \sum_{k=0}^{\infty} q_k = 1$. When the progeny mean $\mu = Q'(1) > 1$, the curve $Q(s)$ intersects $t = s$ at $s = 1$ from below, and there is a second intersection point to the left of $s = 1$. When $\mu < 1$, the second intersection point occurs to the right of $s = 1$.

The cobweb diagram in Figure 9.1 following the progress of the iterates s_n makes it clear that they not only monotonically increase, but they are also bounded above by the smaller of the two roots. Thus, the limit s_∞ of the s_n exists and is bounded above by the smaller root. Taking limits in the recurrence $s_{n+1} = Q(s_n)$ demonstrates that s_∞ coincides with the smaller root of $s = Q(s)$. In other words, extinction is certain for subcritical and critical processes and uncertain for supercritical processes.

TABLE 9.1. Functional Iteration for an Extinction Probability

Iteration n	**Iterate** s_n	**Iteration** n	**Iterate** s_n
0	.000	10	.847
1	.498	20	.873
2	.647	30	.878
3	.719	40	.879
4	.761	50	.880
5	.788		

Example 9.4.1 *Lotka's American Surname Data*

As a numerical example, consider the data of Lotka [138, 159] on the extinction of surnames among white males in the United States. Using 1920 census data, he computed the progeny generating function

$$\begin{aligned} Q(s) &= .4982 + .2103s + .1270s^2 + .0730s^3 + .0418s^4 + .0241s^5 \\ &\quad + .0132s^6 + .0069s^7 + .0035s^8 + .0015s^9 + .0005s^{10} \end{aligned}$$

for the number of sons of a random father. Because the average number of sons per father $\mu = Q'(1) > 1$, we anticipate an extinction probability $s_\infty < 1$. Table 9.1 lists some representative functional iterates. Convergence to the extinction probability .880 is slow but sure. ■

Example 9.4.2 *Extinction Probability for the Geometric Distribution*

Consider once again the geometric distribution of Example 9.3.1. The two roots of the equation $s = \frac{p}{1-qs}$ are 1 and $\frac{p}{q} = \frac{1}{\mu}$. When $\mu > 1$, the extinction probability is $s_\infty = \frac{p}{q}$. When $\mu < 1$, the root $\frac{p}{q}$ lies outside the interval $[0, 1]$, and the extinction probability is $s_\infty = 1$. Finally, when $\mu = 1$, the two roots coincide. Thus, extinction is certain if $\mu \leq 1$ and uncertain otherwise.

The same conclusions can be reached by considering the behavior of $s_n = Q_n(0)$ as suggested by formula (9.4). If $p < q$, then

$$\lim_{n\to\infty} Q_n(0) = \lim_{n\to\infty} \frac{p}{q} \frac{1 - \left(\frac{p}{q}\right)^n}{1 - \left(\frac{p}{q}\right)^{n+1}} = \frac{p}{q}.$$

If $p > q$, then

$$\lim_{n\to\infty} Q_n(0) = \lim_{n\to\infty} \frac{1 - \left(\frac{q}{p}\right)^n}{1 - \left(\frac{q}{p}\right)^{n+1}} = 1.$$

Finally, if $p = q$, then

$$\lim_{n\to\infty} Q_n(0) = \lim_{n\to\infty} \frac{n}{n+1} = 1.$$

The rate of convergence in this critical case is very slow. ■

If T denotes the generation at which extinction occurs in a subcritical process, then $\Pr(T > n) = 1 - Q_n(0) = 1 - s_n$. The tail-probability method of Example 2.5.1 therefore implies

$$\begin{aligned} \mathrm{E}(T) &= \sum_{n=0}^{\infty} (1 - s_n) \\ \mathrm{E}(T^2) &= \sum_{n=0}^{\infty} (2n+1)(1 - s_n). \end{aligned} \tag{9.5}$$

Truncated versions of these sums permit one to approximate the first two moments of T. Problem 10 develops practical error bounds on this procedure. Finally, it is useful to contrast the critical case to the subcritical case. For instance, with a critical geometric generating function, we find that

$$\mathrm{E}(T) = \sum_{n=0}^{\infty} (1 - s_n) = \sum_{n=0}^{\infty} \left(1 - \frac{n}{n+1}\right) = \infty.$$

Now let $Y_n = 1 + \sum_{k=1}^{n} X_k$ be the total number of descendants up to generation n. The limit $Y_\infty = \lim_{n\to\infty} Y_n$ exists because the Y_n form an increasing sequence. The next proposition collects pertinent facts about this interesting random variable.

Proposition 9.4.1 *If $r_k = \Pr(Y_\infty = k)$ and $R(s) = \sum_{k=1}^{\infty} r_k s^k$, then the following hold:*

(a) The extinction probability $s_\infty = \Pr(Y_\infty < \infty) = \sum_{k=1}^{\infty} r_k$. Therefore, in a supercritical process $Y_\infty = \infty$ occurs with positive probability.

(b) The generating function $R(s)$ satisfies $R(s) = sQ(R(s))$.

(c) For $\mu < 1$, the mean and variance of Y_∞ are $\mathrm{E}(Y_\infty) = \frac{1}{1-\mu}$ and $\mathrm{Var}(Y_\infty) = \frac{\sigma^2}{(1-\mu)^3}$.

(d) If the power $Q(s)^j$ has expansion $\sum_{k=0}^{\infty} q_{jk} s^k$, then $r_j = \frac{1}{j} q_{j,j-1}$.

Proof: To prove part (a), note that there is at least one particle per generation if and only if the process does not go extinct. Hence, Y_∞ is finite if and only if the process goes extinct. For part (b), note that we have by analogy to equation (9.1) the representation

$$Y_\infty = 1 + \sum_{k=1}^{X_1} Y_{\infty k}, \tag{9.6}$$

where the $Y_{\infty k}$ are independent, probabilistic replicas of Y_∞. The generating function version of this representation is precisely $R(s) = sQ(R(s))$ as

described in Example 2.4.4. To verify part (c), observe that equation (9.6) implies

$$\begin{aligned} \mathrm{E}(Y_\infty) &= 1 + \mu\,\mathrm{E}(Y_\infty) \\ \mathrm{Var}(Y_\infty) &= \mu\,\mathrm{Var}(Y_\infty) + \sigma^2\,\mathrm{E}(Y_\infty)^2. \end{aligned}$$

Assuming for the sake of simplicity that both $\mathrm{E}(Y_\infty)$ and $\mathrm{Var}(Y_\infty)$ are finite, we can solve the first of these equations to derive the stated expression for $\mathrm{E}(Y_\infty)$. Inserting this solution into the second equation and solving gives $\mathrm{Var}(Y_\infty)$. Proof of part (d) involves the tricky Lagrange inversion formula and is consequently omitted [42]. ■

Example 9.4.3 *Total Descendants for the Geometric Distribution*

When $Q(s) = \frac{p}{1-qs}$, application of part (b) of Proposition 9.4.1 gives the equation $R(s) = ps/[1-qR(s)]$ or $qR(s)^2 - R(s) + ps = 0$. The smaller root

$$R(s) = \frac{1 - \sqrt{1-4pqs}}{2q}$$

of this quadratic is the relevant one. Indeed, based on the representation

$$\begin{aligned} \frac{1+\sqrt{1-4pq}}{2q} &= \frac{p+q+\sqrt{(p+q)^2-4pq}}{2q} \\ &= \frac{p+q+|p-q|}{2q}, \end{aligned}$$

taking the larger root produces the contradictory results $R(1) > 1$ in the subcritical case where $q < p$ and $R(1) = 1$ in the supercritical case where $q > p$. In the subcritical case, Proposition 9.4.1 or differentiation of $R(s)$ shows that $\mathrm{E}(Y_\infty) = \frac{p}{p-q}$ and $\mathrm{Var}(Y_\infty) = \frac{pq}{(p-q)^3}$. ■

9.5 Immigration

We now modify the definition of a branching process to allow for immigration at each generation. In other words, we assume that the number of particles X_n at generation n is the sum $U_n + Z_n$ of the progeny U_n of generation $n-1$ plus a random number of immigrant particles Z_n independent of U_n. To make our theory as simple as possible, we take the Z_n to be independent and identically distributed with common mean α and variance β^2. If the process without immigration is subcritical ($\mu < 1$), then particle numbers eventually reach a stochastic equilibrium between extinction and immigration. The goal of this section is to characterize the equilibrium distribution.

Our point of departure in the subcritical case is the representation

$$U_n = \sum_{k=1}^{X_{n-1}} V_{n-1,k} \tag{9.7}$$

of the progeny of generation $n-1$ partitioned by parent particle k. From this representation it follows by the usual conditioning arguments that

$$\begin{aligned} \mathrm{E}(X_n) &= \mathrm{E}(X_{n-1})\,\mathrm{E}(V_{n-1,1}) + \mathrm{E}(Z_n) \\ &= \mu\,\mathrm{E}(X_{n-1}) + \alpha \\ \mathrm{Var}(X_n) &= \mathrm{E}(X_{n-1})\,\mathrm{Var}(V_{n-1,1}) + \mathrm{Var}(X_{n-1})\,\mathrm{E}(V_{n-1,1})^2 + \mathrm{Var}(Z_n) \\ &= \sigma^2\,\mathrm{E}(X_{n-1}) + \mu^2\,\mathrm{Var}(X_{n-1}) + \beta^2. \end{aligned} \tag{9.8}$$

If we assume $\lim_{n\to\infty} \mathrm{E}(X_n) = \mathrm{E}(X_\infty)$ and $\lim_{n\to\infty} \mathrm{Var}(X_n) = \mathrm{Var}(X_\infty)$ for some random variable X_∞, then taking limits on n in the two equations in (9.8) yields

$$\begin{aligned} \mathrm{E}(X_\infty) &= \mu\,\mathrm{E}(X_\infty) + \alpha \\ \mathrm{Var}(X_\infty) &= \sigma^2\,\mathrm{E}(X_\infty) + \mu^2\,\mathrm{Var}(X_\infty) + \beta^2. \end{aligned}$$

Solving these two equations in succession produces

$$\begin{aligned} \mathrm{E}(X_\infty) &= \frac{\alpha}{1-\mu} \\ \mathrm{Var}(X_\infty) &= \frac{\alpha\sigma^2 + \beta^2(1-\mu)}{(1-\mu)^2(1+\mu)}. \end{aligned} \tag{9.9}$$

It is interesting that $\mathrm{E}(X_\infty)$ is the product of the average number of immigrants α per generation times the average clan size $\frac{1}{1-\mu}$ per particle identified in part (c) of Proposition 9.4.1.

Our next aim is to find the distribution of X_∞. Let $P_n(s)$ be the generating function of X_n and $R(s)$ be the common generating function of the Z_n. Then the decomposition $X_n = U_n + Z_n$ and equation (9.7) imply

$$P_n(s) = P_{n-1}(Q(s))R(s), \tag{9.10}$$

where $Q(s)$ is the progeny generating function. Iterating equation (9.10) yields

$$P_n(s) = P_0(Q_n(s)) \prod_{k=0}^{n-1} R(Q_k(s)), \tag{9.11}$$

where $Q_k(s)$ is again the k-fold functional composition of $Q(s)$ with itself. The finite product (9.11) tends to the infinite product

$$P_\infty(s) = \prod_{k=0}^{\infty} R(Q_k(s)) \tag{9.12}$$

representation of the generating function of X_∞. Observe here that the leading term $P_0(Q_n(s))$ on the right of equation (9.11) tends to 1 because $Q_n(0)$ tends to the extinction probability 1 and $Q_n(0) \leq Q_n(s) \leq 1$ for all $s \in [0,1]$.

The problem now is to recover the coefficients p_k of $P_\infty(s) = \sum_{k=0}^\infty p_k s^k$. This is possible using the values of $P_\infty(s)$ for s on the boundary of the unit circle. Once we reparameterize by setting $s = e^{2\pi it}$ for $t \in [0,1]$ and $i = \sqrt{-1}$, we recognize p_k as the kth Fourier series coefficient of the periodic function $P_\infty(e^{2\pi it}) = \sum_{k=0}^\infty p_k e^{2\pi ikt}$. Obviously,

$$p_k = \int_0^1 P_\infty(e^{2\pi it}) e^{-2\pi ikt} dt$$

can be approximated by the Riemann sum

$$\frac{1}{n} \sum_{j=0}^{n-1} P_\infty\left(e^{\frac{2\pi ij}{n}}\right) e^{-\frac{2\pi ijk}{n}} \tag{9.13}$$

for n large. The n sums in (9.13) for $0 \leq k \leq n-1$ collectively define the finite Fourier transform of the sequence of n numbers $P_\infty(e^{2\pi ij/n})$ for $0 \leq j \leq n-1$. To compute a finite Fourier transform, one can use an algorithm known as the fast Fourier transform [87, 117]. This material is reviewed in Appendix A.4 and Section 13.3.

In summary, we can compute the p_k by

(a) choosing n so large that all p_k with $k \geq n$ can be ignored,

(b) approximating $P_\infty(e^{2\pi ij/n})$ by the finite product

$$\prod_{k=0}^m R(Q_k(e^{2\pi ij/n}))$$

with m large,

(c) taking finite Fourier transforms of the finite product.

To check the accuracy of these approximations, we numerically compute the moments of X_∞ from the resulting p_k for $0 \leq k \leq n-1$ and compare the results to the theoretical moments [120].

Example 9.5.1 *Huntington's Disease*

Huntington's disease is caused by a deleterious dominant gene cloned in 1993 [198]. In the late 1950s, Reed and Neil estimated that carriers of the Huntington gene have a fitness $f \approx 0.81$ [165]. Here f is the ratio of the expected number of offspring of a carrier to the expected number of offspring of a normal person. In a stationary population, each person has

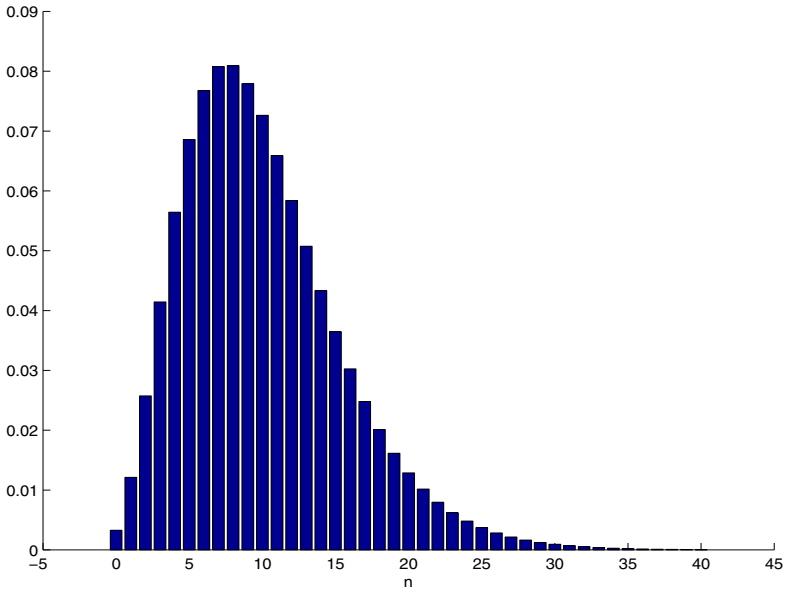


FIGURE 9.2. Equilibrium Distribution $\Pr(X_\infty = n)$ for Huntington's Disease

on average one daughter and one son. Each carrier produces on average $0.81 \times 2 \times \frac{1}{2} = 0.81$ carrier children. If we let the progeny generating function $Q(s)$ be Poisson with mean $\mu = 0.81$, then the ultimate number of people Y_∞ affected starting from a single mutation has mean and variance

$$\begin{aligned} \mathrm{E}(Y_\infty) &= \frac{1}{1-\mu} = 5.24 \\ \mathrm{Var}(Y_\infty) &= \frac{\sigma^2}{(1-\mu)^3} = 118.09 \end{aligned}$$

based on part (c) of Proposition 9.4.1 with $\sigma^2 = \mu$. The generation T at which extinction of the mutation occurs has mean $\mathrm{E}(T) = 3.01$ and variance $\mathrm{Var}(T) = \mathrm{E}(T^2) - \mathrm{E}(T)^2 = 11.9$ [127].

Let us also assume that the surrounding normal population exhibits a mutation rate of $\nu = 2 \times 10^{-6}$ and contains $r = 237{,}500$ females. Immigration into the branching process occurs whenever a parental gene mutates just prior to transmission to a child. By the "law of rare events" explained in Example 14.3.1, the number of new mutants at each generation is approximately Poisson with mean $\alpha = 2r2\nu = 4r\nu$. Here $2r$ is the total number of births per generation, and 2ν is the probability that either of the two parental genes contributed to a child mutate. According to equation (9.9), the mean and variance of the equilibrium distribution are

$$\mathrm{E}(X_\infty) = \frac{4r\nu}{1-\mu}$$

$$\begin{aligned} &= 10 \\ \mathrm{Var}(X_\infty) &= \frac{4r\nu\sigma^2 + 4r\nu(1-\mu)}{(1-\mu)^2(1+\mu)} \\ &= 29.08. \end{aligned}$$

Figure 9.2 depicts the equilibrium distribution as computed by the Fourier series method. ■

9.6 Multitype Branching Processes

In a multitype branching process, one follows a finite number of independently acting particles that reproduce and die. Each particle is classified in one of n possible categories. In a continuous-time process, a type i particle lives an exponentially distributed length of time with death intensity λ_i. At the end of its life, a type i particle reproduces both particles of its own type and particles of other types. Suppose that on average it produces f_{ij} particles of type j.

We would like to calculate the average number of particles $m_{ij}(t)$ of type j at time $t \geq 0$ starting with a single particle of type i at time 0. Since particles of type j at time $t+s$ either arise from particles of type j at time t that do not die during $(t, t+s)$ or from particles of type k that die during $(t, t+s)$ and reproduce particles of type j, we find that

$$m_{ij}(t+s) = m_{ij}(t)(1-\lambda_j s) + \sum_k m_{ik}(t)\lambda_k f_{kj} s + o(s).$$

Forming the corresponding difference quotients and sending s to 0 yield the system of differential equations

$$m'_{ij}(t) = \sum_k m_{ik}(t)\lambda_k(f_{kj} - 1_{\{k=j\}}),$$

which we summarize as the matrix differential equation $M'(t) = M(t)\Omega$ for the $n \times n$ matrices $M(t) = [m_{ij}(t)]$ and $\Omega = [\lambda_i(f_{ij} - 1_{\{i=j\}})]$. The solution is provided by the matrix exponential $M(t) = e^{t\Omega}$ subject to the initial condition $M(0) = I$ [14].

The asymptotic behavior of $M(t)$ is determined qualitatively by the eigenvalue ρ of Ω with largest real part. As shown in Appendix A.2, a dominant eigenvalue exists and is real for an irreducible process. In such a process, a particle of one type can ultimately produce descendants of every other type. A descendant may be a granddaughter, great granddaughter, and so forth rather than a daughter directly. Irreducibility can be checked by examining the off-diagonal elements of Ω. When a process is irreducible, it is said to be subcritical, critical, or supercritical according as $\rho < 0$, $\rho = 0$, or $\rho > 0$, respectively.

The assignment of a chain to one of these three categories can be based entirely on the reproduction matrix F. Proposition A.2.5 of the Appendix demonstrates this fact. The components of $M(t)$ tend to zero for a subcritical process and to infinity for a supercritical process. To gain insight into this claim, consider the vector function $M(t)v$, where v is the unique positive eigenvector corresponding to the dominant eigenvalue ρ. Differentiation of $M(t)v$ gives the vector differential equation

$$\frac{d}{dt}M(t)v \;=\; M(t)\Lambda v \;=\; \rho M(t)v$$

with initial condition $M(0)v = v$. The solution $M(t)v = e^{\rho t}v$ exhibits the claimed behavior of convergence to $\mathbf{0}$ or divergence.

To investigate the phenomenon of extinction, let $e_i(t)$ be the probability that the process is extinct at time t given that it begins with a single particle of type i at time 0. We can characterize the $e_i(t)$ by deriving a system of nonlinear ordinary differential equations. In almost all cases, this system must be solved numerically. In deriving the system of differential equations, suppose a type i particle produces d_1 type 1 daughter particles, d_2 type 2 daughter particles, and so on, to d_n type n daughter particles, with probability $p_{i,(d_1,\ldots,d_n)}$. To ease the notational burden, we write

$$\begin{aligned}\mathbf{d} &= (d_1,\ldots,d_n)\\ p_{i\mathbf{d}} &= p_{i,(d_1,\ldots,d_n)}\\ \mathbf{e}(t)^{\mathbf{d}} &= \prod_{i=1}^{n} e_i(t)^{d_i}.\end{aligned}$$

Given these conventions, we contend that

$$e_i(t+s) \;=\; (1-\lambda_i s)e_i(t) + \lambda_i s\sum_{\mathbf{d}} p_{i\mathbf{d}}\mathbf{e}(t)^{\mathbf{d}} + o(s). \qquad (9.14)$$

The logic behind this expression is straightforward. Either the original particle does not die during the time interval $(0, s)$, or it does and leaves behind a vector of $\mathbf{d}$ daughter particles with probability $p_{i\mathbf{d}}$. Each of the clans emanating from one of the daughter particles goes extinct independently of the remaining clans. Rearranging expression (9.14) into a difference quotient and sending s to 0 yield the differential equation

$$e_i'(t) \;=\; -\lambda_i e_i(t) + \lambda_i\sum_{\mathbf{d}} p_{i\mathbf{d}}\mathbf{e}(t)^{\mathbf{d}}.$$

The probabilities $e_i(t)$ are increasing because once extinction occurs, it is permanent. To find the limit of $e_i(t)$ as t tends to infinity, we set $e_i'(t)$ equal to 0. This action determines the ultimate extinction probabilities $\mathbf{e} = (e_1,\ldots,e_n)$ through the algebraic equations

$$e_i \;=\; \sum_{\mathbf{d}} p_{i\mathbf{d}}\mathbf{e}^{\mathbf{d}}. \qquad (9.15)$$

It is noteworthy that these equations do not depend on the average life expectancies of the different particle types but only on their reproductive patterns. For subcritical and critical irreducible processes, all extinction probabilities are 1; see Problem 17 for the subcritical case. For supercritical irreducible processes, all extinction probabilities are strictly less than 1 [13, 14, 84, 106].

9.7 Viral Reproduction in HIV

As an illustration of the above theory, consider the following branching process model of how the HIV virus infects CD4 cells of the immune system in the first stage of an HIV infection [156]. Particles in this simplified model correspond to either virus particles (virions) in plasma or two types of infected CD4 cells. Virions are type 1 particles, latently infected CD4 cells are type 2 particles, and actively infected CD4 cells are type 3 particles. In actively infected CD4 cells, HIV is furiously replicating. As a first approximation, we will assume that replicated virions are released in a burst when the cell membrane ruptures and a large number of virions spill into the plasma. HIV does not replicate in latently infected cells. Therefore, a latently infected cell either converts to an actively infected cell or quietly dies without replenishing the plasma load of virus.

Let us now consider the fate of each type of particle. Type 1 particles or virions die or are eliminated from plasma at rate σ. A virion enters and infects an uninfected CD4 cell at rate βR, where β is the infection rate per CD4 cell, and where R is the fixed number of uninfected CD4 cells in plasma. (A defect of the model is that as time progresses R should decline as more and more CD4 cells are infected. In the earliest stage of an infection, we can ignore this effect.) Let θ be the probability that a CD4 cell commences its infection in a latent rather than in an active state. In the branching process paradigm, we interpret death broadly to mean either natural virion death, virion elimination, or virion removal by infection. Thus, our verbal description translates into the quantitative assumptions $\lambda_1 = \sigma + \beta R$ and

$$f_{11} = 0, \qquad f_{12} = \frac{\theta\beta R}{\lambda_1}, \qquad f_{13} = \frac{(1-\theta)\beta R}{\lambda_1}.$$

Latently infected cells are eliminated at rate μ or convert to actively infected cells at rate α. Again broadly interpreting death in the branching process sense, we take $\lambda_2 = \mu + \alpha$ and

$$f_{21} = 0, \qquad f_{22} = 0, \qquad f_{23} = \frac{\alpha}{\lambda_2}.$$

Finally, actively infected CD4 cells are eliminated at rate μ or burst due to viral infection at rate δ. When an actively infected CD4 cell bursts, it

dumps an average π virions into the plasma. These new virions start the cycle anew. For type 3 particles we have $\lambda_3 = \mu + \delta$ and

$$f_{31} = \frac{\delta\pi}{\lambda_3}, \qquad f_{32} = 0, \qquad f_{33} = 0.$$

The Ω matrix determining the mean behavior of the model therefore reduces to

$$\Omega = \begin{pmatrix} -\sigma - \beta R & \theta\beta R & (1-\theta)\beta R \\ 0 & -\mu - \alpha & \alpha \\ \delta\pi & 0 & -\mu - \delta \end{pmatrix}.$$

It is impossible to calculate extinction probabilities in this model without specifying the model more fully. One possibility is to make the somewhat more realistic assumption that reproduction of virions by an actively infected cell occurs continuously rather than as a gigantic burst. If the cell sheds single virions with a steady intensity γ, then we need to adjust γ so that the expected number of virions produced continuously over the cell's lifetime matches the expected number of virions produced instantaneously by bursting. This leads to the condition $\gamma/\mu = \delta\pi/\lambda_3$ determining γ. We also must adjust our conception of death. The branching process model requires that reproduction occur simultaneously with death. Hence, in our revised model, an actively infected cell dies with rate $\mu+\gamma$. At its death, it leaves behind no particles with probability $\mu/(\mu+\gamma)$ or an actively infected cell and a virion with probability $\gamma/(\mu+\gamma)$.

With these amendments, the system of extinction equations (9.15) becomes

$$\begin{aligned} e_1 &= \frac{\sigma}{\lambda_1} + \frac{\theta\beta R}{\lambda_1} e_2 + \frac{(1-\theta)\beta R}{\lambda_1} e_3 \\ e_2 &= \frac{\mu}{\lambda_2} + \frac{\alpha}{\lambda_2} e_3 \\ e_3 &= \frac{\mu}{\mu+\gamma} + \frac{\gamma}{\mu+\gamma} e_1 e_3. \end{aligned}$$

This system of equations reduces to a single quadratic equation for e_1 if we (a) substitute for e_2 in the first equation using the second equation, (b) solve the third equation for e_3 in terms of e_1, and (c) substitute the result into the modified first equation. We leave the messy details to the reader.

9.8 Basic Reproduction Numbers

The reproduction number R_0 of an irreducible continuous-time branching process is defined by focusing on a single particle and calculating the average number of offspring of the same type it generates in one cycle of

the process [102, 200]. Without loss of generality, we will take the reference type to be type 1. This arbitrary choice is made merely for the sake of convenience. We then define d_i to be the expected number of particles of type 1 produced in one cycle starting with a single particle of type i. This definition is vague because the term cycle is undefined. In addition to counting a type i particle's immediate daughters of type 1, we count type 1 particles that ultimately issue from her daughters of type $j \neq 1$. These considerations can be summarized by the system of equations

$$d_i = f_{i1} + \sum_{j=2}^{n} f_{ij} d_j. \tag{9.16}$$

The system (9.16) is recursive in the sense that the implied enumeration takes into account every chain of particles starting with the founding type i particle and ending with the first type 1 particle encountered. Furthermore, every nonnegative solution vector $d = (d_1, \ldots, d_n)^t$ is nontrivial because irreducibility forces at least one reproductive value f_{i1} to be positive.

It is straightforward to generate the vector d. The last $n-1$ equations of the system (9.16) do not involve d_1 and can be solved in the form

$$\begin{pmatrix} d_2 \\ \vdots \\ d_n \end{pmatrix} = (I_{n-1} - G)^{-1} \begin{pmatrix} f_{21} \\ \vdots \\ f_{n1} \end{pmatrix}, \tag{9.17}$$

where I_{n-1} is the $n-1$ dimensional identity matrix and G is the corresponding lower-right submatrix of F. Once d_2 through d_n are calculated, it is trivial to recover d_1 via

$$d_1 = f_{11} + \sum_{j=2}^{n} f_{1j} d_j. \tag{9.18}$$

When the process is subcritical, we will show that we can solve for d in this fashion and that all components d_i are nonnegative. Given that d is properly defined with nonnegative entries, we further demonstrate that d_1 is less than 1, equal to 1, or greater than 1, depending on whether the process is, respectively, subcritical, critical, or supercritical.

As an illustration of the above theory, consider the HIV model. If we take type 1 (virions) as the reference type, then the system (9.16) becomes

$$\begin{aligned} d_1 &= \frac{\theta\beta R}{\sigma + \beta R} d_2 + \frac{(1-\theta)\beta R}{\sigma + \beta R} d_3 \\ d_2 &= \frac{\alpha}{\mu + \alpha} d_3 \\ d_3 &= \frac{\delta\pi}{\mu + \delta}. \end{aligned}$$

In this case there is a single solution vector (d_1, d_2, d_3), and its entries are clearly positive. Simple algebra shows that the expected number of virions produced in one cycle is

$$d_1 = \frac{\delta\pi\beta R[\alpha + (1-\theta)\mu]}{(\sigma + \beta R)(\mu + \alpha)(\mu + \delta)}.$$

This basic reproduction number determines the qualitative behavior of the branching process. When the reproduction number $d_1 < 1$, virus numbers keep dropping until extinction. When $d_1 > 1$, extinction is uncertain and virus numbers may grow exponentially.

The reader is urged to consult Appendix A.2 for the details omitted in the following proofs. Consider first the existence of the subvector $(d_2, \ldots, d_n)^t$. This is tied to the convergence of the series

$$(I_{n-1} - G)^{-1} f_{-1} = \sum_{k=0}^{\infty} G^k f_{-1}, \tag{9.19}$$

where f_{-1} denotes the first column of F with entry f_{11} removed. The series expansion of the inverse $(I_{n-1} - G)^{-1}$ is valid whenever the dominant eigenvalue $\rho(G) < 1$. Now $\rho(F - I) < 0$ entails $\rho(F) < 1$, which is sufficient to prove $\rho(G) < 1$. Indeed, suppose that $Gv = rv$ for some vector $v \neq \mathbf{0}$ and scalar r with $|r| > \rho(F)$. If we define the vector w to have entries $w_i = |v_i|$, then $Gw \geq |r|w$. Since all entries of F are nonnegative, it follows that

$$F \begin{pmatrix} 0 \\ w \end{pmatrix} \geq \begin{pmatrix} 0 \\ |r|w \end{pmatrix} = |r| \begin{pmatrix} 0 \\ w \end{pmatrix}.$$

Definition (A.1) of Appendix A.2 now yields the contradiction $|r| \leq \rho(F)$. Given that the geometric series (9.19) converges, it is straightforward to prove that it solves the equation $d_{-1} = f_{-1} + Gd_{-1}$. The representation (9.18) shows that d_1 is nonnegative as well.

Assuming that the system (9.16) possesses a nonnegative solution d, we now undertake the task of classifying the behavior of the branching process based on the entry d_1. If $d_1 \geq 1$, then

$$d_i \leq f_{i1} d_1 + \sum_{j=2}^{n} f_{ij} d_j = \sum_{j=1}^{n} f_{ij} d_j$$

for all i. In view of Proposition A.2.2, we find $\rho(F) \geq 1$. On the other hand, if $d_1 \leq 1$, then for all i

$$d_i \geq f_{i1} d_1 + \sum_{j=2}^{n} f_{ij} d_j = \sum_{j=1}^{n} f_{ij} d_j,$$

and Proposition A.2.4 implies $\rho(F) \leq 1$. This settles the case $d_1 = 1$ because the only possibility is $\rho(F) = 1$. To clarify what happens when $d_1 < 1$ or $d_1 > 1$, we note that either situation leads to the inequality $Fd \neq d$ since at least one entry $f_{i1} > 0$. Application of Propositions A.2.1, A.2.2, and A.2.4 now completes the proof.

9.9 Problems

1. If p and α are constants in the open interval $(0, 1)$, then show that $Q(s) = 1 - p(1 - s)^\alpha$ is a generating function with nth functional iterate

$$Q_n(s) \quad = \quad 1 - p^{1+\alpha+\cdots+\alpha^{n-1}}(1 - s)^{\alpha^n}.$$

   Remember to check that the coefficients of $Q(s)$ are nonnegative and sum to 1.

2. Let X_n be the number of particles at generation n in a supercritical branching process with progeny mean μ and variance σ^2. If $X_0 = 1$ and $Z_n = X_n/\mu^n$, then find $\lim_{n\to\infty} \mathrm{E}(Z_n)$ and $\lim_{n\to\infty} \mathrm{Var}(Z_n)$. The fact that these limits exist and are finite when $\mu > 1$ correctly suggests that Z_n tends to a limiting random variable Z_∞.

3. Consider a supercritical branching process X_n with progeny generating function $Q(s)$ and extinction probability s_∞. Show that

$$\Pr(1 \leq X_n \leq k)s_\infty^k \quad \leq \quad Q_n'(s_\infty)$$

   for all $k \geq 1$ and that

$$Q_n'(s_\infty) \quad = \quad Q'(s_\infty)^n.$$

   Use these results and the Borel-Cantelli lemma to prove that

$$\Pr\left(\lim_{n\to\infty} X_n = \infty\right) \quad = \quad 1 - s_\infty.$$

4. Continuing Example 9.2.5, let $P(s)$ be the generating function for the total number of carrier and normal children born to a carrier of the mutant gene. Express the progeny generating function $Q(s)$ of the mutant-gene branching process in terms of $P(s)$. Find the mean μ and variance σ^2 of $Q(s)$ in terms of the mean α and variance β^2 of $P(s)$. How small must α be in order for extinction of the mutant-gene process to be certain?

5. Continuing Problem 36 of Chapter 7, let $T = \min\{n : S_n = 0\}$ be the epoch of the first visit to 0 given $S_0 = 1$. Define $Z_0 = 1$ and

$$Z_j = \sum_{n=0}^{T-1} 1_{\{S_n=j,\ S_{n+1}=j+1\}}$$

for $j \geq 1$. Thus, Z_j is the number of times the random walk moves from j to $j+1$ before hitting 0. Demonstrate that

$$\Pr(Z_1 = k) = \left(\frac{1}{2}\right)^{k+1}.$$

Given $Z_{j-1} = i$, also prove that Z_j can be represented as the sum of i independent copies of Z_1. Therefore, Z_j is a branching process with progeny distribution equal to the distribution of Z_1. What is the generating function of Z_1? How can this be applied to yield the generating function of Z_j for $j > 1$? (Hints: First calculate the probability $\Pr(Z_1 = 0)$. Then observe that each return to state 1 is followed by a step to 0 or an excursion from state 1 back to itself passing through state 2. To calculate the distribution of Z_j for $j > 1$, note that each of the Z_{j-1} steps from $j-1$ to j is followed by an excursion from state j back to state $j-1$. These arguments rely on the fact that the random walk is certain to visit any state starting from any other state.)

6. The generating function $\frac{p}{1-qs}$ is an example of a fractional linear transformation $\frac{\alpha s+\beta}{\gamma s+\delta}$ [93]. To avoid trivial cases where the fractional linear transformation is undefined or constant, we impose the condition that $\alpha\delta - \beta\gamma \neq 0$. The restricted set of fractional linear transformations (or Möbius functions) forms a group under functional composition. This group is the homomorphic image of the group of invertible 2×2 matrices under the correspondence

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \longrightarrow \frac{\alpha s + \beta}{\gamma s + \delta}. \tag{9.20}$$

A group homomorphism is a function between two groups that preserves the underlying algebraic operation. Show that the correspondence (9.20) qualifies as a group homomorphism in the sense that if $f_i(s) = \frac{\alpha_i s+\beta_i}{\gamma_i s+\delta_i}$ for $i = 1, 2$, then

$$\begin{pmatrix} \alpha_1 & \beta_1 \\ \gamma_1 & \delta_1 \end{pmatrix} \begin{pmatrix} \alpha_2 & \beta_2 \\ \gamma_2 & \delta_2 \end{pmatrix} \longrightarrow f_1(f_2(s)).$$

The homomorphism (9.20) correctly pairs the two identity elements $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and $f(s) = s$ of the groups.

7. Suppose $Q(s)$ is a generating function with mean μ and variance σ^2. If $\mu - 1$ is small and positive, then verify that $Q(s)$ has approximate extinction probability $e^{-2(\mu-1)/\sigma^2}$. Show that this approximation equals 0.892 for Lotka's demographic data. (Hints: Put $t = \ln s$ and $L(t) = \ln Q(s)$. Expand $L(t)$ in a second-order Taylor series around $t = 0$.)

8. Let s_∞ be the extinction probability of a supercritical branching process with progeny generating function $Q(s) = \sum_{k=0}^{\infty} q_k s^k$. If the mean μ of $Q(s)$ is fixed, then one can construct counterexamples showing that s_∞ is not necessarily increasing as a function of q_0. As a case in point, let $Q(s) = P\left(\frac{1}{2} + \frac{1}{2}s\right)$ and consider

$$\begin{aligned} P(s) &= \frac{1}{6} + \frac{5}{6}s^3 \\ P(s) &= \frac{3}{32} + \frac{15}{24}s^2 + \frac{5}{32}s^4 + \frac{3}{24}s^5. \end{aligned}$$

   Check numerically that these two choices lead to the extinction probabilities $s_\infty = 0.569$ and $s_\infty = 0.594$ and coefficients $q_0 = 0.271$ and $q_0 = 0.264$.

9. Newton's method offers an alternative method of finding the extinction probability s_∞ of a supercritical generating function $Q(s)$. Let $s_0 = 0$ and $t_0 = 0$ be the initial values in the iteration schemes

$$s_{n+1} = Q(s_n), \qquad t_{n+1} = t_n + \frac{Q(t_n) - t_n}{1 - Q'(t_n)}.$$

   Prove that t_n is an increasing sequence satisfying $s_n \leq t_n \leq s_\infty$ for all $n \geq 0$. It follows that $\lim_{n\to\infty} t_n = s_\infty$ and that Newton's method converges faster than functional iteration.

10. In a subcritical branching process, let T be the generation at which the process goes extinct starting from a single particle at generation 0. If $s_k = \Pr(T \leq k)$ and $Q(s)$ is the progeny generating function, then prove the error estimates

$$\begin{aligned} \frac{Q'(s_n)}{1 - Q'(s_n)}(1 - s_n) &\leq \mathrm{E}(T) - \sum_{k=0}^{n}(1 - s_k) \\ &\leq \frac{Q'(1)}{1 - Q'(1)}(1 - s_n) \end{aligned}$$

    for $\mathrm{E}(T)$. (Hint: Use the first equation in (9.5).)

11. In a branching process, let $R(s) = \sum_{k=0}^{\infty} r_k s^k$ be the generating function of the total number of particles Y_∞ over all generations starting

from a single particle. If the progeny generating function is $Q(s)$, find r_k and the mean and variance of Y_∞ in each of the cases

$$\begin{aligned} Q(s) &= e^{\lambda(s-1)} \\ Q(s) &= (q+ps)^n \\ Q(s) &= \Big(\frac{p}{1-qs}\Big)^n, \end{aligned}$$

where $q = 1 - p$. Furthermore, show that the Poisson, binomial, or negative binomial form of each of these progeny generating functions is preserved when we substitute $\frac{1}{2} + \frac{1}{2}s$ for s. Comment on the relevance of this last result to Problem 4.

12. Suppose X_n denotes the number of particles in a branching process with immigration. Let μ be the mean number of progeny per particle and α the mean number of new immigrants per generation. An ordinary branching process corresponds to the case $\alpha = 0$. For $k > j$ show that

$$\begin{aligned} \mathrm{E}(X_k \mid X_j) &= \begin{cases} (k-j)\alpha + X_j, & \mu = 1 \\ \frac{\alpha(1-\mu^{k-j})}{1-\mu} + \mu^{k-j} X_j, & \mu \neq 1 \end{cases} \\ \mathrm{Cov}(X_k, X_j) &= \mu^{k-j} \operatorname{Var}(X_j). \end{aligned}$$

13. In a subcritical branching process with immigration, let $Q(s)$ be the progeny generating function and $R(s)$ the generating function of the number of new immigrants at each generation. If the equilibrium distribution has generating function $P_\infty(s)$, then show that

$$P_\infty(s) = P_\infty(Q(s))R(s).$$

For the choices $Q(s) = 1 - p + ps$ and $R(s) = e^{-\lambda(1-s)}$, find $P_\infty(s)$. (Hint: Let $P_\infty(s)$ be a Poisson generating function.)

14. Branching processes can be used to model the formation of polymers [154]. Consider a large batch of identical subunits in solution. Each subunit has $m > 1$ reactive sites that can attach to similar reactive sites on other subunits. For the sake of simplicity, assume that a polymer starts from a fixed ancestral subunit and forms a tree structure with no cross linking of existing subunits. Also assume that each reactive site behaves independently and bonds to another site with probability p. Subunits attached to the ancestral subunit form the first generation of a branching process. Subunits attached to these subunits form the second generation and so forth. In this problem we investigate the possibility that polymers of infinite size form. In this case the solution turns into a gel. Show that the progeny distribution

for the first generation is binomial with m trials and success probability p and that the progeny distribution for subsequent generations is binomial with $m-1$ trials and success probability p. Show that the extinction probability t_∞ satisfies

$$\begin{aligned} t_\infty &= (1-p+ps_\infty)^m \\ s_\infty &= (1-p+ps_\infty)^{m-1}, \end{aligned}$$

where s_∞ is the extinction probability for a line of descent emanating from a first-generation subunit. Prove that polymers of infinite size can occur if and only if $(m-1)p > 1$.

15. Yeast cells reproduce by budding. Suppose at each generation a yeast cell either dies with probability p, survives without budding with probability q, or survives with budding off a daughter cell with probability r. In the ordinary branching process paradigm, a surviving cell is considered a new cell. If we refuse to take this view, then what is the distribution of the number of daughter cells budded off by a single yeast cell before its death? Show that the extinction probability of a yeast cell line is 1 when $p \geq r$ and $\frac{p}{r}$ when $p < r$ [68].

16. At an X-linked recessive disease locus, there are two alleles, the normal allele (denoted $+$) and the disease allele (denoted $-$). Construct a two-type branching process for carrier females (genotype $+/-$) and affected males (genotype $-$). Calculate the expected numbers f_{ij} of offspring of each type assuming that carrier females average 2 children, affected males average $2f$ children, all mates are $+/+$ or $+$, and children occur in a 1:1 sex ratio. Note that a branching process model assumes that all children are born simultaneously with the death of a parent. In a continuous-time model, the sensible choice for the death rate λ of either type parent is the reciprocal of the generation time, say about $\frac{1}{25}$ per year in humans.

17. Let $e(s)$ be the vector of extinction probabilities defined in Section 9.6. If the dominant eigenvalue ρ of the matrix Ω has row eigenvector w^t with positive entries and norm $\|w\|_1 = 1$, then demonstrate that $w^t[\mathbf{1} - e(s)] \leq e^{\rho s}$. In the subcritical case with $\rho < 0$, this inequality implies that $\lim_{s\to\infty} e(s) = \mathbf{1}$. It also has consequences for the mean time to extinction. If T_i is the time to extinction starting from a single particle of type i, then show that $\mathrm{E}(T_i) \leq -(\rho w_i)^{-1}$. (Hints: Write

$$\frac{d}{ds}[1-e_i(s)] = -\lambda_i[1-e_i(s)] + \lambda_i \sum_{\mathbf{d}} p_{i\mathbf{d}} \left[1 - \mathbf{e}(s)^{\mathbf{d}}\right],$$

and apply the bound

$$\sum_{\mathbf{d}} p_{i\mathbf{d}}[1-\mathbf{e}(s)^{\mathbf{d}}] \leq \sum_{j=1}^{n} f_{ij}[1-e_j(s)]$$

based on the mean value theorem. Multiply the resulting inequality by w_i and sum on i. When you invoke the mean value theorem, it is helpful to view f_{ij} as a partial derivative of the multivariate generating function $P_i(t, \mathbf{z})$ introduced in Problem 22.)

18. Consider a continuous-time branching process with two types. If the process is irreducible and has reproduction matrix $F = (f_{ij})$, then demonstrate that comparison of the criterion

$$R_0 = f_{11} + f_{22} + f_{12}f_{21} - f_{11}f_{22}$$

to the number 1 determines whether the process is subcritical, critical, or supercritical, subject to either side condition $f_{22} < 1$ or $f_{11} < 1$. Show that the criterion R_0 arises regardless of whether you use type 1 or type 2 as the reference type.

19. Consider a multitype branching process with immigration. Suppose that each particle of type i has an exponential lifetime with death intensity λ_i and produces on average f_{ij} particles of type j at the moment of its death. Independently of death and reproduction, immigrants of type i enter the population according to a Poisson process with intensity α_i. If the Poisson immigration processes for different types are independent, then show that the mean number $m_i(t)$ of particles of type i satisfies the differential equation

$$m_i'(t) = \alpha_i + \sum_j m_j(t)\lambda_j(f_{ji} - 1_{\{j=i\}}).$$

Collecting the $m_i(t)$ and α_i into row vectors $m(t)$ and α, respectively, and the $\lambda_j(f_{ji} - 1_{\{j=i\}})$ into a matrix Ω, show that

$$m(t) = m(0)e^{t\Omega} + \alpha\Omega^{-1}(e^{t\Omega} - I),$$

assuming that Ω is invertible. If we replace the constant immigration intensity α_i by the exponentially decreasing immigration intensity $\alpha_i e^{-\mu t}$, then verify that

$$m(t) = m(0)e^{t\Omega} + \alpha(\Omega + \mu I)^{-1}(e^{t\Omega} - e^{-t\mu I}).$$

20. In a certain species, females die with intensity μ and males with intensity ν. All reproduction is through females at an intensity of λ per female. At each birth, the mother bears a daughter with probability p and a son with probability $1-p$. Interpret this model as a two-type, continuous-time branching process with X_t representing the number of females and Y_t representing the number of males, and show that

$$\mathrm{E}(X_t) = \mathrm{E}(X_0)e^{(\lambda p - \mu)t}$$

$$\begin{aligned}\mathrm{E}(Y_t) &= \mathrm{E}(X_0)\frac{\lambda(1-p)}{\lambda p+\nu-\mu}e^{(\lambda p-\mu)t} \\ &\quad + \left[\mathrm{E}(Y_0)-\mathrm{E}(X_0)\frac{\lambda(1-p)}{\lambda p+\nu-\mu}\right]e^{-\nu t}.\end{aligned}$$

21. In some applications of continuous-time branching processes, it is awkward to model reproduction as occurring simultaneously with death. Birth-death processes offer an attractive alternative. In a birth-death process, a type i particle experiences death at rate μ_i and reproduction of daughter particles of type j at rate β_{ij}. Each reproduction event generates one and only one daughter particle. Thus, in a birth-death process each particle continually buds off daughter particles until it dies. In contrast, each particle of a multitype continuous-time branching process produces a burst of offspring at the moment of its death. This problem considers how we can reconcile these two modes of reproduction. There are two ways of doing this, one exact and one approximate.

    (a) Show that in a birth-death process, a particle of type i produces the count vector $\mathbf{d} = (d_1, \ldots, d_n)$ of daughter particles with probability

    $$p_{i\mathbf{d}} = \frac{\mu_i}{(\mu_i+\beta_i)^{|\mathbf{d}|+1}}\binom{|\mathbf{d}|}{d_1 \ldots d_n}\prod_{k=1}^{n}\beta_{ik}^{d_k},$$

    where $\beta_i = \sum_{j=1}^{n}\beta_{ij}$ and $|\mathbf{d}| = d_1+\cdots+d_n$. (Hint: Condition on the time of death. The number of daughter particles of a given type produced up to this time follows a Poisson distribution.)

    (b) If we delay all offspring until the moment of death, then we get a branching process approximation to the birth-death process. What is the death rate λ_i in the branching process approximation? Show that the approximate process has progeny generating function

    $$\begin{aligned}P_i(\mathbf{s}) &= \sum_{\mathbf{d}} p_{i\mathbf{d}}\mathbf{s}^{\mathbf{d}} \\ &= \sum_{m=0}^{\infty}\frac{\mu_i}{(\mu_i+\beta_i)^{m+1}}\left(\sum_{j=1}^{n}\beta_{ij}s_j\right)^m \\ &= \frac{\mu_i}{\mu_i+\beta_i-\sum_{j=1}^{n}\beta_{ij}s_j}\end{aligned}$$

    for a type i particle.

    (c) In the branching process approximation, demonstrate that a type i particle averages $f_{ij} = \frac{\partial}{\partial s_j}P_i(\mathbf{1}) = \beta_{ij}/\mu_i$ type j daughter particles.

(d) Explain in laymen's terms the meaning of the ratio defining f_{ij}.

(e) Alternatively, we can view the mother particle as dying in one of two ways. Either it dies in the ordinary way at rate μ_i, or it disappears at a reproduction event and is replaced by an identical substitute and a single daughter particle. Eventually one of the substitute particles dies in the ordinary way before reproducing, corresponding to death in the original birth-death process. What is the death rate λ_i in this exact branching process analog of the birth-death process? Justify in words the progeny generating function

$$P_i(\mathbf{s}) = \frac{\mu_i}{\mu_i + \beta_i} + \sum_{k=1}^{r} \frac{\beta_{ik}}{\mu_i + \beta_i} s_i s_k.$$

(f) We can turn the approximate correspondence discussed in parts (b) and (c) around and seek to mimic a branching process by a birth-death process. The most natural method involves matching the mean number of daughter particles f_{ij} in the branching process to the mean number of daughter particles budded off in the birth-death process. Given the λ_i and the f_{ij}, what are the natural values for the death rates μ_i and the birth rates β_{ij} in the birth-death approximation to the branching process?

22. Consider a multitype continuous-time branching process with n particle types. Let X_{it} be the number of particles of type i at time t. Section 9.6 derives a system of ordinary differential equations describing the extinction probabilities at time t starting from a single particle of any type at time 0. It is possible to derive a similar system of ordinary differential equations describing each of the multivariate generating functions $P_i(t, \mathbf{z}) = \mathrm{E}(\mathbf{z}^{X_t})$ for the full distribution of the random vector $X_t = (X_{1t}, \ldots, X_{nt})$ starting from a single particle of type i. In the notation of Section 9.6, show that

$$\frac{\partial}{\partial t} P_i(t, \mathbf{z}) = -\lambda_i P_i(t, \mathbf{z}) + \lambda_i \sum_{\mathbf{d}} p_{i\mathbf{d}} P(t, \mathbf{z})^{\mathbf{d}}, \qquad (9.21)$$

where $P(t, \mathbf{z}) = [P_1(t, \mathbf{z}), \ldots, P_n(t, \mathbf{z})]$. (Hints: Write the generating function $P_i(t, \mathbf{z})$ as the expectation $\mathrm{E}(\mathbf{z}^{X_t} \mid X_0 = e_i)$, where e_i is the usual ith unit vector. The ancestral particle either lives throughout the short time interval $(0, s)$ or dies and reproduces. In the latter case, each daughter particle initiates an independent clan that evolves according to the rules governing the clan issuing from an ancestral particle of the same type.)

23. Continuing Problem 22, consider calculation of the variance $\mathrm{Var}(X_t)$ for a single-type process starting from a single particle. Suppose that

the progeny generating function $\sum_{d=0}^{\infty} p_d z^d$ has mean μ and variance σ^2 and the death intensity is λ. (Note here that we drop the subscript from p_d denoting the beginning particle type.) If $m_1(t)$ and $m_2(t)$ are the first and second factorial moments of X_t, then derive the ordinary differential equations

$$\begin{aligned}\frac{d}{dt}m_1(t) &= -\lambda m_1(t) + \lambda\mu m_1(t) \\ \frac{d}{dt}m_2(t) &= -\lambda m_2(t) + \lambda(\sigma^2 + \mu^2 - \mu)m_1(t)^2 + \lambda\mu m_2(t).\end{aligned}$$

by differentiating equation (9.21) with respect to z and setting $z = 1$. What are the initial values $m_1(0)$ and $m_2(0)$? Solve these differential equations, and demonstrate that

$$\mathrm{Var}(X_t) = \begin{cases} \frac{\sigma^2+\mu^2-2\mu+1}{\mu-1}\left[e^{2\lambda(\mu-1)t} - e^{\lambda(\mu-1)t}\right], & \mu \neq 1 \\ \sigma^2\lambda t, & \mu = 1. \end{cases}$$

What is $\mathrm{Var}(X_t)$ starting from $j > 1$ particles?

24. Continuing Problem 22, consider a single-type process starting from a single particle. Suppose that the progeny generating function is $\sum_{d=0}^{\infty} p_d z^d = z^k$ for some $k \geq 2$ and the death intensity is λ. Show that the ordinary differential equation

$$\frac{\partial}{\partial t}P(t,z) = -\lambda P(t,z) + \lambda P(t,z)^k$$

for the generating function of X_t has solution

$$P(t,z) = ze^{-\lambda t}\left[1 - z^{k-1} + e^{-\lambda(k-1)t}z^{k-1}\right]^{-1/(k-1)}.$$

Also check the initial condition $P(0,z) = z$. What is $P(t,z)$ if you start from $j > 1$ particles?

25. In the cancer model of Coldman and Goldie [41], cancer cells are of two types. Type 1 cells are ordinary cancer cells. Type 2 particles are cancer cells with resistance to an anti-cancer chemotherapeutic agent. In culture, particles of both types live an exponential length of time of average duration λ^{-1} and then divide. Type 2 cells always produce type 2 cells, but type 1 cells produce two cells of type 1 with probability $1-\alpha$ or the combination of one cell of type 1 and one cell of type 2 with probability α. The process commences with a single type 1 cell. One can make considerable progress understanding this system using the notation and results of Problem 22.

(a) Show that the multivariate generating functions characterizing this process satisfy the differential equations

$$\begin{aligned}\frac{\partial}{\partial t}P_1(t,\mathbf{z}) &= -\lambda P_1(t,\mathbf{z}) + \lambda(1-\alpha)P_1(t,\mathbf{z})^2 \\ &\quad +\lambda\alpha P_1(t,\mathbf{z})P_2(t,\mathbf{z}) \\ \frac{\partial}{\partial t}P_2(t,\mathbf{z}) &= -\lambda P_2(t,\mathbf{z}) + \lambda P_2(t,\mathbf{z})^2.\end{aligned}$$

(b) Subject to the initial conditions $P_1(0,\mathbf{z}) = z_1$ and $P_2(0,\mathbf{z}) = z_2$, demonstrate that these equations have solutions

$$\begin{aligned}P_1(t,\mathbf{z}) &= \frac{z_1 e^{-\lambda t}(e^{-\lambda t}z_2 + 1 - z_2)^{-\alpha}}{1 + z_1[(e^{-\lambda t}z_2 + 1 - z_2)^{1-\alpha} - 1]z_2^{-1}} \\ P_2(t,\mathbf{z}) &= \frac{z_2}{z_2 + (1-z_2)e^{\lambda t}}.\end{aligned}$$

(Hints: Solve for $P_2(t,\mathbf{z})$ first. Use the solution

$$f(t) = \frac{f(0)e^{\int_0^t g(s)ds}}{1 - cf(0)\int_0^t e^{\int_0^s g(r)dr}ds} \tag{9.22}$$

to the Riccati differential equation $f'(t) = g(t)f(t) + cf(t)^2$ to solve for both $P_2(t,\mathbf{z})$ and $P_1(t,\mathbf{z})$.

(c) Derive the solution to the Riccati equation (9.22) by writing a linear differential equation for $h(t) = 1/f(t)$.

(d) Prove that the probability of no type 2 particles at time t is

$$P_1[t,(1,0)] = \frac{1}{1-\alpha+\alpha e^{\lambda t}}.$$

(e) Find the value of t such that $P_1[t,(1,0)] = 1/2$. Note that this time is relatively short. Thus, therapy should be as prompt and as radical as possible.

26. Although null and deleterious mutations commonly occur in human cells, cancer initiation and progression is driven by mutations that increase cell fitness. These fitness advantages take the form of higher than normal birth rates and/or lower than normal death rates (apoptosis). In practical terms the cascade of mutations involves activation of oncogenes, disabling of tumor suppressor genes, and enhancement of chromosome instability. Multitype branching processes can be used to model the evolution of a cancer cell line. For the sake of simplicity, suppose that a cell can accumulate anywhere from 1 to n mutations and that the number of mutations is the sole determinant of fitness.

A cell is labeled by the number of mutations it carries. Assume a type k cell dies at rate δ_k, splits into two daughter cells of type k at rate β_k, and splits into one daughter cell of type k and one daughter cell of type $k+1$ at rate μ_k. To keep the number of cell types finite, take $\mu_n = 0$. All other rates are positive. Show that the $n \times n$ matrix $\Omega = (\omega_{kl})$ determining the mean behavior of the process has entries

$$\omega_{kl} = \begin{cases} \beta_k - \delta_k, & l = k \\ \mu_k, & l = k+1 \\ 0, & \text{otherwise.} \end{cases}$$

Is Ω irreducible? What are its eigenvalues? State conditions under which the mean population size will tend to 0 or grow explosively. You may assume that the process starts with at least one cell of type 1. The most subtle behavior occurs when $\beta_k \leq \delta_k$ for all k and $\beta_k = \delta_k$ for some k. Prove that the mean number of cells of type k never decreases when $\beta_k = \delta_k$.

Now let e_k be the extinction probability of the cancer line starting from a single cell of type k. Show that these probabilities satisfy the equations

$$\begin{aligned} e_n &= \frac{\beta_n}{\lambda_n} e_n^2 + \frac{\delta_n}{\lambda_n} \\ e_k &= \frac{\beta_k}{\lambda_k} e_k^2 + \frac{\mu_k}{\lambda_k} e_k e_{k+1} + \frac{\delta_k}{\lambda_k}, \quad 1 \leq k < n, \end{aligned}$$

where $\lambda_k = \beta_k + \delta_k + \mu_k$. Argue that the e_k can be calculated recursively starting with

$$e_n = \begin{cases} 1, & \beta_n \leq \delta_n \\ \frac{\delta_n}{\beta_n}, & \beta_n > \delta_n. \end{cases}$$

In solving the various quadratics, why should one always take the left root?

10
Martingales

10.1 Introduction

Martingales generalize the notion of a fair game in gambling. Theory to the contrary, many gamblers still believe that they simply need to hone their strategies to beat the house. Probabilists know better. The real payoff with martingales is their practical value throughout probability theory. This chapter introduces martingales, develops some relevant theory, and delves into a few applications. As a prelude, readers are urged to review the material on conditional expectations in Chapter 1. In the current chapter we briefly touch on the convergence properties of martingales, the optional stopping theorem, and large deviation bounds via Azuma's inequality. More extensive treatments of martingale theory appear in the books [23, 24, 53, 80, 106, 118, 208]. Our other referenced sources either provide elementary accounts comparable in difficulty to the current material [129, 170] or interesting special applications [4, 134, 186, 201].

10.2 Definition and Examples

A sequence of integrable random variables X_n forms a martingale relative to a second sequence of random variables Y_n provided

$$\mathrm{E}(X_{n+1} \mid Y_1, \ldots, Y_n) \quad = \quad X_n \tag{10.1}$$

K. Lange, *Applied Probability,*
DOI 10.1007/978-1-4419-7165-4_10, © Springer Science+Business Media, LLC 2010

for all $n \geq 1$. In many applications $X_n = Y_n$. It saves space and is often conceptually simpler to replace the collection of random variables $Y_1, \ldots, Y_n$ by the σ-algebra of events $\mathcal{F}_n$ that it generates. Note that these σ-algebras are increasing in the sense that any $A \in \mathcal{F}_n$ satisfies $A \in \mathcal{F}_{n+1}$. This relationship is written $\mathcal{F}_n \subset \mathcal{F}_{n+1}$, and the sequence $\mathcal{F}_n$ is said to be a filter. This somewhat odd terminology is motivated by the fact that the $\mathcal{F}_n$ contain more information or detail as n increases. In any case, we now rephrase the definition of a martingale to be a sequence of integrable random variables X_n satisfying

$$\mathrm{E}(X_{n+1} \mid \mathcal{F}_n) \;=\; X_n \tag{10.2}$$

relative to a filter $\mathcal{F}_n$. Readers who find this definition unnecessarily abstract are urged to fall back on the original definition (10.1).

Before turning to specific examples, let us deduce a few elementary properties of martingales. First, equation (10.2) implies that the random variable X_n is measurable with respect to $\mathcal{F}_n$. When $\mathcal{F}_n$ is generated by $Y_1, \ldots, Y_n$, then X_n is expressible as a measurable function of $Y_1, \ldots, Y_n$. Second, iteration of the identity

$$\begin{aligned} \mathrm{E}(X_{n+1}) &= \mathrm{E}[\mathrm{E}(X_{n+1} \mid \mathcal{F}_n)] \\ &= \mathrm{E}(X_n) \end{aligned}$$

leads to $\mathrm{E}(X_n) = \mathrm{E}(X_1)$ for all $n > 1$. Third, an obvious inductive argument using the tower property (1.6) gives

$$\begin{aligned} \mathrm{E}(X_{n+k} \mid \mathcal{F}_n) &= \mathrm{E}[\mathrm{E}(X_{n+k} \mid \mathcal{F}_{n+1}) \mid \mathcal{F}_n] \\ &= \mathrm{E}(X_{n+1} \mid \mathcal{F}_n) \\ &= X_n \end{aligned}$$

for all $k > 0$. Fourth, if the X_n are square-integrable and $i \leq j \leq k \leq l$, then

$$\begin{aligned} \mathrm{E}[(X_j - X_i)(X_l - X_k)] &= \mathrm{E}\{\mathrm{E}[(X_j - X_i)(X_l - X_k) \mid \mathcal{F}_j]\} \\ &= \mathrm{E}\{(X_j - X_i)\,\mathrm{E}[(X_l - X_k) \mid \mathcal{F}_j]\} \\ &= \mathrm{E}[(X_j - X_i)(X_j - X_j)] \\ &= 0. \end{aligned}$$

In other words, the increments $X_j - X_i$ and $X_l - X_k$ are orthogonal. This fact allows us to calculate the variance of

$$X_n \;=\; X_1 + \sum_{i=2}^{n} (X_i - X_{i-1})$$

via

$$\mathrm{Var}(X_n) \;=\; \mathrm{Var}(X_1) + \sum_{i=2}^{n} \mathrm{Var}(X_i - X_{i-1}), \tag{10.3}$$

exactly as if we were dealing with sums of independent random variables. The fact that $\mathrm{Var}(X_n)$ is increasing in n follows immediately from the decomposition (10.3). Finally, the special case

$$\mathrm{Var}(X_{m+n} - X_m) \quad = \quad \mathrm{Var}(X_{m+n}) - \mathrm{Var}(X_m) \tag{10.4}$$

of the orthogonal increments property is worth highlighting.

Example 10.2.1 *Sums of Random Variables*

If Y_n is a sequence of independent random variables with common mean $\mu = 0$, then the partial sums $S_n = Y_1 + \cdots + Y_n$ constitute a martingale relative to the filter $\mathcal{F}_n$ generated by $\{Y_1, \ldots, Y_n\}$. The martingale property (10.2) follows from the calculation

$$\begin{aligned} \mathrm{E}(S_{n+1} \mid \mathcal{F}_n) &= \mathrm{E}(Y_{n+1} \mid \mathcal{F}_n) + \mathrm{E}(S_n \mid \mathcal{F}_n) \\ &= \mathrm{E}(Y_{n+1}) + S_n \\ &= S_n. \end{aligned}$$

When $\mu \neq 0$, the modified sequence $S_n - n\mu$ forms a martingale. If the Y_n are dependent random variables, then the sequence

$$\begin{aligned} S_n &= \sum_{i=1}^{n} [Y_i - \mathrm{E}(Y_i \mid \mathcal{F}_{i-1})] \\ &= S_{n-1} + Y_n - \mathrm{E}(Y_n \mid \mathcal{F}_{n-1}) \end{aligned}$$

provides a zero-mean martingale because

$$\begin{aligned} \mathrm{E}(S_{n+1} \mid \mathcal{F}_n) &= S_n + \mathrm{E}(Y_{n+1} \mid \mathcal{F}_n) - \mathrm{E}(Y_{n+1} \mid \mathcal{F}_n) \\ &= S_n. \end{aligned}$$

■

Example 10.2.2 *Products of Independent Random Variables*

Similarly if Y_n is a sequence of independent random variables with common mean $\mu = \mathrm{E}(Y_1) = 1$, then the partial products $X_n = \prod_{i=1}^{n} Y_i$ constitute a martingale relative to the filter $\mathcal{F}_n$ generated by $\{Y_1, \ldots, Y_n\}$. The martingale property (10.2) follows from the calculation

$$\begin{aligned} \mathrm{E}\left(\prod_{i=1}^{n+1} Y_i \,\middle|\, \mathcal{F}_n \right) &= \mathrm{E}(Y_{n+1} \mid \mathcal{F}_n) \prod_{i=1}^{n} Y_i \\ &= \prod_{i=1}^{n} Y_i. \end{aligned}$$

When $\mu \neq 1$, the modified sequence $\mu^{-n} X_n$ is a martingale.

This martingale arises in Wald's theory of sequential testing in statistics. Let $Z_1, Z_2, \ldots$ be an i.i.d. sequence of random variables with density $f(z)$ under the simple null hypothesis H_o and density $g(z)$ under the simple alternative hypothesis H_a. Both of these densities are relative to a common measure μ such as Lebesgue measure or counting measure. If H_o is true, then

$$\int \frac{g(z)}{f(z)} f(z)\, d\mu(z) \;=\; \int g(z)\, d\mu(z) \;=\; 1.$$

It follows that the likelihood ratio statistics

$$X_n \;=\; \prod_{i=1}^{n} \frac{g(Z_i)}{f(Z_i)}$$

constitute a martingale. ■

Example 10.2.3 *Martingale Differences*

If X_n is a martingale with respect to the filter $\mathcal{F}_n$, then the difference sequence $\{X_{m+n} - X_m\}_{n\geq 1}$ is a martingale with respect to the shifted filter $\{\mathcal{F}_{m+n}\}_{n\geq 1}$. This assertion is a consequence of the identity

$$\mathrm{E}(X_{m+n} - X_m \mid \mathcal{F}_{m+n-1}) \;=\; X_{m+n-1} - X_m.$$

■

Example 10.2.4 *Doob's Martingale*

Let Z be an integrable random variable and $\mathcal{F}_n$ be a filter. Then the sequence $X_n = \mathrm{E}(Z \mid \mathcal{F}_n)$ is a martingale because of the tower property

$$\mathrm{E}[\mathrm{E}(Z \mid \mathcal{F}_{n+1}) \mid \mathcal{F}_n] \;=\; \mathrm{E}(Z \mid \mathcal{F}_n).$$

In many examples, $\mathcal{F}_n$ is the σ-algebra defined by n independent random variables $Y_1, \ldots, Y_n$. If the random variable Z is a function of $Y_1, \ldots, Y_m$ alone, then $X_m = Z$, and for $1 \leq n < m$

$$\begin{aligned} X_n(y_1, \ldots, y_n) &= \mathrm{E}(Z \mid Y_1 = y_1, \ldots, Y_n = y_n) \qquad (10.5) \\ &= \int \cdots \int Z(y_1, \ldots, y_m)\, dF_{n+1}(y_{n+1}) \cdots dF_m(y_m), \end{aligned}$$

where $F_k(y_k)$ is the distribution function of Y_k. In other words, the conditional expectation X_n is generated by integrating over the last $m - n$ arguments of Z. This formula for X_n is correct because Fubini's theorem gives

$$\begin{aligned} &\mathrm{E}[1_A(Y_1, \ldots, Y_n) Z] \\ = \;& \int \cdots \int 1_A(y_1, \ldots, y_n) X_n(y_1, \ldots, y_n)\, dF_1(y_1) \cdots dF_n(y_n) \end{aligned}$$

for any event A depending only on $Y_1, \ldots, Y_n$. In some applications, martingales of finite length appear. If $X_1, \ldots, X_m$ is a finite martingale relative to the filter $\mathcal{F}_1 \subset \cdots \subset \mathcal{F}_m$, then $Z = X_m$ satisfies Doob's requirement $X_n = \mathrm{E}(Z \mid \mathcal{F}_n)$ for every $n \leq m$. ■

Example 10.2.5 *Branching Processes*

Suppose Y_n counts the number of particles at generation n in a discrete-time branching process. (Here the index n starts at 0 rather than 1.) Let $\mathcal{F}_n$ be the σ-algebra generated by $Y_0, \ldots, Y_n$, and let μ be the mean of the progeny distribution. We argued in Chapter 9 that

$$\mathrm{E}(Y_{n+1} \mid Y_n) \;=\; \mathrm{E}(Y_{n+1} \mid \mathcal{F}_n) \;=\; \mu Y_n.$$

It follows that $X_n = \mu^{-n} Y_n$ is a martingale relative to the filter $\mathcal{F}_n$. ■

Example 10.2.6 *Wright-Fisher Model of Genetic Drift*

In the Wright-Fisher model of Example 7.3.2, the proportion $X_n = \frac{1}{2m} Y_n$ of a_1 alleles at generation n provides a martingale relative to the filter $\mathcal{F}_n$ determined by the random counts $Y_1, \ldots, Y_n$ of a_1 alleles at generations 1 through n. Indeed,

$$\begin{aligned} \mathrm{E}(X_{n+1} \mid \mathcal{F}_n) &= \frac{1}{2m} \mathrm{E}(Y_{n+1} \mid Y_n) \\ &= X_n \end{aligned}$$

is obvious from the nature of the binomial sampling with success probability X_n in forming the population at generation $n + 1$. ■

10.3 Martingale Convergence

Our purpose in this section is to inquire when a martingale X_n possesses a limit. The theory is much simpler if we stipulate that the X_n have finite second moments. This condition is not sufficient for convergence as Example 10.2.1 shows. However, if the second moments $\mathrm{E}(X_n^2)$ are uniformly bounded, then in general we get a convergent sequence. The next proposition paves the way by generalizing Chebyshev's inequality.

Proposition 10.3.1 (Kolmogorov) *Suppose that relative to the filter $\mathcal{F}_n$ the square-integrable martingale X_n has mean $\mathrm{E}(X_n) = 0$. The inequality*

$$\Pr(\max\{|X_1|, \ldots, |X_n|\} > \epsilon) \;\leq\; \frac{\mathrm{Var}(X_n)}{\epsilon^2} \tag{10.6}$$

then holds for all $\epsilon > 0$ and $n \geq 1$.

Proof: For the purposes of this proof, we can reduce $\mathcal{F}_n$ to the σ-algebra of events generated by $X_1, \ldots, X_n$. Let $M = m$ be the smallest subscript between 1 and n such that $|X_m| > \epsilon$. If no such subscript exists, then set $M = 0$. The calculation

$$\begin{aligned} \mathrm{E}(X_n^2) &= \mathrm{E}(X_n^2 1_{\{M=0\}}) + \sum_{m=1}^{n} \mathrm{E}(X_n^2 1_{\{M=m\}}) \\ &\geq \sum_{m=1}^{n} \mathrm{E}(X_n^2 1_{\{M=m\}}) \end{aligned} \tag{10.7}$$

follows because the events $\{M = 0\}, \ldots, \{M = n\}$ are mutually exclusive and exhaustive. In view of the fact that the event $\{M = m\} \in \mathcal{F}_m$, we have

$$\begin{aligned} \mathrm{E}(X_n^2 1_{\{M=m\}}) &= \mathrm{E}[1_{\{M=m\}} \mathrm{E}(X_n^2 \mid \mathcal{F}_m)] \\ &= \mathrm{E}\{1_{\{M=m\}} \mathrm{E}[(X_n - X_m)^2 + X_m^2 \mid \mathcal{F}_m]\} \\ &\geq \mathrm{E}[1_{\{M=m\}} \mathrm{E}(X_m^2 \mid \mathcal{F}_m)] \\ &= \mathrm{E}(1_{\{M=m\}} X_m^2) \\ &\geq \epsilon^2 \Pr(M = m). \end{aligned} \tag{10.8}$$

In this derivation, we have employed the identities

$$\mathrm{E}[X_m(X_n - X_m) \mid \mathcal{F}_m] = X_m(X_m - X_m) = 0$$

and

$$\mathrm{E}[X_n(X_n - X_m) \mid \mathcal{F}_m] = \mathrm{E}[X_n^2 \mid \mathcal{F}_m] - X_m^2.$$

Combining inequality (10.8) with inequality (10.7) yields

$$\mathrm{E}(X_n^2) \geq \epsilon^2 \sum_{m=1}^{n} \Pr(M = m) = \epsilon^2 \Pr(M > 0),$$

which is clearly equivalent to inequality (10.6). ■

Proposition 10.3.2 *Suppose the martingale X_n relative to the filter $\mathcal{F}_n$ has uniformly bounded second moments. Then $X_\infty = \lim_{n\to\infty} X_n$ exists almost surely and*

$$\lim_{n\to\infty} \mathrm{E}[(X_\infty - X_n)^2] = 0. \tag{10.9}$$

Proof: To prove that X_∞ exists almost surely, it suffices to show that the sequence X_n is almost surely a Cauchy sequence. Let A_{km} be the event $\{|X_{m+n} - X_m| > \frac{1}{k} \text{ for some } n \geq 1\}$. On the complement of the event $A = \bigcup_{k\geq 1} \bigcap_{m\geq 1} A_{km}$, the sequence X_n is Cauchy. The inequality

$$\Pr(A) \leq \sum_{k\geq 1} \inf_{m\geq 1} \Pr(A_{km})$$

suggests that we verify $\liminf_{m\to\infty} \Pr(A_{km}) = 0$. To achieve this goal, we apply Proposition 10.3.1 to the difference $X_{m+n} - X_m$, taking into account Example 10.2.3 and equality (10.4). It follows that

$$\begin{aligned}
\Pr(A_{km}) &= \lim_{l\to\infty} \Pr\Big(\max_{1\le n\le l} |X_{m+n} - X_m| > \frac{1}{k}\Big) \\
&\le \lim_{l\to\infty} k^2 \operatorname{Var}(X_{m+l} - X_m) \\
&= \lim_{l\to\infty} k^2[\operatorname{Var}(X_{m+l}) - \operatorname{Var}(X_m)].
\end{aligned}$$

Because the sequence $\operatorname{Var}(X_n)$ is increasing and uniformly bounded, this last inequality yields

$$\begin{aligned}
0 &\le \liminf_{m\to\infty} \Pr(A_{km}) \\
&\le k^2 \liminf_{m\to\infty} \lim_{l\to\infty} [\operatorname{Var}(X_{m+l}) - \operatorname{Var}(X_m)] \\
&= 0,
\end{aligned}$$

which finishes the proof that X_n converges almost surely.

To establish the limit (10.9), we note that Fatou's lemma and equality (10.4) imply

$$\begin{aligned}
0 &\le \lim_{m\to\infty} \mathrm{E}[(X_\infty - X_m)^2] \\
&= \lim_{m\to\infty} \mathrm{E}\Big[\lim_{n\to\infty} (X_{m+n} - X_m)^2\Big] \\
&\le \lim_{m\to\infty} \liminf_{n\to\infty} \mathrm{E}[(X_{m+n} - X_m)^2] \\
&= \lim_{m\to\infty} \liminf_{n\to\infty} [\operatorname{Var}(X_{m+n}) - \operatorname{Var}(X_m)] \\
&= 0.
\end{aligned}$$

This completes the proof. ■

Example 10.3.1 *Strong Law of Large Numbers*

Let Y_n be a sequence of independent random variables with common mean μ and variance σ^2. According to the analysis of Example 10.2.1, the sums $X_n = \sum_{i=1}^n i^{-1}(Y_i - \mu)$ constitute a martingale. Because $\mathrm{E}(X_n) = 0$ and $\operatorname{Var}(X_n) = \sigma^2 \sum_{i=1}^n i^{-2}$, this martingale has uniformly bounded second moments and therefore converges almost surely by Proposition 10.3.2. In view of the identity $Y_i - \mu = i(X_i - X_{i-1})$, we can write

$$\frac{1}{n}\sum_{i=1}^n (Y_i - \mu) = \frac{1}{n}\Big(\sum_{i=1}^n iX_i - \sum_{i=1}^n iX_{i-1}\Big) = X_n - \frac{1}{n}\sum_{i=1}^n X_{i-1} \quad (10.10)$$

with the convention $X_0 = 0$. Because $\lim_{n\to\infty} X_n = X_\infty$ exists, it is easy to show that

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n X_{i-1} = X_\infty$$

as well. In conjunction with equation (10.10), this yields

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n}(Y_i - \mu) \;=\; X_\infty - X_\infty \;=\; 0 \tag{10.11}$$

and proves the strong law of large numbers. Other proofs exist that do not require the Y_n to have finite variance. ■

Example 10.3.2 *Convergence of a Supercritical Branching Process*

Returning to Example 10.2.5, suppose that the process is supercritical. Taking into account that $\mathrm{E}(X_n) = 1$ by design, we can invoke Proposition 10.3.2 provided the variances $\mathrm{Var}(X_n) = \mu^{-2n}\,\mathrm{Var}(Y_n)$ are uniformly bounded. If σ^2 is the variance of the progeny distribution, then equation (9.3) indicates that

$$\begin{aligned}\mathrm{Var}(X_n) &= \frac{\sigma^2(1-\frac{1}{\mu^n})}{\mu(\mu-1)}\\ &\le \frac{\sigma^2}{\mu(\mu-1)}\end{aligned}$$

for $\mu > 1$. Thus, Proposition 10.3.2 implies that $\lim_{n\to\infty} X_n = X_\infty$ exists.

Finding the distribution of X_∞ is difficult. For a geometric progeny generating function $Q(s) = \frac{p}{1-qs}$, progress can be made by deriving a functional equation characterizing the Laplace transform $L_\infty(t) = \mathrm{E}(e^{-tX_\infty})$. If $Q_n(s)$ is the probability generating function of Y_n, then the Laplace transform of X_n can be expressed as

$$L_n(t) \;=\; \mathrm{E}(e^{-\frac{tY_n}{\mu^n}}) \;=\; Q_n(e^{-\frac{t}{\mu^n}}).$$

In view of the defining equation $Q_{n+1}(s) = Q(Q_n(s))$, this leads directly to $L_{n+1}(\mu t) = Q(L_n(t))$, which produces the functional equation

$$L_\infty(\mu t) \;=\; Q(L_\infty(t)) \tag{10.12}$$

after taking limits on n. Straightforward algebra shows that the fractional linear transformation

$$L_\infty(t) \;=\; \frac{pt-p+q}{qt-p+q}$$

solves equation (10.12) when $Q(s) = \frac{p}{1-qs}$ and $\mu = \frac{q}{p}$. To identify the distribution with Laplace transform $L_\infty(t)$, note that

$$\frac{pt-p+q}{qt-p+q} \;=\; re^0 + (1-r)\int_0^\infty e^{-tx}(1-r)e^{-(1-r)x}dx \tag{10.13}$$

for $r = \frac{p}{q}$. In other words, X_∞ is a mixture of a point mass at 0 and an exponential distribution with intensity $1-r$. The total mass r at 0 is just the extinction probability in this case. ■

10.4 Optional Stopping

Many applications of martingales involve stopping times. A stopping time T is a random variable that is adapted to a filter $\mathcal{F}_n$. The possible values of T are ∞ and the nonnegative integers. The word "adapted" here is a technical term meaning that the event $\{T = n\}$ is in $\mathcal{F}_n$ for all n. In less precise language, a stopping time can only depend on the past and present and cannot anticipate the future. A typical stopping time associated with a martingale X_n is $T_A = \min\{n : X_n \in A\}$, the first entry into a Borel set A of real numbers. The next proposition gives sufficient conditions for the mean value of the stopped process X_T to equal the martingale mean $\mathrm{E}(X_n)$.

Proposition 10.4.1 (Optional Stopping Theorem) *Let X_n be a martingale with common mean μ relative to the filter $\mathcal{F}_n$. If T is a stopping time for X_n satisfying*

(a) $\Pr(T < \infty) = 1$,

(b) $\mathrm{E}(|X_T|) < \infty$,

(c) $\lim_{n\to\infty} \mathrm{E}(X_n 1_{\{T>n\}}) = 0$,

then $\mathrm{E}(X_T) = \mu$. *Condition (c) holds if the second moments* $\mathrm{E}(X_n^2)$ *are uniformly bounded.*

Proof: In view of the fact that the event $\{T = i\} \in \mathcal{F}_i$, we have

$$\begin{aligned}
\mathrm{E}(X_T) &= \mathrm{E}(X_T 1_{\{T>n\}}) + \sum_{i=1}^{n} \mathrm{E}(X_T 1_{\{T=i\}}) \\
&= \mathrm{E}(X_T 1_{\{T>n\}}) + \sum_{i=1}^{n} \mathrm{E}(X_i 1_{\{T=i\}}) \\
&= \mathrm{E}(X_T 1_{\{T>n\}}) + \sum_{i=1}^{n} \mathrm{E}[\mathrm{E}(X_n \mid \mathcal{F}_i) 1_{\{T=i\}}] \\
&= \mathrm{E}(X_T 1_{\{T>n\}}) + \sum_{i=1}^{n} \mathrm{E}(X_n 1_{\{T=i\}}) \\
&= \mathrm{E}(X_T 1_{\{T>n\}}) + \mathrm{E}(X_n 1_{\{T\le n\}}).
\end{aligned}$$

Subtracting $\mu = \mathrm{E}(X_n 1_{\{T>n\}}) + \mathrm{E}(X_n 1_{\{T\le n\}})$ from this identity yields

$$\mathrm{E}(X_T) - \mu = \mathrm{E}(X_T 1_{\{T>n\}}) - \mathrm{E}(X_n 1_{\{T>n\}}). \tag{10.14}$$

The dominated convergence theorem implies $\lim_{n\to\infty} \mathrm{E}(X_T 1_{\{T>n\}}) = 0$. This takes care of the first term on the right-hand side of equation (10.14).

The second term tends to 0 by condition (c). Owing to the Cauchy-Schwarz inequality

$$\mathrm{E}(X_n 1_{\{T>n\}})^2 \quad \leq \quad \mathrm{E}(X_n^2)\Pr(T>n), \tag{10.15}$$

condition (c) holds whenever the second moments $\mathrm{E}(X_n^2)$ are uniformly bounded. ■

Example 10.4.1 *Wald's Identity and the Sex Ratio*

Consider Example 10.2.1 with the understanding that $\sigma^2 = \mathrm{Var}(Y_n)$ is finite, $\mu = \mathrm{E}(Y_n)$ is not necessarily 0, and the Y_n are independent. If T is a stopping time relative to the filter $\mathcal{F}_n$ generated by $Y_1, \ldots, Y_n$, then Wald's identity says that the stopped sum S_T has mean $\mathrm{E}(S_T) = \mu\,\mathrm{E}(T)$ provided $\mathrm{E}(T) < \infty$. We have already visited this problem when T is independent of the Y_n. To prove Wald's identity in general, we apply Proposition 10.4.1 to the martingale $R_n = S_n - n\mu$.

Part (b) of the proposition requires checking that $\mathrm{E}(|R_T|) < \infty$. This inequality follows from the calculation

$$\begin{aligned}
\mathrm{E}(|R_T|) &\leq \mathrm{E}\left(\sum_{i=1}^{T} |Y_i - \mu|\right) \\
&= \mathrm{E}\left(\sum_{n=1}^{\infty}\sum_{i=1}^{n} |Y_i - \mu| 1_{\{T=n\}}\right) \\
&= \mathrm{E}\left(\sum_{i=1}^{\infty} |Y_i - \mu| 1_{\{T\geq i\}}\right) \\
&= \sum_{i=1}^{\infty} \mathrm{E}(|Y_i - \mu|)\Pr(T \geq i) \\
&\leq \sigma\,\mathrm{E}(T).
\end{aligned}$$

Here we have used Schlömilch's inequality $\mathrm{E}(|Y_i - \mu|) \leq \sigma$ and the fact that $1_{\{T\geq i\}} = 1_{\{T>i-1\}}$ depends only on $Y_1, \ldots, Y_{i-1}$ and consequently is independent of Y_i. Part (c) of Proposition 10.4.1 is validated by noting that inequality (10.15) can be continued to

$$\begin{aligned}
\mathrm{E}(R_n 1_{\{T>n\}})^2 &\leq \mathrm{E}(R_n^2)\Pr(T>n) \\
&= n\sigma^2 \Pr(T>n) \\
&\leq \sigma^2 \sum_{i=n+1}^{\infty} i \Pr(T=i).
\end{aligned}$$

Because $\mathrm{E}(T) < \infty$ by assumption, $\lim_{n\to\infty} \sum_{i=n+1}^{\infty} i \Pr(T=i) = 0$. This completes the proof.

The family planning model discussed in Examples 2.3.3 and 6.6.2 involves stopping times $T = N_{sd}$ that could conceivably change the sex ratio of females to males. However, Wald's identity rules this out. If Y_i is the indicator random variable recording whether the ith birth is female, then S_T is the number of daughters born to a couple with stopping time T. Wald's identity $\mathrm{E}(S_T) = q\,\mathrm{E}(T)$ implies that the proportion of daughters

$$\frac{\mathrm{E}(S_T)}{\mathrm{E}(T)} \quad = \quad q$$

over a large number of such families does not deviate from q. ■

Example 10.4.2 *Hitting Probabilities in the Wright-Fisher Model*

Because the proportion $X_n = \frac{1}{2m}Y_n$ of a_1 alleles at generation n is a martingale, Proposition 10.4.1 can be employed to calculate the probability of eventual fixation of the a_1 allele. Condition (a) of the proposition holds because the underlying Markov chain must reach one of the two absorbing states 0 or $2m$. Conditions (b) and (c) are trivial to verify in light of the inequalities $0 \leq X_n \leq 1$. If T is the time of absorption at 0 or 1 and $Y_1 = i$, then Proposition 10.4.1 implies

$$\begin{aligned} \frac{i}{2m} &= \mathrm{E}(X_1) \\ &= \mathrm{E}(X_T) \\ &= 0 \cdot \Pr(X_T = 0) + 1 \cdot \Pr(X_T = 1). \end{aligned}$$

Thus, the a_1 allele is eventually fixed with probability $\frac{i}{2m}$. ■

In many cases checking the conditions of Proposition 10.4.1 is onerous. Hence, the following alternative version of the optional stopping theorem is convenient [209].

Proposition 10.4.2 *In Proposition 10.4.1, suppose that the stopping time T is finite with probability 1 and that for $n \leq T$ we can write either*

$$X_n \quad = \quad B_n + I_n \quad \text{or} \quad X_n \quad = \quad B_n - I_n,$$

where $|B_n| \leq d$ and $0 \leq I_{n-1} \leq I_n$. In other words for all times up to T, the B process is bounded and the I process is increasing. Then the identity $\mathrm{E}(X_T) = \mu$ holds without making assumptions (b) and (c) of the proposition.

Proof: We first note that $T \wedge m = \min\{T, m\}$ is a stopping time because the event $\{T \wedge m \leq n\} = \{T \leq n\} \cup \{m \leq n\}$ is certainly in $\mathcal{F}_n$. For this stopping time, condition (a) of the proposition is obvious, condition (b) follows from the inequality $|X_{T\wedge m}| \leq |X_1| + \cdots + |X_m|$, and condition (c)

is a consequence of the fact that $1_{\{T\wedge m>n\}} = 0$ for $n > m$. The proposition therefore yields

$$\mathrm{E}(X_{T\wedge m}) \;=\; \mathrm{E}(B_{T\wedge m}) \pm \mathrm{E}(I_{T\wedge m}) \;=\; \mu.$$

Since $T \wedge m$ increases to its limit T, the monotone convergence theorem implies that $\mathrm{E}(I_T)$ exists. To rule out an infinite value for the limit, we simply observe that the inequality

$$|\mathrm{E}(I_{T\wedge m}) \mp \mu| \;\leq\; d$$

holds for all m. To conclude the proof, it suffices to apply the dominated convergence theorem to the sequence $X_{T\wedge m}$ with dominating random variable $d + I_T$. ■

Example 10.4.3 *Gambler's Ruin Problem*

Consider a random walk that takes steps of -1 and $+1$ with probabilities p and $q = 1 - p$, respectively. If we start the walk at 0, then it is of considerable interest to calculate the probability r_{ab} that the walk reaches $-a$ before it reaches b, where a and b are nonnegative integers. In the equivalent gambler's ruin problem, one starts at a and conducts the walk until it reaches 0 (ruin of the gambler) or $a + b$ (ruin of the house). Let X_i denote the outcome of trial i. The associated random variable $Y_i = (q/p)^{X_i}$ has mean 1. According to Example 10.2.2, the sequence

$$Z_n \;=\; \prod_{i=1}^{n} Y_i \;=\; \prod_{i=1}^{n} \Big(\frac{q}{p}\Big)^{X_i}$$

is a martingale. Let T be the random epoch at which the random walk hits $-a$ or b. If we assume that $p \neq q$ and apply the strong law of large numbers to the sequence X_i, then it is obvious that the barrier $-a$ is eventually hit when $p < q$ and the barrier b is eventually hit when $p > q$. Therefore, T is finite, and Proposition 10.4.2 with $I_n = 0$ implies

$$r_{ab}\Big(\frac{p}{q}\Big)^{a} + (1 - r_{ab})\Big(\frac{q}{p}\Big)^{b} \;=\; 1.$$

Trivial algebra now yields

$$r_{ab} \;=\; \frac{1 - \Big(\frac{q}{p}\Big)^{b}}{\Big(\frac{p}{q}\Big)^{a} - \Big(\frac{q}{p}\Big)^{b}}. \tag{10.16}$$

Except for notational differences, this reproduces the solution found in Problem 38 of Chapter 7. Problem 16 of this chapter treats the symmetric case $p = q$ using a different martingale. Problems 16 and 17 calculate the mean number of steps until ruin. ■

Example 10.4.4 *Successive Random Permutations*

Consider the following recursive construction. Let Y_1 be the number of matches in a random permutation π_1 of the set $\{1, \ldots, m\}$. Throw out each integer i for which $\pi_1(i) = i$, and relabel the remaining integers $1, \ldots, m - Y_1$. Let Y_2 be the number of matches in an independent random permutation π_2 of the set $\{1, \ldots, m - Y_1\}$. Throw out each integer i for which $\pi_2(i) = i$, and relabel the remaining integers $1, \ldots, m - Y_1 - Y_2$. Continue this process N times until $\sum_{i=1}^{N} Y_i = m$.

We now calculate the mean of the random variable N by exploiting the martingale

$$\begin{aligned} X_n &= \sum_{i=1}^{n} [Y_i - \mathrm{E}(Y_i \mid \mathcal{F}_{i-1})] \\ &= \sum_{i=1}^{n} (Y_i - 1) \end{aligned}$$

relative to the filter $\mathcal{F}_n$ generated by $Y_1, \ldots, Y_n$. Example 10.2.1 establishes the martingale property, and Example 2.2.1 proves the identity

$$\mathrm{E}(Y_i \mid \mathcal{F}_{i-1}) = 1.$$

This is a perfect situation in which to apply Proposition 10.4.2. We merely observe that the sequence $X_n = B_n - I_n$ is a difference of a bounded sequence $B_n = \sum_{i=1}^{n} Y_i$ of random variables with bound m and an increasing sequence $I_n = n$ of constants. Proposition 10.4.2 therefore permits us to conclude that

$$\begin{aligned} 0 &= \mathrm{E}(X_N) \\ &= \mathrm{E}\Big(\sum_{i=1}^{N} Y_i\Big) - \mathrm{E}(N) \\ &= m - \mathrm{E}(N), \end{aligned}$$

which obviously entails $\mathrm{E}(N) = m$. ■

Example 10.4.5 *Sequential Testing in Statistics*

Consider Wald's likelihood ratio martingale of Example 10.2.2. Suppose we quit sampling at the first epoch T with either $X_T \geq \alpha^{-1}$ or $X_T \leq \beta$. In the former case, we decide in favor of the alternative hypothesis H_a, and in the latter case, we decide in favor of the null hypothesis H_o. We claim that the type I and type II errors satisfy

$$\begin{aligned} \Pr(\text{reject } H_o \mid H_o) &\leq \alpha \\ \Pr(\text{reject } H_a \mid H_a) &\leq \beta. \end{aligned} \tag{10.17}$$

To prove this claim, we follow the arguments of Williams [209], taking for granted that $\Pr(T < \infty) = 1$. If we set $T \wedge n = \min\{T, n\}$ and assume H_o, then applying Proposition 10.4.1 to $T \wedge n$ and invoking Fatou's lemma yield

$$\begin{aligned}
\alpha^{-1} \Pr(X_T \geq \alpha^{-1}) &\leq \mathrm{E}(X_T) \\
&= \mathrm{E}\left(\lim_{n\to\infty} X_{T\wedge n}\right) \\
&\leq \liminf_{n\to\infty} \mathrm{E}(X_{T\wedge n}) \\
&= \mathrm{E}(X_1) \\
&= 1.
\end{aligned}$$

The extremes of this last string of inequalities produce inequality (10.17). In general, sequential testing is more efficient in reaching a decision than testing with a fixed sample size. ■

10.5 Large Deviation Bounds

In Chapter 1 we investigated various classical inequalities such as Chebyshev's inequality. For well-behaved random variables, much sharper results are possible. The next proposition gives Azuma's tail-probability bound for martingales with bounded differences. Hoeffding originally established the bound for sums of independent random variables.

Proposition 10.5.1 (Azuma-Hoeffding) *Suppose the sequence of random variables X_n forms a martingale with mean 0 relative to the filter $\mathcal{F}_n$. If under the convention $X_0 = 0$ there exists a sequence of constants c_n such that $\Pr(|X_n - X_{n-1}| \leq c_n) = 1$, then*

$$\mathrm{E}\left(e^{\beta X_n}\right) \leq e^{(\beta^2/2)\sum_{k=1}^n c_k^2} \tag{10.18}$$

for all $\beta > 0$. Inequality (10.18) entails the further inequalities

$$\Pr(X_n \geq \lambda) \leq e^{-\lambda^2/(2\sum_{k=1}^n c_k^2)} \tag{10.19}$$

and

$$\Pr(|X_n| \geq \lambda) \leq 2e^{-\lambda^2/(2\sum_{k=1}^n c_k^2)} \tag{10.20}$$

for all $\lambda > 0$.

Proof: Because the function e^u is convex, $e^{\alpha u+(1-\alpha)v} \leq \alpha e^u + (1-\alpha)e^v$ for any $\alpha \in [0,1]$ and pair u and v. Putting $u = -\beta c$, $v = \beta c$, and $\alpha = \frac{c-x}{2c}$ for $x \in [-c, c]$ therefore yields

$$e^{\beta x} \leq \frac{c-x}{2c} e^{-\beta c} + \frac{c+x}{2c} e^{\beta c}.$$

If we substitute $X_n - X_{n-1}$ for x and c_n for c in this inequality and take conditional expectations, then the hypothesis $|X_n - X_{n-1}| \leq c_n$ and the fact $\mathrm{E}(X_n - X_{n-1} \mid \mathcal{F}_{n-1}) = 0$ imply

$$\begin{aligned}
\mathrm{E}\left(e^{\beta X_n}\right) &= \mathrm{E}\left[\mathrm{E}\left(e^{\beta X_n} \mid \mathcal{F}_{n-1}\right)\right] \\
&= \mathrm{E}\left\{e^{\beta X_{n-1}} \mathrm{E}\left[e^{\beta(X_n - X_{n-1})} \mid \mathcal{F}_{n-1}\right]\right\} \\
&\leq \mathrm{E}\left(e^{\beta X_{n-1}}\right)\left(\frac{c_n - 0}{2c_n} e^{-\beta c_n} + \frac{c_n + 0}{2c_n} e^{\beta c_n}\right) \\
&= \mathrm{E}\left(e^{\beta X_{n-1}}\right) \frac{e^{-\beta c_n} + e^{\beta c_n}}{2}.
\end{aligned}$$

Induction on n and the convention $X_0 = 0$ now prove inequality (10.18) provided we can show that

$$\frac{e^u + e^{-u}}{2} \leq e^{\frac{u^2}{2}}. \tag{10.21}$$

However, inequality (10.21) follows by expanding its right and left sides in Taylor's series and noting that the corresponding coefficients of u^{2n} satisfy $\frac{1}{(2n)!} \leq \frac{1}{2^n n!}$. Of course, the coefficients of the odd powers u^{2n+1} on both sides of inequality (10.21) vanish.

To prove inequality (10.19), we apply Markov's inequality in the form

$$\begin{aligned}
\Pr(X_n \geq \lambda) &\leq \mathrm{E}\left(e^{\beta X_n}\right) e^{-\beta\lambda} \\
&\leq e^{(\beta^2/2)\sum_{k=1}^n c_k^2 - \beta\lambda}
\end{aligned} \tag{10.22}$$

for an arbitrary $\beta > 0$. The particular β minimizing the exponent on the right-hand side of (10.22) is clearly $\beta = \lambda/(\sum_{k=1}^n c_k^2)$. Substituting this choice in inequality (10.22) yields inequality (10.19). Because $-X_n$ also fulfills the hypotheses of the proposition, inequality (10.20) follows from inequality (10.19). ■

In applying Proposition 10.5.1 to a martingale X_n with nonzero mean μ, we replace X_n by the recentered martingale $X_n - \mu$. Recentering has no impact on the differences $(X_n - \mu) - (X_{n-1} - \mu) = X_n - X_{n-1}$, so the above proof holds without change.

Example 10.5.1 *Tail Bound for the Binomial Distribution*

Suppose Y_n is a sequence of independent random variables with common mean $\mu = \mathrm{E}(Y_n)$. If $|Y_n - \mu| \leq c_n$ with probability 1, then Proposition 10.5.1 applies to the martingale $S_n - n\mu = \sum_{i=1}^n Y_i - n\mu$. In particular, when each Y_n is a Bernoulli random variable with success probability μ, then $c_n = 1$ and

$$\Pr(|S_n - n\mu| \geq \lambda) \leq 2e^{-\frac{\lambda^2}{2n}}. \tag{10.23}$$

Inequality (10.23) has an interesting implication for the sequence of Bernstein polynomials that approximate a continuous function $f(x)$ satisfying a Lipschitz condition $|f(u) - f(v)| \leq d|u - v|$ for all $u, v \in [0, 1]$ [82]. As in Example 3.5.1, we put $\mu = x$ and argue that

$$\begin{aligned} & \left| \mathrm{E}\left[f\Big(\frac{S_n}{n}\Big)\right] - f(x)\right| \\ \leq \;& d\frac{\lambda}{n} \Pr\left(\left|\frac{S_n}{n} - x\right| < \frac{\lambda}{n}\right) + 2||f||_\infty \Pr\left(\left|\frac{S_n}{n} - x\right| \geq \frac{\lambda}{n}\right). \end{aligned}$$

Invoking the large deviation bound (10.23) for the second probability gives the inequality

$$\left| \mathrm{E}\left[f\Big(\frac{S_n}{n}\Big)\right] - f(x)\right| \leq d\frac{\lambda}{n} + 4||f||_\infty e^{-\frac{\lambda^2}{2n}}.$$

For the choice $\lambda = \sqrt{2n \ln n}$, this yields the uniform bound

$$\begin{aligned} \left| \mathrm{E}\left[f\Big(\frac{S_n}{n}\Big)\right] - f(x)\right| &\leq d\sqrt{\frac{2\ln n}{n}} + \frac{4||f||_\infty}{n} \\ &= O\left(\sqrt{\frac{\ln n}{n}}\right). \end{aligned}$$

This is an improvement over the uniform bound $O(n^{-1/3})$ for continuously differentiable functions noted in Problem 21 of Chapter 3. It is worse than the best available uniform bound $O(n^{-1})$ for functions that are twice continuously differentiable rather than merely Lipschitz [162]. ■

Many applications of Proposition 10.5.1 involve a more complicated combination of independent random variables $Y_1, \ldots, Y_n$ than a sum. If Z is such a combination and $\mathcal{F}_i$ is the σ-algebra generated by $Y_1, \ldots, Y_i$, then Example 10.2.4 implies that the conditional expectations $X_i = \mathrm{E}(Z \mid \mathcal{F}_i)$ constitute a finite-length martingale with $X_n = Z$. To apply Proposition 10.5.1, we observe that equation (10.5) implies

$$\begin{aligned} & X_i(y_1, \ldots, y_i) \\ = \;& \int \cdots \int Z(y_1, \ldots, y_{i-1}, y_i, v_{i+1}, \ldots, v_n)\, d\omega_i(v_i) \cdots d\omega_n(v_n) \\ & X_{i-1}(y_1, \ldots, y_{i-1}) \qquad (10.24) \\ = \;& \int \cdots \int Z(y_1, \ldots, y_{i-1}, v_i, v_{i+1}, \ldots, v_n)\, d\omega_i(v_i) \cdots d\omega_n(v_n), \end{aligned}$$

where ω_k is the multivariate distribution of Y_k. Obviously, the extra integration on ω_i introduced in the multiple integral defining X_i has no effect. Taking differences in (10.24) produces an integral expression for $X_i - X_{i-1}$ whose integrand is

$$\begin{aligned} & Z_i - Z_{i-1} \qquad (10.25) \\ = \;& Z(y_1, \ldots, y_{i-1}, y_i, v_{i+1}, \ldots, v_n) - Z(y_1, \ldots, y_{i-1}, v_i, v_{i+1}, \ldots, v_n). \end{aligned}$$

If $|Z_i - Z_{i-1}| \leq c_i$ for some constant c_i, then the inequality $|X_i - X_{i-1}| \leq c_i$ follows after integration against the independent probability distributions $\omega_i, \ldots, \omega_n$. These considerations set the stage for applying Proposition 10.5.1 to $Z = X_n$.

Example 10.5.2 *Longest Common Subsequence*

In the longest common subsequence problem considered in Example 5.7.1, we can derive tail-probability bounds for M_n. Let Y_i be the random pair of letters chosen for position i of the two strings. Conditioning on the σ-algebra $\mathcal{F}_i$ generated by $Y_1, \ldots, Y_i$ creates a martingale $X_i = \mathrm{E}(M_n \mid \mathcal{F}_i)$ with $X_n = M_n$. We now bound $|X_i - X_{i-1}|$ by considering what happens when exactly one argument of $M_n(y_1, \ldots, y_n)$ changes. If this is argument i, then the pair $y_i = (u_i, v_i)$ becomes the pair $y_i^* = (u_i^*, v_i^*)$. In a longest common subsequence of $y = (y_1, \ldots, y_n)$, changing u_i to u_i^* creates or destroys at most one match. Likewise, changing v_i to v_i^* creates or destroys at most one match. Thus, the revised string y^* has a common subsequence whose length differs from $M_n(y)$ by at most 2, and the pertinent inequality $M_n(y) - 2 \leq M_n(y^*)$ follows. By the same token, $M_n(y^*) - 2 \leq M_n(y)$. Hence $|M_n(y) - M_n(y^*)| \leq 2$, and Proposition 10.5.1 yields

$$\Pr[|M_n - \mathrm{E}(M_n)| \geq \lambda\sqrt{n}] \quad \leq \quad 2e^{-\frac{\lambda^2}{8}}.$$

Further problems of this sort are treated in the references [186, 201]. ■

Example 10.5.3 *Euclidean Traveling Salesman Problem*

In some cases it is possible to calculate a more subtle bound on the martingale differences $X_{i+1} - X_i$. The Euclidean traveling salesman problem is typical in this regard. In applying Proposition 10.5.1 to Example 5.7.2, we take Z to be $D_n(\{Y_1, \ldots, Y_n\})$ and $\mathcal{F}_i$ to be the σ-algebra generated by $Y_1, \ldots, Y_i$. Consider the integrand (10.25) determining the martingale difference $X_i - X_{i-1}$. If S denotes the set $S = \{y_1, \ldots, y_{i-1}, v_{i+1}, \ldots, v_n\}$, then the reasoning that produces inequality (5.12) in Example 5.7.2 also leads to the inequalities

$$\begin{aligned} D_{n-1}(S) \quad &\leq \quad D_n(S \cup \{y_i\}) \quad \leq D_{n-1}(S) + 2\min_{w \in S} ||w - y_i|| \\ D_{n-1}(S) \quad &\leq \quad D_n(S \cup \{v_i\}) \quad \leq D_{n-1}(S) + 2\min_{w \in S} ||w - v_i|| \end{aligned}$$

involving $Z_i = D_n(S \cup \{y_i\})$ and $Z_{i-1} = D_n(S \cup \{v_i\})$. It follows that

$$\begin{aligned} |Z_i - Z_{i-1}| \quad &\leq \quad 2\min_{w \in S} ||w - y_i|| + 2\min_{w \in S} ||w - v_i|| \\ &\leq \quad 2\min_{j > i} ||v_j - y_i|| + 2\min_{j > i} ||v_j - v_i|| \end{aligned}$$

and consequently that

$$|X_i - X_{i-1}| \quad \leq \quad 2\,\mathrm{E}\left(\min_{j > i} ||Y_j - y_i||\right) + 2\,\mathrm{E}\left(\min_{j > i} ||Y_j - Y_i||\right).$$

We now estimate the right-tail probability $\Pr(\min_{j>i} ||Y_j - y|| \geq r)$ for a generic point y representing either y_i or Y_i. The smallest area of the unit square at a distance of r or less from y is a quarter-circle of radius r. This extreme case occurs when y occupies a corner of the square. In view of this result and the inequality $(1-x)^k \leq e^{-kx}$ for $x \in (0,1)$ and $k > 0$, we have

$$\Pr\left(\min_{j>i} ||Y_j - y|| \geq r\right) \leq \left(1 - \frac{\pi r^2}{4}\right)^{n-i} \leq e^{-\frac{(n-i)\pi r^2}{4}}.$$

Application of Example 2.5.1 therefore yields

$$\begin{aligned} \mathrm{E}\left(\min_{j>i} ||Y_j - y||\right) &\leq \int_0^\infty e^{-\frac{(n-i)\pi r^2}{4}} dr \\ &= \frac{1}{2}\int_{-\infty}^\infty e^{-\frac{(n-i)\pi r^2}{4}} dr \\ &= \frac{1}{\sqrt{n-i}} \end{aligned}$$

and

$$|X_i - X_{i-1}| \leq \frac{4}{\sqrt{n-i}}.$$

The case $i = n$ must be considered separately. If we use the crude inequality

$$|X_n - X_{n-1}| = |D_n - X_{n-1}| \leq 2\sqrt{2},$$

then the sum $\sum_{i=1}^n c_i^2$ figuring in Proposition 10.5.1 can be bounded by

$$\sum_{i=1}^n c_i^2 \leq (2\sqrt{2})^2 + 4^2 \sum_{i=1}^{n-1} \frac{1}{n-i} \leq 8 + 16(\ln n + 1).$$

This in turn translates into the Azuma-Hoeffding bound

$$\Pr[|D_n - \mathrm{E}(D_n)| \geq \lambda] \leq 2e^{-\frac{\lambda^2}{48+32\ln n}}.$$

Problem 25 asks the reader to check some details of this argument. ■

10.6 Problems

1. Define the random variables Y_n inductively by taking $Y_0 = 1$ and Y_{n+1} to be uniformly distributed on the interval $(0, Y_n)$. Show that the sequence $X_n = 2^n Y_n$ is a martingale.

2. An urn contains b black balls and w white balls. Each time we randomly withdraw a ball, we replace it by $c+1$ balls of the same color. Let X_n be the fraction of white balls after n draws. Demonstrate that X_n is a martingale.

3. Let $Y_1, Y_2, \ldots$ be a sequence of independent random variables with zero means and common variance σ^2. If $X_n = Y_1 + \cdots + Y_n$, then show that $X_n^2 - n\sigma^2$ is a martingale.

4. Let $Y_1, Y_2, \ldots$ be a sequence of i.i.d. random variables with common moment generating function $M(t) = \mathrm{E}(e^{tY_1})$. Prove that

$$X_n \quad = \quad M(t)^{-n} e^{t(Y_1+\cdots+Y_n)}$$

is a martingale whenever $M(t) < \infty$.

5. Let Y_n be a finite-state, discrete-time Markov chain with transition matrix $P = (p_{ij})$. If v is a column eigenvector for P with nonzero eigenvalue λ, then verify that $X_n = \lambda^{-n} v_{Y_n}$ is a martingale, where v_{Y_n} is coordinate Y_n of v.

6. Suppose Y_t is a continuous-time Markov chain with infinitesimal generator Ω. Let v be a column eigenvector of Ω with eigenvalue λ. Show that $X_t = e^{-\lambda t} v_{Y_t}$ is a martingale in the sense that

$$\mathrm{E}(X_{t+s} \mid Y_r,\, r \in [0,t]) \quad = \quad X_t.$$

Here v_{Y_t} is coordinate Y_t of v.

7. Let $\{X_n\}_{n\geq 0}$ be a family of random variables with finite expectations that satisfy

$$\mathrm{E}(X_{n+1} \mid X_1, \ldots, X_n) \quad = \quad \alpha X_n + (1-\alpha) X_{n-1}$$

for $n \geq 1$ and some constant $\alpha \neq 1$. Find a second constant β so that the random variables $Y_n = \beta X_n + X_{n-1}$ for $n \geq 1$ constitute a martingale relative to $\{X_n\}_{n\geq 0}$.

8. Suppose Y_n is the number of particles at the nth generation of a branching process. If s_∞ is the extinction probability, prove that $X_n = s_\infty^{Y_n}$ is a martingale. (Hint: If $Q(s)$ is the progeny generating function, then $Q(s_\infty) = s_\infty$.)

9. Let N_t denote the number of random points that occur by time t in a Poisson process on $[0, \infty)$ with intensity λ. Show that the following stochastic processes

$$\begin{aligned} X_t &= N_t - \lambda t \\ X_t &= (N_t - \lambda t)^2 - \lambda t \\ X_t &= e^{-\theta N_t + \lambda t(1-e^{-\theta})} \end{aligned}$$

enjoy the martingale property $\mathrm{E}(X_{t+s} |\, N_r,\, r \in [0,t]) = X_t$ for $s > 0$. (Hint: $N_{t+s} - N_t$ is independent of N_t and distributed as N_s.)

10. In Example 10.3.2, show that $\text{Var}(X_\infty) = \frac{\sigma^2}{\mu(\mu-1)}$ by differentiating equation (10.12) twice. This result is consistent with the mean square convergence displayed in equation (10.9).

11. In Example 10.3.2, show that the fractional linear transformation

$$L_\infty(t) = \frac{pt-p+q}{qt-p+q}$$

solves equation (10.12) when $Q(s) = \frac{p}{1-qs}$ and $\mu = \frac{q}{p}$. Also verify equation (10.13).

12. In Example 10.2.2, suppose that each Y_n is equally likely to assume the values $\frac{1}{2}$ and $\frac{3}{2}$. Show that $\prod_{i=1}^\infty Y_i \equiv 0$ but $\prod_{i=1}^\infty E(Y_i) = 1$ [24]. (Hint: Apply the strong law of large numbers to the sequence $\ln Y_n$.)

13. Given $X_0 = \mu \in (0,1)$, define X_n inductively by

$$X_{n+1} = \begin{cases} \alpha + \beta X_n, & \text{with probability } X_n \\ \beta X_n, & \text{with probability } 1 - X_n\,, \end{cases}$$

where $\alpha, \beta > 0$ and $\alpha+\beta = 1$. Prove that X_n is a martingale with (a) $X_n \in (0,1)$, (b) $\text{E}(X_n) = \mu$, and (c) $\text{Var}(X_n) = [1-(1-\alpha^2)^n]\mu(1-\mu)$. Also prove that Proposition 10.3.2 implies that $\lim_{n\to\infty} X_n = X_\infty$ exists with $\text{E}(X_\infty) = \mu$ and $\text{Var}(X_\infty) = \mu(1-\mu)$. (Hint: Derive a recurrence relation for $\text{Var}(X_{n+1})$ by conditioning on X_n.)

14. In Proposition 10.3.2, prove that $X_n = \text{E}(X_\infty \mid \mathcal{F}_n)$. If X_n is defined by $X_n = \text{E}(X \mid \mathcal{F}_n)$ to begin with, then one can also show that $X_\infty = \text{E}(X \mid \mathcal{F}_\infty)$, where $\mathcal{F}_\infty$ is the smallest σ-algebra containing $\cup_n \mathcal{F}_n$. Omit this further technicality. (Hint: For $C \in \mathcal{F}_n$ take the limit of $\text{E}(X_{n+m}1_C) = \text{E}(X_n 1_C)$ as m tends to ∞.)

15. Let $Y_1, Y_2, \ldots$ be a sequence of independent random variables. The tail σ-algebra $\mathcal{T}$ generated by the sequence can be expressed as $\mathcal{T} = \cap_n \mathcal{T}_n$, where $\mathcal{T}_n$ is the σ-algebra generated by $Y_n, Y_{n+1}, \ldots$. It is easy to construct events in $\mathcal{T}$. For instance in an infinite sequence of coin tosses, the event that a finite number of heads occurs belongs to $\mathcal{T}$. The zero-one law says that any $C \in \mathcal{T}$ has $\Pr(C) = 0$ or $\Pr(C) = 1$. Use Proposition 10.3.2 and Problem 14 to prove the zero-one law. (Hints: Let $\mathcal{F}_n$ be the σ-algebra generated by $Y_1, \ldots, Y_n$. The martingale $X_n = \text{E}(1_C \mid \mathcal{F}_n)$ is constant and converges to X_∞, which is measurable with respect to $\mathcal{F}_\infty$. But C is a member of $\mathcal{F}_\infty$.)

16. Let $S_n = X_1 + \cdots + X_n$ be a symmetric random walk on the integers $\{-a, \ldots, b\}$ starting at $S_0 = 0$. For the stopping time

$$T = \min\{n : S_n = -a \text{ or } S_n = b\},$$

prove that $\Pr(S_T = b) = a/(a+b)$ by considering the martingale S_n and that $\mathrm{E}(T) = ab$ by considering the martingale $S_n^2 - n$. (Hints: Apply Proposition 10.4.2 and Problem 3.)

17. Continuing Problem 16, assume that the random walk is asymmetric and moves to the right with probability p and to the left with probability $q = 1 - p$. Show that the stopping time T has mean

$$\mathrm{E}(T) \quad = \quad \frac{1}{p-q}\left[(1-r_{ab})b - r_{ab}a\right],$$

where r_{ab} is the ruin probability (10.16). (Hint: Apply Proposition 10.4.2 to the martingale $S_n - n(p-q)$.)

18. In the Wright-Fisher model of Example 10.2.6, show that

$$Z_n \quad = \quad \frac{X_n(1-X_n)}{\left(1-\frac{1}{2m}\right)^n}$$

is a martingale. Assuming that $\lim_{n\to\infty} Z_n = Z_\infty$ exists, we have $X_n(1-X_n) \approx \left(1-\frac{1}{2m}\right)^n Z_\infty$ for n large. In other words, X_n approaches either 0 or 1 at rate $1 - \frac{1}{2m}$.

19. Continuing Problem 18, let T be the time of absorption at 0 or 1 starting from $Y_0 = i$ copies of the a_1 allele. Demonstrate that

$$\Pr(T > n) \quad \leq \quad i(2m-i)\left(1-\frac{1}{2m}\right)^n \quad \leq \quad \epsilon$$

for $\epsilon \in (0,1)$ and $n = 2m\ln[i(2m-i)] - 2m\ln\epsilon$.

20. Continuing Problem 9, let T_n be the time at which N_t first equals the positive integer n. Assuming the optional stopping theorem holds for the stopping time T_n and the martingales identified in Problem 9, show that

$$\begin{aligned} \mathrm{E}(T_n) &= \frac{n}{\lambda} \\ \mathrm{Var}(T_n) &= \frac{n}{\lambda^2} \\ \mathrm{E}\left(e^{-\beta T_n}\right) &= \left(\frac{\lambda}{\lambda+\beta}\right)^n \end{aligned}$$

for $\beta > 0$. These results agree with our earlier findings concerning the mean, variance, and Laplace transform of T_n. (Hints: Use $N_{T_n} = n$ and set $\beta = -\lambda(1-e^{-\theta})$ for the third equality.)

21. Let $Y_1, \ldots, Y_n$ be independent Bernoulli random variables with success probability μ. Graphically compare the large deviation bound (10.23) to Chebyshev's bound

$$\Pr(|S_n - n\mu| \geq \lambda) \leq \frac{n\mu(1-\mu)}{\lambda^2}$$

when $\mu = 1/2$. Which bound is better? If neither is uniformly better than the other, determine which combinations of values of n and λ favor Chebyshev's bound.

22. Suppose that $v_1, \ldots, v_n \in \mathsf{R}^m$ have Euclidean norms $\|v_i\|_2 \leq 1$. Let $Y_1, \ldots, Y_n$ be independent random variables uniformly distributed on the two-point set $\{-1, 1\}$. If $Z = \|Y_1 v_1 + \cdots + Y_n v_n\|_2$, then prove that

$$\Pr[Z - \mathrm{E}(Z) \geq \lambda\sqrt{n}] \quad \leq \quad e^{-\frac{\lambda^2}{8}}.$$

23. Consider a random graph with n nodes. Between every pair of nodes, independently introduce an edge with probability p. The graph is said to be k colorable if it is possible to assign each of its nodes one of k colors so that no pair of adjacent nodes share the same color. The chromatic number X of the graph is the minimum value of k. Demonstrate that $\Pr[|X - \mathrm{E}(X)| \geq \lambda] \leq 2e^{-\lambda^2/(2n)}$. (Hint: Consider the martingale $X_i = \mathrm{E}(X \mid Y_1, \ldots, Y_i)$, where Y_i is the random set of edges connecting node i to nodes $1, \ldots, i-1$.)

24. Consider a multinomial experiment with n trials, m possible cells, and success probability p_i for cell i. Let S_k be the number of cells with exactly k successes. Show that

$$\mathrm{E}(S_k) \quad = \quad \sum_{i=1}^{m} \binom{n}{k} p_i^k (1-p_i)^{n-k}.$$

Apply the Azuma-Hoeffding theorem and prove that

$$\begin{aligned} \Pr[|S_0 - \mathrm{E}(S_0)| \geq \lambda] &\leq 2e^{-\lambda^2/(2n)} \\ \Pr[|S_k - \mathrm{E}(S_k)| \geq \lambda] &\leq 2e^{-\lambda^2/(8n)}, \qquad k > 0. \end{aligned}$$

(Hint: Let X_i be the martingale $\mathrm{E}(S_k \mid Y_1, \ldots, Y_i)$, where Y_i is the outcome of trial i.)

25. Example 10.5.3 relies on some unsubstantiated claims. Prove that: (a) $(1-x)^k \leq e^{-kx}$ for $x \in (0,1)$ and $k > 0$, (b) $|D_n - X_{n-1}| \leq 2\sqrt{2}$, and (c) $1 + \cdots + \frac{1}{n-1} \leq \ln n + 1$.

11
Diffusion Processes

11.1 Introduction

Despite their reputation for sophistication, diffusion processes are widely applied throughout science and engineering. Here we survey the theory at an elementary level, stressing intuition rather than rigor. Readers with the time and mathematical inclination should follow up this brief account by delving into serious presentations of the mathematics [80, 107]. A good grounding in measure theory is indispensable in understanding the theory. At the highest level of abstraction, diffusion processes can be treated via the Ito stochastic integral [30, 38]. As background for this chapter, the reader is advised to review the material in Section 1.8 on the multivariate normal distribution.

Because of the hard work involved in mastering the abstract theory, there is a great deal to be said for starting with specific applications and heuristic arguments. Brownian motion is the simplest and most pervasive diffusion process. Several more complicated processes can be constructed from standard Brownian motion. The current chapter follows a few selected applications from population biology, neurophysiology, and population genetics. These serve to illustrate concrete techniques for calculating moments, first passage times, and equilibrium distributions. Ordinary and partial differential equations play a prominent role in these computations.

K. Lange, *Applied Probability,*
DOI 10.1007/978-1-4419-7165-4_11, © Springer Science+Business Media, LLC 2010

11.2 Basic Definitions and Properties

A diffusion process X_t is a continuous-time Markov process with approximately Gaussian increments over small time intervals. Its sample paths $t \mapsto X_t$ are continuous functions of t over a large interval I, either finite or infinite. The process X_t is determined by the Markovian assumption and the distribution of its increments. For small s and $X_t = x$, the increment $X_{t+s} - X_t$ is nearly Gaussian (normal) with mean and variance

$$\begin{aligned} \mathrm{E}(X_{t+s} - X_t \mid X_t = x) &= \mu(t,x)s + o(s) && (11.1) \\ \mathrm{E}[(X_{t+s} - X_t)^2 \mid X_t = x] &= \sigma^2(t,x)s + o(s). && (11.2) \end{aligned}$$

The functions $\mu(t,x)$ and $\sigma^2(t,x) \geq 0$ are called the infinitesimal mean and variance, respectively. Here the term "infinitesimal variance" is used rather than "infinitesimal second moment" because the approximation

$$\begin{aligned} & \mathrm{Var}(X_{t+s} - X_t \mid X_t = x) \\ = \; & \mathrm{E}[(X_{t+s} - X_t)^2 \mid X_t = x] - [\mu(t,x)s + o(s)]^2 \\ = \; & \mathrm{E}[(X_{t+s} - X_t)^2 \mid X_t = x] + o(s) \end{aligned}$$

follows directly from approximation (11.1). If the infinitesimal mean and variance do not depend on time t, then the process is time homogeneous. The choices $\mu(t,x) = 0$, $\sigma^2(t,x) = 1$, and $X_0 = 0$ characterize standard Brownian motion.

To begin our nonrigorous, intuitive discussion of diffusion processes, we note that the normality assumption implies

$$\begin{aligned} \mathrm{E}(|X_{t+s} - X_t|^m \mid X_t = x) &= \mathrm{E}\left(\left|\frac{X_{t+s} - X_t}{\sigma(t,x)\sqrt{s}}\right|^m \;\middle|\; X_t = x\right)\left[\sigma(t,x)\sqrt{s}\right]^m \\ &= o(s) && (11.3) \end{aligned}$$

for $m > 2$. (See Problem 1.) This insight is crucial in various arguments involving Taylor series expansions. For instance, it allows us to deduce how X_t behaves under a smooth, invertible transformation. If $Y_t = g(t, X_t)$ denotes the transformed process, then

$$\begin{aligned} Y_{t+s} - y &= \frac{\partial}{\partial t}g(t,x)s + \frac{\partial}{\partial x}g(t,x)(X_{t+s} - x) + \frac{1}{2}\frac{\partial^2}{\partial t^2}g(t,x)s^2 \\ &\quad + \frac{\partial^2}{\partial t \partial x}g(t,x)s(X_{t+s} - x) + \frac{1}{2}\frac{\partial^2}{\partial x^2}g(t,x)(X_{t+s} - x)^2 \\ &\quad + O[(|X_{t+s} - x| + s)^3] \end{aligned}$$

for $X_t = x$ and $y = g(t,x)$. Taking conditional expectations and invoking equation (11.3) produce

$$\begin{aligned} \mathrm{E}(Y_{t+s} - Y_t \mid Y_t = y) &= \frac{\partial}{\partial t}g(t,x)s + \frac{\partial}{\partial x}g(t,x)\mu(t,x)s \\ &\quad + \frac{1}{2}\frac{\partial^2}{\partial x^2}g(t,x)\sigma^2(t,x)s + o(s). \end{aligned}$$

Similarly,

$$\mathrm{Var}(Y_{t+s} - Y_t \mid Y_t = y) \quad = \quad \left[\frac{\partial}{\partial x} g(t,x)\right]^2 \sigma^2(t,x)s + o(s).$$

It follows that the transformed diffusion process Y_t has infinitesimal mean and variance

$$\begin{aligned} \mu_Y(t,y) &= \frac{\partial}{\partial t} g(t,x) + \frac{\partial}{\partial x} g(t,x)\mu(t,x) + \frac{1}{2}\frac{\partial^2}{\partial x^2} g(t,x)\sigma^2(t,x) \\ \sigma_Y^2(t,y) &= \left[\frac{\partial}{\partial x} g(t,x)\right]^2 \sigma^2(t,x) \end{aligned} \tag{11.4}$$

at $y = g(t,x)$.

In many cases of interest, the random variable X_t has a density function $f(t,x)$ that depends on the initial point $X_0 = x_0$. To characterize $f(t,x)$, we now give a heuristic derivation of Kolmogorov's forward partial differential equation. Our approach exploits the notion of probability flux. Here it helps to imagine a large ensemble of diffusing particles, each independently executing the same process. We position ourselves at some point x and record the rate at which particles pass through x from left to right minus the rate at which they pass from right to left. This rate, normalized by the total number of particles, is the probability flux at x. We can express the flux more formally as the negative derivative $-\frac{\partial}{\partial t}\Pr(X_t \le x)$.

To calculate this time derivative, we rewrite the difference

$$\begin{aligned} & \Pr(X_t \le x) - \Pr(X_{t+s} \le x) \\ = \; & \Pr(X_t \le x, X_{t+s} > x) + \Pr(X_t \le x, X_{t+s} \le x) \\ & - \Pr(X_t \le x, X_{t+s} \le x) - \Pr(X_t > x, X_{t+s} \le x) \\ = \; & \Pr(X_t \le x, X_{t+s} > x) - \Pr(X_t > x, X_{t+s} \le x). \end{aligned}$$

The first of the resulting probabilities, $\Pr(X_t \le x, X_{t+s} > x)$, can be expressed as

$$\Pr(X_t \le x, X_{t+s} > x) \quad = \quad \int_0^\infty \int_{x-z}^{x} f(t,y)\phi_s(y,z)\, dy\, dz,$$

where the increment $Z = X_{t+s} - X_t$ has density $\phi_s(y,z)$ when $X_t = y$. The limits on the inner integral reflect the inequalities $x \ge y$ and $y + z > x$. In similar fashion, the second probability becomes

$$\Pr(X_t > x, X_{t+s} \le x) \quad = \quad \int_{-\infty}^{0} \int_{x}^{x-z} f(t,y)\phi_s(y,z)\, dy\, dz,$$

producing overall

$$\Pr(X_t \le x) - \Pr(X_{t+s} \le x) \quad = \quad \int_{-\infty}^{\infty} \int_{x-z}^{x} f(t,y)\phi_s(y,z)\, dy\, dz. \tag{11.5}$$

For small values of s, only values of y near x should contribute to the flux. Therefore, we can safely substitute the first-order expansion

$$f(t,y)\phi_s(y,z) \approx f(t,x)\phi_s(x,z) + \frac{\partial}{\partial x}\Big[f(t,x)\phi_s(x,z)\Big](y-x)$$

in equation (11.5). In light of equations (11.1) and (11.2), this yields

$$\begin{aligned}
& \Pr(X_t \le x) - \Pr(X_{t+s} \le x) \\
\approx & \int_{-\infty}^{\infty}\int_{x-z}^{x} \Big\{ f(t,x)\phi_s(x,z) + \frac{\partial}{\partial x}\Big[f(t,x)\phi_s(x,z)\Big](y-x)\Big\}\, dy\, dz \\
= & \int_{-\infty}^{\infty} \Big\{ z f(t,x)\phi_s(x,z) - \frac{z^2}{2}\frac{\partial}{\partial x}\Big[f(t,x)\phi_s(x,z)\Big]\Big\}\, dz \\
\approx & \mu(t,x)f(t,x)s - \frac{1}{2}\frac{\partial}{\partial x}\left[\int_{-\infty}^{\infty} z^2\phi_s(x,z)\, dz f(t,x)\right] \\
\approx & \mu(t,x)f(t,x)s - \frac{1}{2}\frac{\partial}{\partial x}\Big[\sigma^2(t,x)f(t,x)\Big]s.
\end{aligned}$$

Using equation (11.3), one can show that these approximations are good to order $o(s)$. Dividing by s and sending s to 0 give the flux

$$-\frac{\partial}{\partial t}\Pr(X_t \le x) = \mu(t,x)f(t,x) - \frac{1}{2}\frac{\partial}{\partial x}\Big[\sigma^2(t,x)f(t,x)\Big].$$

A final differentiation with respect to x now produces the Kolmogorov forward equation

$$\frac{\partial}{\partial t}f(t,x) = -\frac{\partial}{\partial x}\Big[\mu(t,x)f(t,x)\Big] + \frac{1}{2}\frac{\partial^2}{\partial x^2}\Big[\sigma^2(t,x)f(t,x)\Big]. \quad (11.6)$$

As t tends to 0, the density $f(t,x)$ concentrates all of its mass around the initial point x_0.

11.3 Examples Involving Brownian Motion

Example 11.3.1 *Standard Brownian Motion*

If $\mu(t,x) = 0$ and $\sigma^2(t,x) = 1$, then the forward equation (11.6) becomes

$$\frac{\partial}{\partial t}f(t,x) = \frac{1}{2}\frac{\partial^2}{\partial x^2}f(t,x).$$

For $X_0 = 0$ one can check the solution

$$f(t,x) = \frac{1}{\sqrt{2\pi t}}e^{-\frac{x^2}{2t}}$$

by straightforward differentiation. Thus, X_t has a Gaussian density with mean 0 and variance t. As t tends to 0, X_t becomes progressively more concentrated around its starting point 0. Because X_t and the increment $X_{t+s} - X_t$ are effectively independent for $s > 0$ small, and because the sum of independent Gaussian random variables is Gaussian, X_t and $X_{t+s} - X_t$ are Gaussian and independent for large s as well as for small s. Of course, rigorous proof of this fact is more subtle. In general, we cannot expect a diffusion process X_t to be normally distributed just because its short-time increments are approximately normal.

The independent increments property of standard Brownian motion facilitates calculation of covariances. For instance, writing

$$X_{t+s} \quad = \quad X_{t+s} - X_t + X_t$$

makes it clear that $\text{Cov}(X_{t+s}, X_t) = \text{Var}(X_t) = t$. We will use this formula in finding the infinitesimal mean and variance of the Brownian bridge diffusion process described in Example 11.3.4. The independent increments property also shows that a random vector $(X_{t_1}, \ldots, X_{t_n})$ with $t_1 < \cdots < t_n$ is multivariate normal. For example when $n = 3$, the representation

$$\begin{aligned} Y &= c_1 X_{t_1} + c_2 X_{t_2} + c_3 X_{t_3} \\ &= c_3(X_{t_3} - X_{t_2}) + (c_2 + c_3)(X_{t_2} - X_{t_1}) + (c_1 + c_2 + c_3) X_{t_1} \end{aligned}$$

demonstrates that the arbitrary linear combination Y is univariate normal.

Although the sample paths of Brownian motion are continuous, they are extremely rough and nowhere differentiable [106]. Nondifferentiability is plausible because the difference quotient at any time t satisfies

$$\text{Var}\left(\frac{X_{t+s} - X_t}{s}\right) \quad = \quad \frac{1}{s},$$

which tends to ∞ as s tends to 0. The lack of smoothness of Brownian paths makes stochastic integration such a subtle subject. ■

Example 11.3.2 *The Dirichlet Problem*

The Dirichlet problem arises in electrostatics, heat conduction, and other branches of physics. In two dimensions, it involves finding a function $u(x, y)$ that satisfies Laplace's equation

$$\frac{\partial^2}{\partial x^2} u(x, y) + \frac{\partial^2}{\partial y^2} u(x, y) \quad = \quad 0$$

on a bounded open domain Ω and that has prescribed values

$$u(x, y) \quad = \quad f(x, y)$$

on the boundary $\partial\Omega$ of Ω. The function $f(x, y)$ is assumed continuous. In the heat conduction setting, $u(x, y)$ represents steady-state temperature.

A solution to the Dirichlet problem is said to be a harmonic function. Harmonic functions are characterized by the averaging property

$$u(z) = \frac{1}{2\pi}\int_0^{2\pi} u[z + re^{i\theta}]\,d\theta, \tag{11.7}$$

where $z = (x, y)$ and r is the radius of the circle centered at z. This property must hold for all circles contained in Ω.

Although it took a long time for scientists to make the connection with Brownian motion, the basic idea is fairly simple. Two independent copies X_t and Y_t of standard Brownian motion define a curve $Z_t = (X_t, Y_t)$ in the plane. This bivariate Brownian process is isotropic, that is, exhibits no preferred direction of motion. At each point z in $\Omega \cup \partial\Omega$, we define a stopping time T_z measuring how long it takes the process Z_t to reach $\partial\Omega$ starting from z. The solution to Dirichlet's problem can then be written as the expected value

$$u(z) = \mathrm{E}[f(Z_{T_z})]$$

of $f(w)$ at the exit point. For a point z on $\partial\Omega$, $T_z = 0$ and $u(z) = f(z)$. For z in Ω, the averaging property (11.7) is intuitively clear once we condition on the first point reached on the boundary of the circle. The books [118, 160, 189] tell the whole story, including necessary qualifications on the nature of Ω and extensions to more than two dimensions. ■

Example 11.3.3 *Transformations of Standard Brownian Motion*

The transformed Brownian process $Y_t = \sigma X_t + \alpha t + x_0$ has infinitesimal mean and variance $\mu_Y(t, x) = \alpha$ and $\sigma_Y^2(t, x) = \sigma^2$. It is clear that Y_t is normally distributed with mean $\alpha t + x_0$ and variance $\sigma^2 t$. The further transformation $Z_t = e^{Y_t}$ leads to a process with infinitesimal mean and variance $\mu_Z(t, z) = z\alpha + \frac{1}{2}z\sigma^2$ and $\sigma_Z^2(t, z) = z^2\sigma^2$. Because Y_t is normally distributed, Z_t is lognormally distributed. ■

Example 11.3.4 *Brownian Bridge*

To construct the Brownian bridge Y_t from standard Brownian motion X_t, we restrict t to the interval $[0, 1]$ and condition on the event $X_1 = 0$. This ties down X_t at the two time points 0 and 1. The Brownian bridge diffusion process is important in evaluating the asymptotic distribution of the Kolmogorov-Smirnov statistic in nonparametric statistics [23].

To find the infinitesimal mean and variance of the Brownian bridge, we note that the vector (X_t, X_1, X_{t+s}) follows a multivariate normal distribution with mean vector $\mathbf{0} = (0, 0, 0)$ and covariance matrix

$$\begin{pmatrix} t & t & t \\ t & 1 & t+s \\ t & t+s & t+s \end{pmatrix}.$$

If Y_t equals X_t conditional on the event $X_1 = 0$, then Y_{t+s} conditional on the event $Y_t = y$ is just X_{t+s} conditional on the joint event $X_t = y$ and $X_1 = 0$. It follows from the conditioning formulas in Section 1.8 that Y_{t+s} given $Y_t = y$ is normally distributed with mean and variance

$$\begin{aligned}
\mathrm{E}(Y_{t+s} \mid Y_t = y) &= (t, t+s) \begin{pmatrix} t & t \\ t & 1 \end{pmatrix}^{-1} \begin{pmatrix} y-0 \\ 0-0 \end{pmatrix} + 0 \\
\mathrm{Var}(Y_{t+s} \mid Y_t = y) &= t + s - (t, t+s) \begin{pmatrix} t & t \\ t & 1 \end{pmatrix}^{-1} \begin{pmatrix} t \\ t+s \end{pmatrix}.
\end{aligned}$$

In view of the matrix inverse

$$\begin{pmatrix} t & t \\ t & 1 \end{pmatrix}^{-1} = \frac{1}{t(1-t)} \begin{pmatrix} 1 & -t \\ -t & t \end{pmatrix},$$

straightforward algebra demonstrates that

$$\begin{aligned}
\mathrm{E}(Y_{t+s} \mid Y_t = y) &= y - \frac{ys}{1-t} \\
\mathrm{Var}(Y_{t+s} \mid Y_t = y) &= s - \frac{s^2}{1-t}. \qquad (11.8)
\end{aligned}$$

It follows that the Brownian bridge has infinitesimal mean and variance $\mu(t, y) = -y/(1-t)$ and $\sigma^2(t, y) = 1$. ■

Example 11.3.5 *Bessel Process*

Consider a random process $X_t = (X_{1t}, \ldots, X_{nt})$ in R^n whose n components are independent standard Brownian motions. Let

$$Y_t = \sum_{i=1}^{n} X_{it}^2$$

be the squared distance from the origin. To calculate the infinitesimal mean and variance of this diffusion process, we write

$$Y_{t+s} - Y_t = \sum_{i=1}^{n} \left[2X_{it}(X_{i,t+s} - X_{it}) + (X_{i,t+s} - X_{it})^2 \right].$$

It follows from this representation, independence, and equation (11.3) that

$$\begin{aligned}
\mathrm{E}(Y_{t+s} - Y_t \mid X_t = x) &= ns \\
\mathrm{E}[(Y_{t+s} - Y_t)^2 \mid X_t = x] &= \sum_{i=1}^{n} 4x_i^2 \, \mathrm{E}[(X_{i,t+s} - X_{it})^2 \mid X_t = x] + o(s) \\
&= 4ys + o(s)
\end{aligned}$$

for $y = \sum_{i=1}^n x_i^2$. Because

$$\begin{aligned} \mathrm{E}(Y_{t+s} - Y_t \mid Y_t = y) &= \mathrm{E}[\mathrm{E}(Y_{t+s} - Y_t \mid X_t = x) \mid Y_t = y] \\ \mathrm{Var}(Y_{t+s} - Y_t \mid Y_t = y) &= \mathrm{E}[(Y_{t+s} - Y_t)^2 \mid Y_t = y] + o(s) \\ &= \mathrm{E}\{\mathrm{E}[(Y_{t+s} - Y_t)^2 \mid X_t = x] \mid Y_t = y\} + o(s), \end{aligned}$$

we conclude that Y_t has infinitesimal mean and variance

$$\begin{aligned} \mu_Y(t,y) &= n \\ \sigma_Y^2(t,y) &= 4y. \end{aligned}$$

The random distance $R_t = \sqrt{Y_t}$ is known as the Bessel process. Its infinitesimal mean and variance

$$\begin{aligned} \mu_R(t,r) &= \frac{n-1}{2r} \\ \sigma_R^2(t,r) &= 1 \end{aligned}$$

are immediate consequences of formula (11.4). ■

11.4 Other Examples of Diffusion Processes

Example 11.4.1 *Diffusion Approximation to Kendall's Process*

Suppose in Kendall's model that the birth, death, and immigration rates α, δ, and ν are constant. (Here we have substituted δ for μ to avoid a collision with the symbol for the infinitesimal mean.) Given x particles at time t, there is one birth with probability $\alpha x s$ during the very short time interval $(t, t+s)$, one death with probability $\delta x s$, and one immigrant with probability νs. The probability of more than one event is $o(s)$. These considerations imply that

$$\begin{aligned} \mu(t,x) &= (\alpha - \delta)x + \nu \\ \sigma^2(t,x) &= (\alpha + \delta)x + \nu. \end{aligned}$$

This diffusion approximation is apt to be good for a moderate number of particles and poor for a small number of particles. ■

Example 11.4.2 *Neuron Firing and the Ornstein-Uhlenbeck Process*

Physiologists have long been interested in understanding how neurons fire [108, 199]. Firing involves the electrical potentials across the cell membranes of these basic cells of the nervous system. The typical neuron is composed of a compact cell body or soma into which thousands of small dendrites feed incoming signals. The soma integrates these signals and occasionally fires.

When the neuron fires, a pulse, or action potential, is sent down the axon connected to the soma. The axon makes contact with other nerve cells or with muscle cells through junctions known as synapses. An action potential is a transient electrical depolarization of the cell membrane that propagates from the soma along the axon to the synapses. In a neuron's resting state, there is a potential difference of about -70 mV (millivolts) across the soma membrane. This is measured by inserting a microelectrode through the membrane. A cell is said to be excited, or depolarized, if the soma potential exceeds the resting potential; it is inhibited, or hyperpolarized, in the reverse case. When the soma potential reaches a threshold of from 10 to 40 mV above the resting potential, the neuron fires. After the axon fires, the potential is reset to a level below the resting potential. An all or nothing action potential converts an analog potential difference into a digital pulse of information.

To model the firing of a neuron, we let X_t be the soma membrane potential at the time t. Two independent processes, one excitatory and one inhibitory, drive X_t. The soma receives excitatory pulses of magnitude ϵ according to a Poisson process with intensity α and inhibitory pulses of magnitude $-\delta$ according to an independent Poisson process with intensity β. The potential is subject to exponential decay with time constant γ to the resting value x_r. When X_t reaches the fixed threshold s, the neuron fires. Afterwards, X_t is reset to the level x_0.

Because typical values of ϵ and δ range from 0.5 to 1 mV, a diffusion approximation is appropriate. In view of the Poisson nature of the inputs, we have

$$\begin{aligned} \mu(t,x) &= \alpha\epsilon - \beta\delta - \gamma(x - x_r) \\ \sigma^2(t,x) &= \alpha\epsilon^2 + \beta\delta^2. \end{aligned}$$

The transformed process $Y_t = X_t - \eta/\gamma$ with $\eta = \alpha\epsilon - \beta\delta + \gamma x_r$ is known as the Ornstein-Uhlenbeck process. It has infinitesimal mean and variance

$$\begin{aligned} \mu_Y(t,y) &= -\gamma\Big(y + \frac{\eta}{\gamma}\Big) + \eta \\ &= -\gamma y \\ \sigma_Y^2(t,y) &= \alpha\epsilon^2 + \beta\delta^2 \\ &= \sigma^2 \end{aligned}$$

and is somewhat easier to study.

Fortunately, it is possible to relate the Ornstein-Uhlenbeck process to Brownian motion. Consider the complicated transformation

$$Y_t = \nu e^{-\gamma t}\left(W_{e^{\zeta t}} - W_1 + \nu^{-1} y_0\right)$$

of standard Brownian motion W_t. If we write

$$\begin{aligned} Y_{t+s} - Y_t &= \nu e^{-\gamma(t+s)}\left[W_{e^{\zeta(t+s)}} - W_{e^{\zeta t}}\right] \\ &\quad + \left(e^{-\gamma s} - 1\right)\nu e^{-\gamma t}\left(W_{e^{\zeta t}} - W_1 + \nu^{-1} y_0\right), \end{aligned}$$

then it is clear that this increment is Gaussian. Furthermore, setting $\zeta = 2\gamma$ gives

$$\begin{aligned}
\mathrm{E}\left[Y_{t+s} - Y_t \mid Y_t = y\right] &= \left(e^{-\gamma s} - 1\right) y \\
&= -\gamma y s + o(s) \\
\mathrm{Var}\left[Y_{t+s} - Y_t \mid Y_t = y\right] &= \nu^2 e^{-2\gamma(t+s)} \left[e^{\zeta(t+s)} - e^{\zeta t}\right] \\
&= \nu^2 \left[1 - e^{-2\gamma s}\right] \\
&= 2\nu^2 \gamma s + o(s).
\end{aligned}$$

Therefore, if we define ν so that $\sigma^2 = 2\nu^2\gamma$, then the infinitesimal mean $-\gamma y$ and variance σ^2 of Y_t coincide with those of the Ornstein-Uhlenbeck process. One interesting dividend of the definition of Y_t and the assumption $\zeta = 2\gamma$ is that we can immediately conclude that Y_t is Gaussian with mean and variance

$$\begin{aligned}
\mathrm{E}(Y_t) &= y_0 e^{-\gamma t} \\
\mathrm{Var}(Y_t) &= \nu^2 e^{-2\gamma t} \left(e^{2\gamma t} - 1\right) \\
&= \frac{\sigma^2}{2\gamma} \left(1 - e^{-2\gamma t}\right).
\end{aligned} \tag{11.9}$$

These are clearly compatible with the initial value $Y_0 = y_0$. ■

Example 11.4.3 *Wright-Fisher Model with Mutation and Selection*

The Wright-Fisher model for the evolution of a deleterious or neutral gene postulates (a) discrete generations, (b) finite population size, (c) no immigration, and (d) random sampling from a gamete pool. In assumption (d), each current population member contributes to the infinite pool of potential gametes (sperm and eggs) in proportion to his or her fitness.

Mutation from the normal allele A_2 to the deleterious allele A_1 takes place at this stage with mutation rate η; back mutation is not permitted. Once the pool of potential gametes is formed, actual gametes are sampled randomly. In the neutral model introduced in Examples 7.3.2, 10.2.6, and 10.4.2, we neglect mutation and selection and treat the two alleles symmetrically.

The population frequencies (proportions) p and q of the two alleles A_1 and A_2 are the primary focus of the Wright-Fisher model [44, 56, 107]. These frequencies change over time in response to the forces of mutation, selection, and genetic drift (random sampling of gametes). Selection operates through fitness differences. Denote the average fitnesses of the genotypes A_1/A_1, A_1/A_2, and A_2/A_2 by w_{A_1/A_1}, w_{A_1/A_2}, and w_{A_2/A_2}, respectively. Because only relative fitness is important in formulating the dynamics of the Wright-Fisher model, for a recessive disease we set $w_{A_1/A_1} = f < 1$

and $w_{A_1/A_2} = w_{A_2/A_2} = 1$. Similarly for a dominant disease, we set $w_{A_1/A_1} = w_{A_1/A_2} = f < 1$ and $w_{A_2/A_2} = 1$.

In a purely deterministic model, the frequency $p_n = 1 - q_n$ of the disease allele A_1 for a dominant disease satisfies the recurrence

$$p_{n+1} = \frac{fp_n^2 + (1+\eta)fp_nq_n + \eta q_n^2}{fp_n^2 + f2p_nq_n + q_n^2}. \tag{11.10}$$

Here individuals bearing the three genotypes contribute gametes in proportion to the product of their population frequencies and relative fitnesses. Because gametes are drawn independently from the pool at generation n, the three genotypes A_1/A_1, A_1/A_2, and A_2/A_2 occur in the Hardy-Weinberg proportions p_{n+1}^2, $2p_{n+1}q_{n+1}$, and q_{n+1}^2, respectively, at generation $n+1$.

Given that we expect p_n to be of order η, equation (11.10) radically simplifies if we expand and drop all terms of order η^2 and higher. The resulting linear recurrence

$$p_{n+1} = \eta + fp_n \tag{11.11}$$

can be motivated by arguing that A_1 genes at generation $n+1$ arise either through mutation (probability ν) or by descent from an existing A_1 gene (probability fp_n under reduced fitness). The recurrence (11.11) has fixed point $p_\infty = \frac{\eta}{1-f}$. Furthermore, because $p_{n+1} - p_\infty = f(p_n - p_\infty)$, the iterates p_n converge to p_∞ at linear rate f.

In contrast, the frequency p_n of the disease allele A_1 for a recessive disease satisfies the deterministic recurrence

$$p_{n+1} = \frac{fp_n^2 + (1+\eta)p_nq_n + \eta q_n^2}{fp_n^2 + 2p_nq_n + q_n^2}.$$

Now we expect p_n to be of order $\sqrt{\eta}$. Expanding the recurrence and dropping all terms of order $\eta^{3/2}$ and higher yields the quadratic recurrence

$$p_{n+1} = \eta + p_nq_n + fp_n^2. \tag{11.12}$$

In this case, A_1 genes arise from mutation (probability ν), transmission by a normal heterozygote (probability $\frac{1}{2}2p_nq_n$), or transmission from an affected homozygote (probability fp_n^2). The recurrence (11.12) has fixed point $p_\infty = \sqrt{\eta/(1-f)}$. To determine the rate of convergence, we rewrite equation (11.12) as

$$p_{n+1} - p_\infty = [1 - (1-f)(p_n + p_\infty)](p_n - p_\infty).$$

This makes it clear that p_n converges to p_∞ at linear rate

$$1 - (1-f)2p_\infty = 1 - 2\sqrt{\eta(1-f)}.$$

These two special cases both entail a disease prevalence on the order of η, which typically falls in the range 10^{-7} to 10^{-5}. The two cases differ

markedly in their rates of convergence to equilibrium. A dominant disease reaches equilibrium quickly unless the average fitness f of affecteds is very close to 1. By comparison, a recessive disease reaches equilibrium extremely slowly.

Random fluctuations in the frequency of the disease allele can be large in small populations. Let N_n be the size of the surrounding population at generation n. In some cases we will take N_n constant to simplify our mathematical development. In the stochastic theory, the deterministic frequency p_n of allele A_1 at generation n is replaced by the random frequency X_n of A_1. This frequency is the ratio of the total number Y_n of A_1 alleles present to the total number of genes $2N_n$. The Wright-Fisher model specifies that Y_n is binomially distributed with $2N_n$ trials and success probability $p(X_{n-1})$ determined by the proportion $p(X_{n-1})$ of A_1 alleles in the pool of potential gametes for generation n. In passing to a diffusion approximation, we take one generation as the unit of time and substitute

$$\begin{aligned} \mu(n, x_n) &= \mathrm{E}(X_{n+1} - X_n \mid X_n = x_n) \qquad (11.13) \\ &= p(x_n) - x_n \\ \sigma^2(n, x_n) &= \mathrm{Var}(X_{n+1} - X_n \mid X_n = x_n) \qquad (11.14) \\ &= \frac{p(x_n)[1 - p(x_n)]}{2N_{n+1}} \end{aligned}$$

for the infinitesimal mean $\mu(t, x)$ and variance $\sigma^2(t, x)$ of the diffusion process evaluated at time $t = n$ and position $x = x_n$.

Under neutral evolution, the gamete pool probability $p(x) = x$. This formula for $p(x)$ entails no systematic tendency for either allele to expand at the expense of the other allele. For a dominant disease, $p(x) = \eta + fx$, while for a recessive disease, $p(x) = \eta + x - (1 - f)x^2$. Most population geneticists substitute $p(x) = x$ in formula (11.14) defining the infinitesimal variance $\sigma^2(t, x)$. This action is justified for neutral and recessive inheritance, but less so for dominant inheritance, where the allele frequency x is typically on the order of magnitude of the mutation rate η. It is also fair to point out that in the presence of inbreeding or incomplete mixing of a population, the effective population size is less than the actual population size [44]. For the sake of simplicity, we will ignore this evolutionary fact. ■

11.5 Process Moments

Taking unconditional expectations in expression (11.1) gives

$$\mathrm{E}(X_{t+s}) = \mathrm{E}(X_t) + \mathrm{E}[\mu(t, X_t)]s + o(s).$$

Forming the obvious difference quotient and sending s to 0 therefore provide the ordinary differential equation

$$\frac{d}{dt}\,\mathrm{E}(X_t) \;=\; \mathrm{E}[\mu(t, X_t)] \tag{11.15}$$

characterizing $\mathrm{E}(X_t)$. In the special case $\mu(t,x) = \alpha x + \beta$, it is easy to check that equation (11.15) has solution

$$\mathrm{E}(X_t) \;=\; \mathrm{E}(X_0)e^{\alpha t} + \frac{\beta}{\alpha}\left(e^{\alpha t} - 1\right) \tag{11.16}$$

unless $\alpha = 0$, in which case $\mathrm{E}(X_t) = \mathrm{E}(X_0) + \beta t$.

Taking unconditional variances in expression (11.2) yields in a similar manner

$$\begin{aligned}
\mathrm{Var}(X_{t+s}) &= \mathrm{E}[\mathrm{Var}(X_t + \Delta X_t \mid X_t)] + \mathrm{Var}[\mathrm{E}(X_t + \Delta X_t \mid X_t)] \\
&= \mathrm{E}[\sigma^2(t, X_t)s + o(s)] + \mathrm{Var}[X_t + \mu(t, X_t)s + o(s)] \\
&= \mathrm{E}[\sigma^2(t, X_t)]s + \mathrm{Var}(X_t) + 2\,\mathrm{Cov}[X_t, \mu(t, X_t)]s + o(s)
\end{aligned}$$

for $\Delta X_t = X_{t+s} - X_t$. In this case taking the difference quotient and sending s to 0 give the ordinary differential equation

$$\frac{d}{dt}\,\mathrm{Var}(X_t) \;=\; \mathrm{E}[\sigma^2(t, X_t)] + 2\,\mathrm{Cov}[X_t, \mu(t, X_t)] \tag{11.17}$$

rigorously derived in reference [58].

Example 11.5.1 *Moments of the Wright-Fisher Diffusion Process*

In the diffusion approximation to the Wright-Fisher model for a dominant disease, equation (11.13) implies

$$\mu(t,x) \;=\; \eta - (1-f)x.$$

It follows from equation (11.16) that

$$\mathrm{E}(X_t) \;=\; \left[x_0 - \frac{\eta}{1-f}\right] e^{-(1-f)t} + \frac{\eta}{1-f}$$

for $X_0 = x_0$. The limiting value of $\eta/(1-f)$ is the same as the deterministic equilibrium. In the case of neutral evolution with $f = 1$ and $\eta = 0$, the mean $\mathrm{E}(X_t) = x_0$ is constant. With constant population size N, equations (11.14) and (11.17) therefore yield the differential equation

$$\begin{aligned}
\frac{d}{dt}\,\mathrm{Var}(X_t) &= \frac{\mathrm{E}(X_t) - \mathrm{E}(X_t^2)}{2N} \\
&= \frac{x_0 - x_0^2 - \mathrm{Var}(X_t)}{2N},
\end{aligned}$$

with solution

$$\mathrm{Var}(X_t) \;=\; x_0(1-x_0)\left[1-e^{-\frac{t}{2N}}\right].$$

This expression for $\mathrm{Var}(X_t)$ tends to $x_0(1-x_0)$ as t tends to ∞, which is the variance of the limiting random variable

$$X_\infty \;=\; \begin{cases} 1 & \text{with probability } x_0 \\ 0 & \text{with probability } 1-x_0\,. \end{cases}$$

Fan and Lange [57] calculate $\mathrm{Var}(X_t)$ for the dominant case. In the recessive case, this approach to $\mathrm{E}(X_t)$ and $\mathrm{Var}(X_t)$ breaks down because $\mu(t,x)$ is quadratic rather than linear in x. ■

Example 11.5.2 *Moments of the Brownian Bridge*

For the Brownian bridge, equation (11.15) becomes

$$\frac{d}{dt}\,\mathrm{E}(X_t) \;=\; -\frac{1}{1-t}\,\mathrm{E}(X_t).$$

The unique solution with $\mathrm{E}(X_0)=0$ is $\mathrm{E}(X_t)=0$ for all t. Equation (11.17) becomes

$$\frac{d}{dt}\,\mathrm{Var}(X_t) \;=\; 1-\frac{2}{1-t}\,\mathrm{Var}(X_t).$$

This has solution

$$\mathrm{Var}(X_t) \;=\; t(1-t)$$

subject to the initial value $\mathrm{Var}(X_0)=0$. These results match the results in formula (11.8) if we replace y by 0, s by t, and t by 0. ■

11.6 First Passage Problems

Let $c<d$ be two points in the interior of the range I of a diffusion process X_t. Define T_c to be the first time t that $X_t=c$ starting from $X_0\geq c$. If eventually $X_t<c$, then the continuity of the sample paths guarantees that $X_t=c$ at some first time T_c. It may be that $T_c=\infty$ with positive probability. Similar considerations apply to T_d, the first time t that $X_t=d$ starting from $X_0\leq d$. The process X_t exits (c,d) at the time $T=\min\{T_c,T_d\}$. We consider two related problems involving these first passage times. One problem is to calculate the probability $u(x)=\Pr(T_d<T_c\mid X_0=x)$ that the process exits via d starting from $x\in[c,d]$. It is straightforward to derive a differential equation determining $u(x)$ given the boundary conditions

$u(c) = 0$ and $u(d) = 1$. With this end in mind, we assume that X_t is time homogeneous.

For $s > 0$ small and $x \in (c, d)$, the probability that X_t reaches either c or d during the time interval $[0, s]$ is $o(s)$. Thus,

$$u(x) = \mathrm{E}[u(X_s) \mid X_0 = x] + o(s).$$

If we let $\Delta X_s = X_s - X_0$ and expand $u(X_s)$ in a second-order Taylor series, then we find that

$$\begin{aligned} u(X_s) &= u(x + \Delta X_s) \\ &= u(x) + u'(x)\Delta X_s + \frac{1}{2}\Big[u''(x) + r(\Delta X_s)\Big]\Delta X_s^2, \quad (11.18) \end{aligned}$$

where the relative error $r(\Delta X_s)$ tends to 0 as ΔX_s tends to 0. Invoking equations (11.1), (11.2), and (11.18) therefore yields

$$\begin{aligned} u(x) &= \mathrm{E}[u(X_s)] + o(s) \\ &= u(x) + \mu(x)u'(x)s + \frac{1}{2}\sigma^2(x)u''(x)s + o(s), \end{aligned}$$

which, upon rearrangement and sending s to 0, gives the differential equation

$$0 = \mu(x)u'(x) + \frac{1}{2}\sigma^2(x)u''(x). \quad (11.19)$$

It is a simple matter to check that equation (11.19) can be solved explicitly by defining

$$v(x) = \int_l^x e^{-\int_l^y \frac{2\mu(z)}{\sigma^2(z)}dz} dy$$

and setting

$$u(x) = \frac{v(x) - v(c)}{v(d) - v(c)}. \quad (11.20)$$

Here the lower limit of integration l can be any point in the interval $[c, d]$. This particular solution also satisfies the boundary conditions.

Example 11.6.1 *Exit Probabilities in the Wright-Fisher Model*

In the diffusion approximation to the neutral Wright-Fisher model with constant population size N, we calculate

$$v(x) = \int_l^x e^{-\int_l^y 0\, dz} dy = x - l.$$

Thus, starting at a frequency of x for allele A_1, allele A_2 goes extinct before allele A_1 with probability

$$u(x) = \lim_{c \to 0,\, d \to 1} \frac{x - l - (c - l)}{d - l - (c - l)} = x.$$

This example is typical in the sense that $u(x) = (x - c)/(d - c)$ for any diffusion process with $\mu(x) = 0$. ■

Example 11.6.2 *Exit Probabilities in the Bessel Process*

Consider two fixed radii $0 < r_0 < r_1$ in the Bessel process. It is straightforward to calculate

$$\begin{aligned} v(x) &= \int_l^x e^{-\int_l^y \frac{n-1}{z}\, dz} dy \\ &= \int_l^x \frac{l^{n-1}}{y^{n-1}}\, dy \\ &= -\frac{l^{n-1}}{(n-2)x^{n-2}} + \frac{l^{n-1}}{(n-2)l^{n-2}} \end{aligned}$$

for $n > 2$. This yields

$$u(x) = \frac{r_0^{-n+2} - x^{-n+2}}{r_0^{-n+2} - r_1^{-n+2}},$$

from which it follows that $u(x)$ tends to 1 as r_0 tends to 0. This fact is intuitively obvious because the Brownian path X_t is much more likely to hit the surface of a large outer sphere of radius r_1 before it hits the surface of a small inner sphere of radius r_0. In R^2 we have

$$u(x) = \frac{\ln x - \ln r_0}{\ln r_1 - \ln r_0},$$

and the same qualitative comments apply. ■

Another important problem is to calculate the expectation

$$w(x) = \mathrm{E}[g(T) \mid X_0 = x]$$

of a function of the exit time T from $[c, d]$. For instance, $g(t) = t^n$ gives the nth moment of T, and $g(t) = e^{-\theta t}$ gives the Laplace transform of T. We again derive an ordinary differential equation determining $w(x)$, but now the pertinent boundary conditions are $w(c) = w(d) = g(0)$. To emphasize the dependence of T on the initial position x, let us write T_x in place of T.

We commence our derivation with the expansion

$$w(x) = \mathrm{E}[g(T_x) \mid X_0 = x]$$

$$\begin{aligned}
&= \mathrm{E}[g(T_{X_s} + s) \mid X_0 = x] + o(s) \\
&= \mathrm{E}[g(T_{X_s}) + g'(T_{X_s})s \mid X_0 = x] + o(s) \\
&= \mathrm{E}\{\mathrm{E}[g(T_{X_s}) \mid X_s] \mid X_0 = x\} + \mathrm{E}[g'(T_{X_s}) \mid X_0 = x]s + o(s) \\
&= \mathrm{E}[w(X_s) \mid X_0 = x] + \mathrm{E}[g'(T_x) \mid X_0 = x]s + o(s).
\end{aligned}$$

Employing the same reasoning used in deriving the differential equation (11.19) for $u(x)$, we deduce that

$$\mathrm{E}[w(X_s) \mid X_0 = x] = w(x) + \mu(x)w'(x)s + \frac{1}{2}\sigma^2(x)w''(x)s + o(s).$$

It follows that

$$\begin{aligned}
w(x) &= w(x) + \mu(x)w'(x)s + \frac{1}{2}\sigma^2(x)w''(x)s \\
&\quad + \mathrm{E}[g'(T_x) \mid X_0 = x]s + o(s).
\end{aligned}$$

Rearranging this and sending s to 0 produce the differential equation

$$0 = \mu(x)w'(x) + \frac{1}{2}\sigma^2(x)w''(x) + \mathrm{E}[g'(T_x) \mid X_0 = x].$$

The special cases $g(t) = t$ and $g(t) = e^{-\theta t}$ correspond to the differential equations

$$0 = \mu(x)w'(x) + \frac{1}{2}\sigma^2(x)w''(x) + 1 \tag{11.21}$$

$$0 = \mu(x)w'(x) + \frac{1}{2}\sigma^2(x)w''(x) - \theta w(x), \tag{11.22}$$

respectively. The pertinent boundary conditions are $w(c) = w(d) = 0$ and $w(c) = w(d) = 1$.

Example 11.6.3 *Mean Exit Times in the Wright-Fisher Model*

In the diffusion approximation to the neutral Wright-Fisher model with constant population size N, equation (11.21) becomes

$$0 = \frac{x(1-x)}{4N}w''(x) + 1. \tag{11.23}$$

If we take $c = 0$ and $d = 1$, then $w(x)$ represents the expected time until fixation of one of the two alleles. To solve equation (11.23), observe that

$$\begin{aligned}
w'(x) &= -4N \int_{\frac{1}{2}}^{x} \frac{1}{y(1-y)}\, dy + k_1 \\
&= -4N \int_{\frac{1}{2}}^{x} \left[\frac{1}{y} + \frac{1}{(1-y)}\right] dy + k_1 \\
&= -4N \left[\ln x - \ln(1-x)\right] + k_1
\end{aligned}$$

for some constant k_1. Integrating again yields

$$\begin{aligned} w(x) &= -4N \int_{\frac{1}{2}}^{x} [\ln y - \ln(1-y)]\, dy + k_1 x + k_2 \\ &= -4N\,[x \ln x + (1-x)\ln(1-x)] + k_1 x + k_2 \end{aligned}$$

for some constant k_2. The boundary condition $w(0) = 0$ implies $k_2 = 0$, and the boundary condition $w(1) = 0$ implies $k_1 = 0$. It follows that

$$w(x) = -4N\,[x \ln x + (1-x)\ln(1-x)]\,.$$

This is proportional to N and attains a maximum of $4N \ln 2$ at $x = 1/2$. ■

Example 11.6.4 *Mean Exit Times in the Bessel Process*

Again fix two radii $0 < r_0 < r_1$. If we let $v(x) = w'(x)$, then equation (11.21) becomes

$$0 = \frac{n-1}{2x} v(x) + \frac{1}{2} v'(x) + 1$$

for the Bessel process. This inhomogeneous differential equation has the particular solution $v(x) = -2x/n$. The general solution is the sum of the particular solution and an arbitrary multiple of a solution to the homogeneous differential equation

$$0 = \frac{n-1}{2x} v(x) + \frac{1}{2} v'(x).$$

This makes it clear that the general solution is

$$v(x) = -\frac{2x}{n} + a x^{-n+1}$$

for an arbitrary constant a. Integrating the general solution for $v(x)$ produces the general solution

$$w(x) = -\frac{x^2}{n} - \frac{a x^{-n+2}}{n-2} + b$$

for $w(x)$ when $n > 2$.

The arbitrary constants a and b are determined by the boundary conditions $w(r_0) = w(r_1) = 0$. Thus,

$$\begin{aligned} 0 &= w(r_1) - w(r_0) \\ &= \frac{r_0^2 - r_1^2}{n} + \frac{a(r_0^{-n+2} - r_1^{-n+2})}{n-2} \end{aligned}$$

gives

$$a = \frac{(n-2)(r_0^2 - r_1^2)}{n(r_1^{-n+2} - r_0^{-n+2})},$$

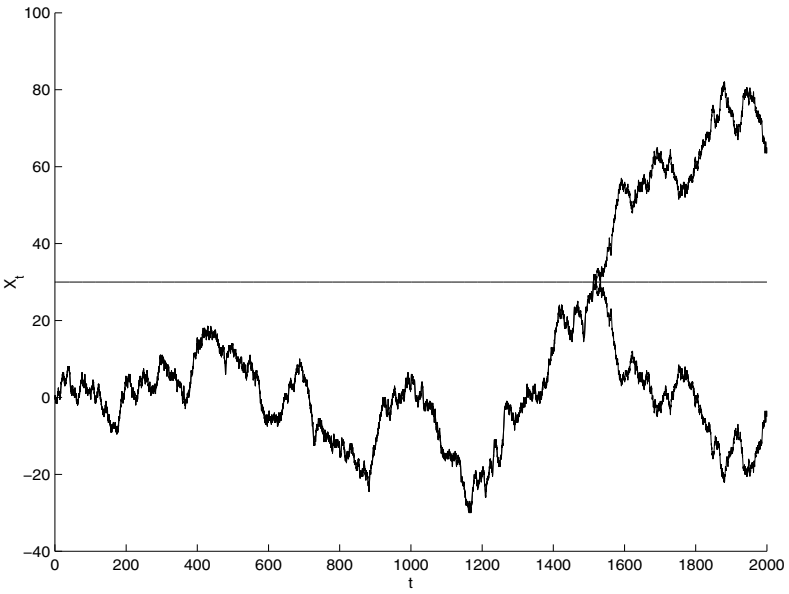


FIGURE 11.1. Reflection of X_t around the Level $b = 30$.

which tends to 0 as r_0 tends to 0. For x fixed and $r_0 = 0$, it follows that

$$b = \frac{r_1^2}{n}.$$

This gives the expected time

$$w(x) = \frac{r_1^2 - x^2}{n} \tag{11.24}$$

to reach the surface of the sphere of radius r_1 starting from the surface of the sphere of radius x. This formula for $w(x)$ also holds in R^2, but the derivation is slightly different. ■

11.7 The Reflection Principle

For standard Brownian motion, we can find the entire distribution of the first passage time T_b to level $b > 0$ when no lower exit level comes into play. The reflection principle operates by reflecting part of a Brownian path crossing the horizontal line of height b. Figure 11.1 depicts a Brownian path from $t = 0$ to $t = 2000$ crossing the level $b = 30$ at about $t = 1532$. After the crossing, the path and its reflection diverge. The principle establishes a one-to-one correspondence between paths that exceed b at time t and paths that exceed b at some time before t but fall below b at time t. The event

$\{T_b > t\}$ is the union of these two kinds of events. By symmetry a reflected path is as likely to be taken as an original path. Hence,

$$\begin{aligned}
\Pr(T_b < t) &= \Pr(X_t > b) + \Pr(T_b < t, X_t \le b) \\
&= 2\Pr(X_t > b) \\
&= \frac{2}{\sqrt{2\pi t}} \int_b^\infty e^{-\frac{x^2}{2t}}\, dx \\
&= \sqrt{\frac{2}{\pi}} \int_{b/\sqrt{t}}^\infty e^{-\frac{y^2}{2}}\, dy.
\end{aligned}$$

Sending t to ∞ shows that $\Pr(T_b < \infty) = 1$. However, the mean of T_b is infinite as verified by the calculation

$$\begin{aligned}
\mathrm{E}(T_b) &= \int_0^\infty \Pr(T_b > t)\, dt \\
&= \int_0^\infty \sqrt{\frac{2}{\pi}} \int_0^{b/\sqrt{t}} e^{-\frac{y^2}{2}}\, dy\, dt \\
&= \sqrt{\frac{2}{\pi}} \int_0^\infty \int_o^{b^2/y^2} dt\, e^{-\frac{y^2}{2}}\, dy \\
&= \sqrt{\frac{2}{\pi}} \int_0^\infty \frac{b^2}{y^2} e^{-\frac{y^2}{2}}\, dy \\
&= \infty.
\end{aligned}$$

It is possible to extend the reflection principle to the transformed Brownian process $Y_t = \sigma X_t + \alpha t$ with infinitesimal mean $\alpha > 0$ and infinitesimal variance σ^2 [203]. Although a path and its reflection are no longer equally likely, they occur in predictable ratios. Consider a path executed by Y_t that crosses the barrier b and terminates at the point $x > b$. If we partition the interval $[T_b, t]$ into the points $T_b = t_0 < t_1 < \cdots < t_n = t$, then reflection replaces every increment $\Delta y_i = Y_{t_{i+1}} - Y_{t_i}$ by its negative $-\Delta y_i$. Because the two increments are approximately normally distributed with mean $\alpha\Delta t_i$ and variance $\sigma^2\Delta t_i$ with $\Delta t_i = t_{i+1} - t_i$, their likelihood ratio is approximately

$$R_i = \frac{e^{-\frac{(-\Delta y_i - \alpha\Delta t_i)^2}{2\sigma^2\Delta t_i}}}{e^{-\frac{(\Delta y_i - \alpha\Delta t_i)^2}{2\sigma^2\Delta t_i}}} = e^{-\frac{2\alpha\Delta y_i}{\sigma^2}}.$$

In view of the independent increments property of Brownian motion, the likelihood ratio for the reflected path versus the original path is

$$R(x) = \prod_{i=0}^{n-1} R_i = e^{-\frac{2\alpha(x-b)}{\sigma^2}},$$

which does not depend on the particular partition.

As with standard Brownian motion, we have

$$\Pr(X_t > b) = \frac{1}{\sqrt{2\pi\sigma^2 t}} \int_b^\infty e^{-\frac{(x-\alpha t)^2}{2\sigma^2 t}} dx.$$

This is one of the two probabilities needed for the distribution function of the first passage time T_b. To find the other probability, we exploit the likelihood ratio $R(x)$ in the calculation.

$$\begin{aligned} \Pr(T_b < t, X_t \leq b) &= \frac{1}{\sqrt{2\pi\sigma^2 t}} \int_b^\infty R(x) e^{-\frac{(x-\alpha t)^2}{2\sigma^2 t}} dx \\ &= \frac{e^{\frac{2\alpha b}{\sigma^2}}}{\sqrt{2\pi\sigma^2 t}} \int_b^\infty e^{-\frac{(x+\alpha t)^2}{2\sigma^2 t}} dx. \end{aligned}$$

It follows that

$$\Pr(T_b < t) = \frac{1}{\sqrt{2\pi\sigma^2 t}} \int_b^\infty e^{-\frac{(x-\alpha t)^2}{2\sigma^2 t}} dx + \frac{e^{\frac{2\alpha b}{\sigma^2}}}{\sqrt{2\pi\sigma^2 t}} \int_b^\infty e^{-\frac{(x+\alpha t)^2}{2\sigma^2 t}} dx.$$

The first passage time T_b is said to have an inverse Gaussian distribution. Problem 10 supplies its mean and variance. Straightforward differentiation of the distribution function provides the density function. In addition to its obvious application to the neuron-firing model, the inverse Gaussian distribution appears in many statistical models involving waiting time data.

11.8 Equilibrium Distributions

In certain situations, a time-homogeneous diffusion process will tend to equilibrium. To find the equilibrium distribution, we set the left-hand side of Kolmogorov's equation (11.6) equal to 0 and solve for the equilibrium distribution $f(x) = \lim_{t\to\infty} f(t,x)$. Integrating the equation

$$0 = -\frac{d}{dx}\Big[\mu(x)f(x)\Big] + \frac{1}{2}\frac{d^2}{dx^2}\Big[\sigma^2(x)f(x)\Big] \tag{11.25}$$

once gives

$$k_1 = -\mu(x)f(x) + \frac{1}{2}\frac{d}{dx}\Big[\sigma^2(x)f(x)\Big]$$

for some constant k_1. The choice $k_1 = 0$ corresponds to the intuitively reasonable condition of no probability flux at equilibrium. Dividing the no flux equation by $\sigma^2(x)f(x)$ yields

$$\frac{d}{dx}\ln[\sigma^2(x)f(x)] = \frac{2\mu(x)}{\sigma^2(x)}.$$

If we now choose l in the interior of the range I of X_t and integrate a second time, then we deduce that

$$\ln[\sigma^2(x)f(x)] = k_2 + \int_l^x \frac{2\mu(y)}{\sigma^2(y)}\,dy,$$

from which Wright's formula

$$f(x) = \frac{k_3 e^{\int_l^x \frac{2\mu(y)}{\sigma^2(y)}\,dy}}{\sigma^2(x)} \tag{11.26}$$

for the equilibrium distribution emerges. An appropriate choice of the constant $k_3 = e^{k_2}$ serves to make $\int_I f(x)\,dx = 1$ when the equilibrium distribution exists and is unique.

Example 11.8.1 *Equilibrium for the Ornstein-Uhlenbeck Process*

Wright's formula (11.26) gives

$$\begin{aligned} f(y) &= \frac{k_3 e^{-\int_0^y \frac{2\gamma z}{\sigma^2}\,dz}}{\sigma^2} \\ &= \sqrt{\frac{\gamma}{\pi\sigma^2}}e^{-\gamma y^2/\sigma^2} \end{aligned}$$

for the Ornstein-Uhlenbeck process. This is exactly the normal density one would predict by sending t to ∞ in the moment equations (11.9). The neuron firing process $X_t = Y_t + \eta/\gamma$ has the same equilibrium distribution shifted by the amount η/γ. ■

Example 11.8.2 *Equilibrium for a Recessive Disease Gene*

Equilibrium for a disease gene is maintained by the balance between selection and mutation. To avoid fixation of the deleterious allele and to ensure existence of the equilibrium distribution, back mutation of the deleterious allele to the normal allele must be incorporated into the model. In reality, the chance of fixation is so remote that back mutation does not enter into the following approximation of the equilibrium distribution $f(x)$. Because only small values of the disease gene frequency are likely, $f(x)$ is concentrated near 0. In the vicinity of 0, the approximation $x(1-x) \approx x$ holds. For a recessive disease, these facts suggest that we use

$$\begin{aligned} \frac{2\mu(y)}{\sigma^2(y)} &= \frac{2[\eta - (1-f)y^2]}{\frac{y(1-y)}{2N}} \\ &\approx 4N\left[\frac{\eta}{y} - (1-f)y\right] \end{aligned}$$

in Wright's formula (11.26) when the surrounding population size N is constant.

With this understanding,

$$\begin{aligned} f(x) &\approx \frac{2Nk_3}{x} e^{4N\eta \ln(x/l) - 2N(1-f)(x^2-l^2)} \\ &= k_4 x^{4N\eta-1} e^{-2N(1-f)x^2} \end{aligned}$$

for some constant $k_4 > 0$. The change of variables $z = 2N(1-f)x^2$ shows that the mth moment of $f(x)$ is

$$\begin{aligned} \int_I x^m f(x)\,dx &\approx k_4 \int_0^1 x^{m+4N\eta-1} e^{-2N(1-f)x^2} dx \\ &= \frac{k_4}{4N(1-f)} \int_0^1 x^{m+4N\eta-2} e^{-2N(1-f)x^2} 4N(1-f)x\,dx \\ &= \frac{k_4}{4N(1-f)} \int_0^{2N(1-f)} \left[\frac{z}{2N(1-f)}\right]^{\frac{m+4N\eta-2}{2}} e^{-z} dz \\ &\approx \frac{k_4}{2[2N(1-f)]^{\frac{m}{2}+2N\eta}} \int_0^\infty z^{\frac{m}{2}+2N\eta-1} e^{-z} dz \\ &= \frac{k_4 \Lambda(\frac{m}{2}+2N\eta)}{2[2N(1-f)]^{\frac{m}{2}+2N\eta}}. \end{aligned}$$

Taking $m = 0$ identifies the normalizing constant

$$k_4 = \frac{2[2N(1-f)]^{2N\eta}}{\Gamma(2N\eta)}.$$

With this value of k_4 in hand, the mean of $f(x)$ is

$$\int_I x f(x)\,dx \approx \frac{\Gamma(2N\eta+\frac{1}{2})}{\sqrt{2N(1-f)}\Gamma(2N\eta)}.$$

When $N\eta$ is large, application of Stirling's formula implies that the mean is close to the deterministic equilibrium value $\sqrt{\eta/(1-f)}$. In practice, one should be wary of applying the equilibrium theory because the approach to equilibrium is so slow. ■

11.9 Problems

1. Assuming that the increment $X_{t+s} - X_t$ is normally distributed with mean and variance given by equations (11.1) and (11.2), check the approximation (11.3) by taking conditional expectations in the inequality

$$\left|\frac{Y}{\sigma(t,x)\sqrt{s}}\right|^m \leq \sum_{i=1}^m \binom{m}{i} \left|\frac{Y-\omega}{\sigma(t,x)\sqrt{s}}\right|^i \left|\frac{\omega}{\sigma(t,x)\sqrt{s}}\right|^{m-i}$$

for $Y = X_{t+s} - X_t$ and $\omega = \mathrm{E}(X_{t+s} - X_t \mid X_t = x)$.

2. Demonstrate that standard Brownian motion is a Markov process. It suffices to check that

$$\begin{aligned} & \Pr(X_{t_n} \le u_n \mid X_{t_1} = u_1, \ldots, X_{t_{n-1}} = u_{n-1}) \\ = \; & \Pr(X_{t_n} \le u_n \mid X_{t_{n-1}} = u_{n-1}) \end{aligned}$$

for any two sequences $0 \le t_1 < \cdots < t_n$ and $u_1, \ldots, u_n$. (Hint: Use the independent increments property.)

3. Let X_t be standard Brownian motion. Calculate the mean and variance functions of the stochastic processes $|X_t|$ and e^{X_t}.

4. Standard Brownian motion X_t can be characterized by four postulates: (a) $\mathrm{E}(X_t) = 0$, (b) $\mathrm{Var}(X_t) = t$, (c) $\mathrm{Cov}(X_s, X_t) = \min\{s, t\}$, and (d) the random vector $(X_{t_1}, \ldots, X_{t_n})$ is multivariate normal for every finite collection $t_1 < \cdots < t_n$. Prove that conditions (c) and (d) can be replaced by (e) for $s < t$ the difference $X_t - X_s$ is univariate normal and (f) for every finite collection $t_1 < \cdots < t_n$, the random variables $X_{t_1}, X_{t_2} - X_{t_1}, \ldots, X_{t_n} - X_{t_{n-1}}$ are independent.

5. For standard Brownian motion X_t, prove that the stochastic process

$$Y_t = \begin{cases} tX_{1/t} & t > 0 \\ 0 & t = 0 \end{cases}$$

also furnishes a version of standard Brownian motion. (Hint: Demonstrate that Y_t satisfies either set of postulates mentioned in Problem 4.)

6. For standard Brownian motion X_t, it makes sense to define the integral $Y_t = \int_0^t X_s ds$ because sample paths are continuous functions of the time parameter. Argue that the stochastic process Y_t satisfies

$$\mathrm{E}(Y_t) = 0 \quad \text{and} \quad \mathrm{Cov}(Y_s, Y_t) = s^2\Big(\frac{t}{2} - \frac{s}{6}\Big)$$

for $0 \le s \le t$. Also show the random vector $(Y_{t_1}, \ldots, Y_{t_n})$ is multivariate normal for every finite collection $t_1 < \cdots < t_n$. (Hint: Approximate integrals by finite Riemann sums and pass to the limit.)

7. Let T_b be the first passage time to level $b > 0$ for standard Brownian motion. Verify that T_b and $b^2 T_1$ have the same distribution.

8. For standard Brownian motion X_t, show that the stochastic processes $Y_t = X_t$ and $Y_t = X_t^2 - t$ enjoy the martingale property

$$\mathrm{E}(Y_{t+s} \mid X_r,\ r \in [0, t]) = Y_t$$

for $s > 0$. (Hint: $X_{t+s} - X_t$ is independent of X_t and distributed as X_s.)

9. Continuing Problem 8, let T be the first time at which X_t attains the value $-a < 0$ or the value $b > 0$. Assuming the optional stopping theorem holds for the stopping time T and the martingales identified in Problem 8, demonstrate that

$$\Pr(X_T = -a) = \frac{b}{a+b} \quad \text{and} \quad \mathrm{E}(T) = ab.$$

10. Continuing Problem 8, let $Y_t = \sigma X_t + \alpha t$. Prove that $Y_t - \alpha t$ and $(Y_t - \alpha t)^2 - \sigma^2 t$ are martingales. Let T_b be the first passage time to level $b > 0$ for $\alpha > 0$. Assuming that the optional stopping theorem holds for T_b and these martingales, demonstrate that

$$\mathrm{E}(T_b) = \frac{b}{\alpha} \quad \text{and} \quad \mathrm{Var}(T_b) = \frac{b\sigma^2}{\alpha^3}.$$

11. Consider a diffusion process X_t with infinitesimal mean $\mu(t,x)$ and infinitesimal variance $\sigma^2(t,x)$. If the function $f(t)$ is strictly increasing and continuously differentiable, then argue that $Y_t = X_{f(t)}$ is a diffusion process with infinitesimal mean and variance

$$\begin{aligned} \mu_Y(t,y) &= \mu[f(t),y]f'(t) \\ \sigma_Y^2(t,y) &= \sigma^2[f(t),y]f'(t). \end{aligned}$$

Apply this result to the situation where Y_t equals y_0 at $t = 0$ and has $\mu_Y(t,y) = 0$ and $\sigma_Y^2(t,y) = \sigma^2(t)$. Show that Y_t is normally distributed with mean and variance

$$\begin{aligned} \mathrm{E}(Y_t) &= y_0 \\ \mathrm{Var}(Y_t) &= \int_0^t \sigma^2(s)\, ds. \end{aligned}$$

(Hint: Let X_t be standard Brownian motion.)

12. Show that

$$\mathrm{Cov}(Y_{t+s}, Y_t) = \frac{\sigma^2 e^{-\gamma s}(1 - e^{-2\gamma t})}{2\gamma}$$

in the Ornstein-Uhlenbeck process when s and t are nonnegative.

13. In the diffusion approximation to a branching process with immigration, we set $\mu(t,x) = (\alpha - \delta)x + \nu$ and $\sigma^2(t,x) = (\alpha + \delta)x + \nu$, where α and δ are the birth and death rates per particle and ν is the immigration rate. Demonstrate that

$$\begin{aligned} \mathrm{E}(X_t) &= x_0 e^{\beta t} + \frac{\nu}{\beta}\left[e^{\beta t} - 1\right] \\ \mathrm{Var}(X_t) &= \frac{\gamma x_0(e^{2\beta t} - e^{\beta t})}{\beta} + \frac{\gamma\nu(e^{2\beta t} - e^{\beta t})}{\beta^2} \\ &\quad - \frac{\gamma\nu(e^{2\beta t} - 1)}{2\beta^2} + \frac{\nu(e^{2\beta t} - 1)}{2\beta} \end{aligned}$$

for $\beta = \alpha - \delta$, $\gamma = \alpha + \delta$, and $X_0 = x_0$. When $\alpha < \delta$, the process eventually reaches equilibrium. Find the limits of $\mathrm{E}(X_t)$ and $\mathrm{Var}(X_t)$.

14. In Problem 13 suppose $\nu = 0$. Verify that the process goes extinct with probability $\min\{1, e^{-2\frac{\alpha-\delta}{\alpha+\delta}x_0}\}$ by using equation (11.20) and sending c to 0 and d to ∞.

15. Consider the transformed Brownian motion with infinitesimal mean α and infinitesimal variance σ^2 described in Example 11.3.3. If the process starts at $x \in [c, d]$, then prove that it reaches d before c with probability

$$u(x) \;=\; \frac{e^{-\beta x} - e^{-\beta c}}{e^{-\beta d} - e^{-\beta c}} \text{ for } \beta \;=\; \frac{2\alpha}{\sigma^2}.$$

Verify that $u(x)$ reduces to $(x - c)/(d - c)$ when $\alpha = 0$. As noted in the text, this simplification holds for any diffusion process with $\mu(x) = 0$.

16. Suppose the transformed Brownian motion with infinitesimal mean α and infinitesimal variance σ^2 described in Example 11.3.3 has $\alpha \geq 0$. If $c = -\infty$ and $d < \infty$, then demonstrate that equation (11.22) has solution

$$w(x) \;=\; e^{\gamma(d-x)} \text{ for } \gamma \;=\; \frac{\alpha - \sqrt{\alpha^2 + 2\sigma^2\theta}}{\sigma^2}.$$

Simplify $w(x)$ when $\alpha = 0$, and show by differentiation of $w(x)$ with respect to θ that the expected time $\mathrm{E}(T)$ to reach the barrier d is infinite. When $\alpha < 0$, show that

$$\Pr(T < \infty) \;=\; e^{\frac{2\alpha}{\sigma^2}(d-x)}.$$

(Hints: The variable γ is a root of a quadratic equation. Why do we discard the other root? In general, $\Pr(T < \infty) = \lim_{\theta \downarrow 0} \mathrm{E}\left(e^{-\theta T}\right)$.)

17. In Problem 16 find $w(x)$ and $\mathrm{E}(T)$ when c is finite. The value $\alpha < 0$ is allowed.

18. Show that formula (11.24) holds in R^2.

19. Consider a diffusion process X_t with infinitesimal mean

$$\mu(t, x) \;=\; \begin{cases} 1, & x < 0 \\ 0, & x = 0 \\ -1, & x > 0 \end{cases}$$

and infinitesimal variance 1. Find the equilibrium distribution $f(x)$ of X_t.

20. In Problem 13 suppose $\nu > 0$ and $\alpha < \delta$. Show that Wright's formula leads to the equilibrium distribution

$$f(x) \quad = \quad k\left[(\alpha+\delta)x+\nu\right]^{\frac{4\nu\delta}{(\alpha+\delta)^2}-1} e^{\frac{2(\alpha-\delta)x}{\alpha+\delta}}$$

for some normalizing constant $k > 0$ and $x > 0$.

21. Use Stirling's formula to demonstrate that

$$\frac{\Gamma(2N\eta+\frac{1}{2})}{\sqrt{2N(1-f)}\Gamma(2N\eta)} \quad \approx \quad \sqrt{\frac{\eta}{1-f}}$$

when N is large in the Wright-Fisher model for a recessive disease.

22. Consider the Wright-Fisher model with no selection but with mutation from allele A_1 to allele A_2 at rate η_1 and from A_2 to A_1 at rate η_2. With constant population size N, prove that the frequency of the A_1 allele follows the beta distribution

$$f(x) \quad = \quad \frac{\Gamma[4N(\eta_1+\eta_2)]}{\Gamma(4N\eta_2)\Gamma(4N\eta_1)} x^{4N\eta_2-1}(1-x)^{4N\eta_1-1}$$

at equilibrium. (Hint: Substitute $p(x) = x$ in formula (11.14) defining the infinitesimal variance $\sigma^2(t,x)$.)

12
Asymptotic Methods

12.1 Introduction

Long before computers revolutionized numerical analysis, applied mathematicians devised many clever techniques for finding approximate answers to hard problems. Because approximate solutions focus on dominant contributions, they often provide more insight than exact solutions. In this chapter we take up the subject of asymptotic analysis. Although this material is old, dating back centuries in some cases, it still has its charms and utility. Our choice of topics differs from the typical syllabus of mathematical statistics, where the emphasis is on large sample theory and convergence in distribution [61, 132, 181]. Here we stress advanced calculus and combinatorics.

The next section begins by reviewing order relations and asymptotic equivalence. The basic definitions are then illustrated by some concrete examples involving Taylor expansions, summation by parts, and integration by parts. Laplace's method and Watson's lemma are more subtle and demanding subjects. These two pillars of asymptotic analysis seek to approximate otherwise intractable integrals. Our examples include Stirling's formula, the birthday problem, the socks in the laundry problem, and approximation of Catalan numbers. The Euler-Maclaurin summation formula bridges the gap between series and integrals of the same function. In practice, integrals are usually easier to evaluate. The section on generating functions and partial fraction decompositions hints at the wider application of analytic function theory in deriving asymptotic expansions. Our

K. Lange, *Applied Probability,*
DOI 10.1007/978-1-4419-7165-4_12, © Springer Science+Business Media, LLC 2010

concluding section mentions a few highlights of large sample theory and more general forms of the central limit theorem.

To keep the chapter within reasonable bounds, we have moved some of the theoretical background to Appendices A.5 and A.6. Our diluted discussion of analytic functions is adequate for our limited needs. Readers who seriously want to pursue asymptotic methods should study analytic function theory (complex variables) in greater depth. Two readable accounts are [93, 188], but there are a host of other good choices. For follow-up on asymptotic methods, the books [19, 20, 46, 50, 78, 147, 207] are excellent.

12.2 Asymptotic Expansions

Asymptotic analysis is the branch of mathematics dealing with the order of magnitude and limiting behavior of functions, particularly at boundary points of their domains of definition [19, 20, 46, 78, 147]. Consider, for instance, the function

$$f(x) \;=\; \frac{x^2+1}{x+1}.$$

It is obvious that $f(x)$ resembles the function x as $x \to \infty$. However, one can be more precise. The expansion

$$\begin{aligned} f(x) &= \frac{x^2+1}{x(1+\frac{1}{x})} \\ &= \left(x+\frac{1}{x}\right)\sum_{k=0}^{\infty}\left(\frac{-1}{x}\right)^k \\ &= x-1-2\sum_{k=1}^{\infty}\left(\frac{-1}{x}\right)^k \end{aligned}$$

indicates that $f(x)$ more closely resembles $x-1$ for large x. Furthermore, $f(x)-x+1$ behaves like $2/x$ for large x. We can refine the precision of the approximation by taking more terms in the infinite series. How far we continue in this and other problems is usually dictated by the application at hand.

12.2.1 Order Relations

Order relations are central to the development of asymptotic analysis. Suppose we have two functions $f(x)$ and $g(x)$ defined on a common interval I, which may extend to ∞ on the right or to $-\infty$ on the left. Let x_0 be either an internal point or a boundary point of I with $g(x) \neq 0$ for x close, but not equal, to x_0. Then the function $f(x)$ is said to be $O(g(x))$

if there exists a constant M such that $|f(x)| \leq M|g(x)|$ as $x \to x_0$. If $\lim_{x\to x_0} f(x)/g(x) = 0$, then $f(x)$ is said to be $o(g(x))$. Obviously, the relation $f(x) = o(g(x))$ implies the weaker relation $f(x) = O(g(x))$. Finally, if $\lim_{x\to x_0} f(x)/g(x) = 1$, then $f(x)$ is said to be asymptotic to $g(x)$. This is usually written $f(x) \asymp g(x)$. In many problems, the functions $f(x)$ and $g(x)$ are defined on the integers $\{1, 2, \ldots\}$ instead of on an interval I, and x_0 is taken as ∞.

For example, on $I = (1, \infty)$ one has $e^x = O(\sinh x)$ as $x \to \infty$ because

$$\frac{e^x}{\frac{e^x - e^{-x}}{2}} \quad = \quad \frac{2}{1 - e^{-2x}} \quad \leq \quad \frac{2}{1 - e^{-2}}.$$

On $(0, \infty)$ one has $\sin^2 x = o(x)$ as $x \to 0$ because

$$\lim_{x\to 0} \frac{\sin^2 x}{x} \quad = \quad \lim_{x\to 0} \sin x \lim_{x\to 0} \frac{\sin x}{x} \quad = \quad 0 \times 1.$$

On $I = (0, \infty)$, our initial example can be rephrased as $(x^2+1)/(x+1) \asymp x$ as $x \to \infty$.

If $f(x)$ is bounded in a neighborhood of x_0, then we write $f(x) = O(1)$ as $x \to x_0$, and if $\lim_{x\to x_0} f(x) = 0$, we write $f(x) = o(1)$ as $x \to x_0$. The notation $f(x) = g(x)+O(h(x))$ means $f(x)-g(x) = O(h(x))$ and similarly for the o notation. For example,

$$\frac{x^2 + 1}{x + 1} \quad = \quad x - 1 + O\Big(\frac{1}{x}\Big).$$

If $f(x)$ is differentiable at point x_0, then

$$f(x_0 + h) - f(x_0) \quad = \quad f'(x_0)h + o(h).$$

There are a host of miniature theorems dealing with order relations. Among these are

$$\begin{aligned}
O(g) + O(g) &= O(g) \\
o(g) + o(g) &= o(g) \\
O(g_1)O(g_2) &= O(g_1 g_2) \\
o(g_1)O(g_2) &= o(g_1 g_2) \\
|O(g)|^\lambda &= O(|g|^\lambda), \quad \lambda > 0 \\
|o(g)|^\lambda &= o(|g|^\lambda), \quad \lambda > 0.
\end{aligned}$$

12.2.2 Finite Taylor Expansions

One easy way of generating approximations to a function is via finite Taylor expansions. Suppose $f(x)$ has $n + 1$ continuous derivatives near $x_0 = 0$.

Then

$$f(x) = \sum_{k=0}^{n} \frac{1}{k!} f^{(k)}(0) x^k + O(x^{n+1})$$

as $x \to 0$. This order relation is validated by l'Hôpital's rule applied $n+1$ times to the quotient

$$\frac{f(x) - \sum_{k=0}^{n} \frac{1}{k!} f^{(k)}(0) x^k}{x^{n+1}}.$$

Of course, it is more informative to write the Taylor expansion with an explicit error term; for instance,

$$f(x) = \sum_{k=0}^{n} \frac{1}{k!} f^{(k)}(0) x^k + \frac{x^{n+1}}{n!} \int_0^1 f^{(n+1)}(tx)(1-t)^n dt. \quad (12.1)$$

This integral $\frac{x^{n+1}}{n!} \int_0^1 f^{(n+1)}(tx)(1-t)^n dt$ form of the remainder $R_n(x)$ after n terms can be derived by noting the recurrence relation

$$R_n(x) = -\frac{x^n}{n!} f^{(n)}(0) + R_{n-1}(x)$$

and the initial condition

$$R_0(x) = f(x) - f(0),$$

both of which follow from integration by parts. One virtue of formula (12.1) emerges when the derivatives of $f(x)$ satisfy $(-1)^k f^{(k)}(x) \geq 0$ for all $k > 0$. If this condition holds, then

$$\begin{aligned} 0 &\leq (-1)^{n+1} R_n(x) \\ &= \frac{x^{n+1}}{n!} \int_0^1 (-1)^{n+1} f^{(n+1)}(tx)(1-t)^n dt \\ &\leq \frac{x^{n+1}}{n!} (-1)^{n+1} f^{(n+1)}(0) \int_0^1 (1-t)^n dt \\ &= \frac{x^{n+1}}{(n+1)!} (-1)^{n+1} f^{(n+1)}(0) \end{aligned}$$

for any $x > 0$. In other words, the remainders $R_n(x)$ alternate in sign and are bounded in absolute value by the next term of the expansion. As an example, the function $f(x) = -\ln(1+x)$ satisfies the inequalities $(-1)^k f^{(k)}(x) \geq 0$ and consequently also an infinity of Taylor expansion inequalities beginning with

$$-x \leq -\ln(1+x) \leq -x + x^2/2.$$

Another appropriate function is e^{-x}. The analogous inequalities are now

$$1 - x \leq e^{-x} \leq 1 - x + x^2/2$$

for $x > 0$.

12.2.3 Exploitation of Nearby Exact Results

One of the tactics of asymptotic analysis is to evaluate nearby exact expressions and then estimate the error of the difference.

Example 12.2.1 *Right-Tail Probability of the Logarithmic Distribution*

Suppose we want to estimate the right-tail probability

$$r_n = -\frac{1}{\ln(1-\theta)} \sum_{k=n}^{\infty} \frac{\theta^k}{k}$$

of a logarithmic distribution, where $n > 1$ and $0 < \theta < 1$. Fortunately, the exact geometric sum

$$\sum_{k=n}^{\infty} \frac{\theta^k}{n} = \frac{\theta^n}{n(1-\theta)}$$

is available. To take advantage of this result, we employ the easily checked summation by parts formula

$$\sum_{k=n}^{m} a_k b_k = \sum_{k=n}^{m-1} c_k(b_k - b_{k+1}) + c_m b_m, \tag{12.2}$$

where $c_k = \sum_{j=n}^{k} a_j$. If c_k converges and b_k converges to 0, then equation (12.2) implies

$$\sum_{k=n}^{\infty} a_k b_k = \sum_{k=n}^{\infty} c_k(b_k - b_{k+1}),$$

provided either side converges. Now set $a_k = \theta^k$ and $b_k = k^{-1}$. With these choices

$$c_k = \frac{\theta^n - \theta^{k+1}}{1-\theta}$$

and

$$\begin{aligned} \sum_{k=n}^{\infty} \frac{\theta^k}{k} &= \sum_{k=n}^{\infty} \frac{\theta^n - \theta^{k+1}}{1-\theta}\Big(\frac{1}{k} - \frac{1}{k+1}\Big) \\ &= \frac{\theta^n}{n(1-\theta)} - \sum_{k=n}^{\infty} \frac{\theta^{k+1}}{1-\theta} \frac{1}{k(k+1)}. \end{aligned}$$

This representation makes it obvious that

$$r_n = -\frac{1}{\ln(1-\theta)} \frac{\theta^n}{n(1-\theta)} \Big[1 + O\Big(\frac{1}{n}\Big)\Big]$$

since the series $\sum_{k=0}^{\infty} \theta^k$ converges absolutely and the factor $[k(k+1)]^{-1}$ is bounded above by n^{-2} for $k \geq n$. ■

12.2.4 Expansions via Integration by Parts

Integration by parts often works well as a formal device for generating asymptotic expansions. Here are two examples.

Example 12.2.2 *Exponential Integral*

Suppose Y has exponential density e^{-y} with unit mean. Given Y, let a point X be chosen uniformly from the interval $[0, Y]$. Then it is easy to show that X has density $E_1(x) = \int_x^\infty e^{-y}y^{-1}dy$ and distribution function $1 - e^{-x} + xE_1(x)$. To generate an asymptotic expansion of the exponential integral $E_1(x)$ as $x \to \infty$, one can repeatedly integrate by parts. This gives

$$\begin{aligned} E_1(x) &= -\frac{e^{-y}}{y}\Big|_x^\infty - \int_x^\infty \frac{e^{-y}}{y^2}dy \\ &= \frac{e^{-x}}{x} + \frac{e^{-y}}{y^2}\Big|_x^\infty + 2\int_x^\infty \frac{e^{-y}}{y^3}dy \\ &\vdots \\ &= e^{-x}\sum_{k=1}^{n}(-1)^{k-1}\frac{(k-1)!}{x^k} + (-1)^n n!\int_x^\infty \frac{e^{-y}}{y^{n+1}}dy. \end{aligned}$$

This is emphatically not a convergent series in powers of $1/x$. In fact, for any fixed x, we have $\lim_{k\to\infty}|(-1)^{(k-1)}(k-1)!/x^k| = \infty$.

Fortunately, the remainders $R_n(x) = (-1)^n n!\int_x^\infty e^{-y}y^{-n-1}dy$ alternate in sign and are bounded in absolute value by

$$\begin{aligned} |R_n(x)| &\le \frac{n!}{x^{n+1}}\int_x^\infty e^{-y}dy \\ &= \frac{n!}{x^{n+1}}e^{-x}, \end{aligned}$$

the absolute value of the next term of the expansion. This suggests that we truncate the expansion when n is the largest integer with

$$\frac{\frac{n!}{x^{n+1}}e^{-x}}{\frac{(n-1)!}{x^n}e^{-x}} \le 1.$$

In other words, we should choose $n \approx x$. If we include more terms, then the approximation degrades. This is in striking contrast to what happens with a convergent series.

Table 12.1 illustrates these remarks by tabulating a few representative values of the functions

$$\begin{aligned} I(x) &= xe^x E_1(x) \\ S_n(x) &= \sum_{k=1}^{n}(-1)^{k-1}\frac{(k-1)!}{x^{k-1}}. \end{aligned}$$

TABLE 12.1. Asymptotic Approximation of the Exponential Integral

x	$I(x)$	$S_1(x)$	$S_2(x)$	$S_3(x)$	$S_4(x)$	$S_5(x)$	$S_6(x)$
1	0.59634	1.0	0.0	2.0	-4.0		
2	0.72266	1.0	0.5	1.0	0.25	1.75	
3	0.78625	1.0	0.667	0.8999	0.6667	0.9626	0.4688
5	0.85212	1.0	0.8	0.88	0.8352	0.8736	0.8352

For large x, the approximation noticeably improves. Thus, $I(10) = 0.91563$ while $S_{10}(10) = 0.91544$, and $I(100) = 0.99019 = S_4(100)$. ■

Example 12.2.3 *Incomplete Gamma Function*

Repeated integration by parts of the right-tail probability of a gamma distributed random variable produces in the same manner

$$\begin{aligned} & \frac{1}{\Gamma(a)} \int_x^\infty y^{a-1} e^{-y} dy \\ = \; & x^a e^{-x} \sum_{k=1}^{n} \frac{1}{x^k \Gamma(a-k+1)} + \frac{1}{\Gamma(a-n)} \int_x^\infty y^{a-n-1} e^{-y} dy. \end{aligned}$$

If a is a positive integer, then the expansion stops at $n = a$ with remainder 0. Otherwise, if n is so large that $a - n - 1$ is negative, then the remainder satisfies

$$\left| \frac{1}{\Gamma(a-n)} \int_x^\infty y^{a-n-1} e^{-y} dy \right| \leq \left| \frac{1}{\Gamma(a-n)} \right| x^{a-n-1} e^{-x}.$$

Reasoning as above, we deduce that it is optimal to truncate the expansion when $|a - n|/x \approx 1$. The right-tail probability

$$\frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-\frac{y^2}{2}} dy = \frac{1}{2\Gamma(\frac{1}{2})} \int_{\frac{x^2}{2}}^\infty z^{\frac{1}{2}-1} e^{-z} dz$$

of the standard normal random variable is covered by the special case $a = 1/2$ for $x > 0$; namely,

$$\frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-\frac{y^2}{2}} dy = \frac{e^{-\frac{x^2}{2}}}{x\sqrt{2\pi}} \left(1 - \frac{1}{x^2} + \frac{3}{x^4} - \frac{3 \cdot 5}{x^6} + \cdots\right).$$

Problem 14 bounds these right-tail probabilities. ■

The previous examples suggest Poincaré's definition of an asymptotic expansion. Let $\phi_n(x)$ be a sequence of functions such that $\phi_{n+1}(x) = o(\phi_n(x))$ as $x \to x_0$. Then $\sum_{k=1}^\infty c_k \phi_k(x)$ is an asymptotic expansion for $f(x)$ if

$f(x) = \sum_{k=1}^{n} c_k\phi_k(x) + o(\phi_n(x))$ holds as $x \to x_0$ for every $n \geq 1$. The constants c_n are uniquely determined by the limits

$$c_n = \lim_{x\to x_0} \frac{f(x) - \sum_{k=1}^{n-1} c_k\phi_k(x)}{\phi_n(x)}$$

taken recursively starting with $c_1 = \lim_{x\to x_0} f(x)/\phi_1(x)$. Implicit in this definition is the assumption that $\phi_n(x) \neq 0$ for x close, but not equal, to x_0.

12.3 Laplace's Method

Laplace's method gives asymptotic approximations for integrals

$$\int_c^d f(y)e^{-xg(y)}dy \tag{12.3}$$

depending on a parameter x as $x \to \infty$. Here the boundary points c and d can be finite or infinite. There are two cases of primary interest. If c is finite, and the minimum of $g(y)$ occurs at c, then the contributions to the integral around c dominate as $x \to \infty$. Without loss of generality, let us take $c = 0$ and $d = \infty$. (If d is finite, then we can extend the range of integration by defining $f(x) = 0$ to the right of d.) Now the supposition that the dominant contributions occur around 0 suggests that we can replace $f(y)$ by $f(0)$ and $g(y)$ by its first-order Taylor expansion $g(y) \approx g(0) + g'(0)y$. Making these substitutions leads us to conjecture that

$$\begin{aligned} \int_0^\infty f(y)e^{-xg(y)}dy &\asymp f(0)e^{-xg(0)} \int_0^\infty e^{-xyg'(0)}dy \\ &= \frac{f(0)e^{-xg(0)}}{xg'(0)}. \end{aligned} \tag{12.4}$$

In essence, we have reduced the integral to integration against the exponential density with mean $[xg'(0)]^{-1}$. As this mean approaches 0, the approximation becomes better and better. Under the weaker assumption that $f(y) \asymp ay^{b-1}$ as $y \to 0$ for $b > 0$, the integral (12.3) can be replaced by an integral involving a gamma density. In this situation,

$$\int_0^\infty f(y)e^{-xg(y)}dy \asymp \frac{a\Gamma(b)e^{-xg(0)}}{[xg'(0)]^b} \tag{12.5}$$

as $x \to \infty$.

The other case occurs when $g(y)$ assumes its minimum at an interior point, say 0, between, say, $c = -\infty$ and $d = \infty$. Now we replace $g(y)$ by its

second-order Taylor expansion $g(y) = g(0) + \frac{1}{2}g''(0)y^2 + o(y^2)$. Assuming that the region around 0 dominates, we conjecture that

$$\begin{aligned}\int_{-\infty}^{\infty} f(y)e^{-xg(y)}dy &\asymp f(0)e^{-xg(0)}\int_{-\infty}^{\infty} e^{-\frac{xg''(0)y^2}{2}}dy \\ &= f(0)e^{-xg(0)}\sqrt{\frac{2\pi}{xg''(0)}}. \qquad (12.6)\end{aligned}$$

In other words, we reduce the integral to integration against the normal density with mean 0 and variance $[xg''(0)]^{-1}$. As this variance approaches 0, the approximation improves. The asymptotic equivalences (12.5) and (12.6) and their generalizations constitute Laplace's method. Appendix A.6 rigorously states and proves the second of these conjectures.

Laplace's method has some interesting applications to order statistics. Let $X_1, \ldots, X_n$ be i.i.d. positive random variables with common distribution function $F(x)$. We assume that $F(x) \asymp ax^b$ as $x \to 0$. Now consider the first order statistic $X_{(1)} = \min_{1 \le i \le n} X_i$. One can express the kth moment of $X_{(1)}$ in terms of its right-tail probability

$$\Pr(X_{(1)} > x) = [1 - F(x)]^n$$

as

$$\begin{aligned}\mathrm{E}(X_{(1)}^k) &= k\int_0^{\infty} x^{k-1}[1 - F(x)]^n dx \\ &= k\int_0^{\infty} x^{k-1}e^{n\ln[1-F(x)]}dx \\ &= \frac{k}{b}\int_0^{\infty} u^{\frac{k}{b}-1}e^{n\ln[1-F(u^{\frac{1}{b}})]}du,\end{aligned}$$

where the last integral arises from the change of variable $u = x^b$. Now the function $g(u) = -\ln[1 - F(u^{\frac{1}{b}})]$ has its minimum at $u = 0$, and an easy calculation invoking the assumption $F(x) \asymp ax^b$ yields $g(u) \asymp au$ as $u \to 0$. Hence, the first form (12.5) of Laplace's method implies

$$\mathrm{E}(X_{(1)}^k) \asymp \frac{k\Gamma(\frac{k}{b})}{b(na)^{\frac{k}{b}}}. \qquad (12.7)$$

Example 12.3.1 *Asymptotics of the Birthday Problem*

This result has an amusing consequence for a birthday problem. Suppose that people are selected one by one from a large crowd until two of the chosen people share a birthday. We would like to know how many people are selected on average before a match occurs. One way of conceptualizing this problem is to imagine drawing people at random times dictated by a Poisson process with unit intensity. The expected time until the first match

then coincides with the expected number of people drawn [26]. Example 6.6.2 takes this Poissonization approach in studying the family planning model. Since the choice of a birthday from the available $n = 365$ days of the year is made independently for each random draw, we are in effect watching the evolution of n independent Poisson processes, each with intensity $1/n$.

Let X_i be the time when the second random point happens in the ith process. The time when the first birthday match occurs in the overall process is $X_{(1)} = \min_{1 \le i \le n} X_i$. Now X_i has right-tail probability

$$\Pr(X_i > x) = \left(1 + \frac{x}{n}\right) e^{-\frac{x}{n}}$$

because zero or one random points must occur on $[0, x]$ in order for $X_i > x$. It follows that X_i has distribution function

$$\begin{aligned} \Pr(X_i \le x) &= 1 - \left(1 + \frac{x}{n}\right) e^{-\frac{x}{n}} \\ &\asymp \frac{x^2}{2n^2}, \end{aligned}$$

and according to our calculation (12.7) with $a = 1/(2n^2)$, $b = 2$, and $k = 1$,

$$\mathrm{E}(X_{(1)}) \asymp \frac{\Gamma(\frac{1}{2})}{2(n\frac{1}{2n^2})^{\frac{1}{2}}} = \frac{1}{2}\sqrt{2\pi n}.$$

For $n = 365$ we get $\mathrm{E}(X_{(1)}) \approx 23.9$, a reasonably close approximation to the true value of 24.6. ■

12.3.1 Stirling's Formula

The behavior of the gamma function

$$\Gamma(\lambda) = \int_0^\infty y^{\lambda-1} e^{-y} dy$$

as $\lambda \to \infty$ can be ascertained by Laplace's method. If we define $z = y/\lambda$, then

$$\Gamma(\lambda + 1) = \lambda^{\lambda+1} \int_0^\infty e^{-\lambda g(z)} dz$$

for the function $g(z) = z - \ln z$, which has its minimum at $z = 1$. Applying Laplace's second approximation (12.6) at $z = 1$ gives Stirling's asymptotic formula

$$\Gamma(\lambda + 1) \asymp \sqrt{2\pi}\lambda^{\lambda+\frac{1}{2}} e^{-\lambda}$$

as $\lambda \to \infty$. We will rederive Stirling's formula in the next section using the machinery of the Euler-Maclaurin summation technique. ■

Example 12.3.2 *Asymptotics of Socks in the Laundry*

In the socks in the laundry problem introduced in Section 4.7, the independence of the uniform processes entails

$$\Pr[X_{(1)} > t] \;=\; \Pr\Big(\min_{1\le i\le n} X_i > t\Big) \;=\; (1-t^2)^n$$

because any given pair of socks arrives after time t with probability $1-t^2$. Integrating this tail probability with respect to t produces

$$\begin{aligned} \mathrm{E}[X_{(1)}] &= \int_0^1 (1-t^2)^n dt \\ &= \int_0^1 e^{n\ln(1-t^2)} dt \\ &\asymp \int_0^\infty e^{-nt^2} dt \\ &= \frac{\sqrt{\pi}}{2\sqrt{n}} \end{aligned} \tag{12.8}$$

and therefore $\mathrm{E}(N_1) = (2n+1)\,\mathrm{E}[X_{(1)}] \asymp \sqrt{\pi n}$, where N_1 is the number of socks extracted until the first matching pair. The essence of Laplace's approximation is simply the recognition that the vast majority of the mass of the integral $\int_0^1 (1-t^2)^n\,dt$ occurs near $t=0$, where $\ln(1-t^2) \approx -t^2$. Similar reasoning leads to

$$\begin{aligned} \mathrm{Var}(N_1) &= (2n+1)(2n+2)\,\mathrm{E}[X_{(1)}^2] - \mathrm{E}(N_1) - \mathrm{E}(N_1)^2 \\ &\asymp (4-\pi)n \end{aligned}$$

for the variance. ■

12.3.2 Watson's Lemma

Watson's lemma is a valuable addition to Laplace's method. Suppose the continuous function $f(x)$ defined on $[0,\infty)$ possesses the asymptotic expansion

$$f(x) \;\asymp\; \sum_{k=0}^{\infty} a_k x^{\lambda_k - 1}$$

as $x \to 0$ and $f(x) = O(e^{cx})$ as $x \to \infty$. The sequence λ_k is taken to be positive and strictly increasing. Then the Laplace transform $\tilde{f}(t)$ exists for $t > c$, and $\tilde{f}(t)$ has the asymptotic expansion

$$\tilde{f}(t) \;\asymp\; \sum_{k=0}^{\infty} a_n \frac{\Gamma(\lambda_k)}{t^{\lambda_k}}$$

as $t \to \infty$. This expansion is validated in Appendix A.6.

Example 12.3.3 *Beta Normalizing Constant*

The beta distribution normalizing constant is defined by

$$B(\alpha,\beta) \;=\; \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx \;=\; \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}.$$

If we restrict α to be an integer, then $B(\alpha,\beta) \asymp \Gamma(\beta)\alpha^{-\beta}$ for α large. It is reasonable to conjecture that this relation continues to hold for all large α. If we make the change of variables $x = e^{-y}$, then

$$B(\alpha,\beta) \;=\; \int_0^\infty e^{-\alpha y}(1-e^{-y})^{\beta-1}dy \;=\; \int_0^\infty e^{-\alpha y}y^{\beta-1}\sum_{k=0}^\infty a_k y^k dy$$

for some sequence of coefficients a_k with $a_0 = 1$. The first term of Watson's expansion says $B(\alpha,\beta) \asymp \Gamma(\beta)\alpha^{-\beta}$ for β fixed. Obviously this also entails the relation $\Gamma(\alpha)/\Gamma(\alpha+\beta) \asymp \alpha^{-\beta}$. ■

Example 12.3.4 *Catalan Asymptotics*

The Catalan numbers c_n introduced in Section 4.5 have generating function

$$\begin{aligned} c(x) &= \frac{1-\sqrt{1-4x}}{2x} \\ &= -\frac{1}{2x}\sum_{n=1}^\infty \binom{\frac{1}{2}}{n}(-4x)^n \\ &= \sum_{n=1}^\infty \frac{(n-\frac{3}{2})\cdots\frac{1}{2}}{n!}4^{n-1}x^{n-1} \\ &= \sum_{n=1}^\infty \frac{\Gamma(n-\frac{1}{2})}{\Gamma(\frac{1}{2})\Gamma(n+1)}4^{n-1}x^{n-1}. \end{aligned}$$

Because $\Gamma(\frac{1}{2}) = \sqrt{\pi}$ and $\Gamma(n-\frac{1}{2})/\Gamma(n+1) \asymp n^{-3/2}$, it follows that

$$c_n \;\asymp\; \frac{4^n n^{-3/2}}{\sqrt{\pi}}$$

for large n. ■

12.4 Euler-Maclaurin Summation Formula

We now turn to the Euler-Maclaurin summation formula [27, 78, 101, 204], another useful tool in asymptotic analysis. The summation formula and its proof depend on the Bernoulli numbers B_n, the Bernoulli polynomials

$B_n(x)$, and the periodic Bernoulli functions $b_n(x)$. The first few Bernoulli numbers $B_n = B_n(0)$ are $B_0 = 1$, $B_1 = -1/2$, $B_2 = 1/6$, $B_3 = 0$, and $B_4 = -1/30$. The odd terms B_{2n+1} vanish for $n \geq 1$. Example A.5.1 derives the Bernoulli polynomials from the three simple properties displayed in equation (A.10). The first two Bernoulli polynomials are $B_0(x) = 1$ and $B_1(x) = x - \frac{1}{2}$. The Bernoulli functions periodically extend the Bernoulli polynomials beyond the base domain [0,1]. Although $B_1(x)$ does not satisfy the periodic property $B_n(0) = B_n(1)$, all subsequent $B_n(x)$ do. Problems 19 through 25 explore further aspects of this fascinating nook of calculus. The next proposition presents one form of the Euler-Maclaurin summation formula based on this background material.

Proposition 12.4.1 *Suppose $f(x)$ has $2m$ continuous derivatives on the interval $[1, n]$ for some positive integer n. Then*

$$\begin{aligned}\sum_{k=1}^{n} f(k) &= \int_1^n f(x)dx + \frac{1}{2}\Big[f(n) + f(1)\Big] + \sum_{j=1}^{m} \frac{B_{2j}}{(2j)!} f^{(2j-1)}(x)\Big|_1^n \\ &\quad - \frac{1}{(2m)!}\int_1^n b_{2m}(x) f^{(2m)}(x)dx, \qquad (12.9)\end{aligned}$$

where B_k is a Bernoulli number and $b_k(x)$ is a Bernoulli function. The remainder in this expansion is bounded by

$$\left|\frac{1}{(2m)!}\int_1^n b_{2m}(x) f^{(2m)}(x)dx\right| \leq C_{2m}\int_1^n |f^{(2m)}(x)|dx, \quad (12.10)$$

where

$$C_{2m} = \frac{2}{(2\pi)^{2m}}\sum_{k=1}^{\infty}\frac{1}{k^{2m}}.$$

Proof: Consider an arbitrary function $g(x)$ defined on $[0, 1]$ with $2m$ continuous derivatives. In view of the property $\frac{d}{dx}B_n(x) = nB_{n-1}(x)$, repeated integration by parts gives

$$\begin{aligned}\int_0^1 g(x)dx &= \int_0^1 B_0(x)g(x)dx \\ &= B_1(x)g(x)|_0^1 - \int_0^1 B_1(x)g'(x)dx \\ &= \sum_{i=1}^{2m}\frac{(-1)^{i-1}B_i(x)}{i!}g^{(i-1)}(x)\Big|_0^1 \\ &\quad + \frac{(-1)^{2m}}{(2m)!}\int_0^1 B_{2m}(x)g^{(2m)}(x)dx.\end{aligned}$$

This formula can be simplified by noting that (a) $B_{2m}(x) = b_{2m}(x)$ on $[0,1]$, (b) $B_1(x) = x - 1/2$, (c) $B_i(0) = B_i(1) = B_i$ when $i > 1$, and (d) $B_i = 0$ when $i > 1$ and i is odd. Hence,

$$\begin{aligned}\int_0^1 g(x)dx &= \frac{1}{2}\Big[g(1)+g(0)\Big] - \sum_{j=1}^{m} \frac{B_{2j}}{(2j)!} g^{(2j-1)}(x)\Big|_0^1 \\ &\quad + \frac{1}{(2m)!}\int_0^1 b_{2m}(x) g^{(2m)}(x)dx.\end{aligned}$$

If we apply this result successively to $g(x) = f(x+k)$ for $k = 1, \ldots, n-1$ and add the results, then cancellation of successive terms produces formula (12.9). The bound (12.10) follows immediately from the Fourier series representation (A.13) of $b_{2m}(x)$. ■

Example 12.4.1 *Harmonic Series*

The harmonic series $\sum_{k=1}^{n} k^{-1}$ can be approximated by taking $f(x)$ to be x^{-1} in Proposition 12.4.1. For example with $m = 2$, we find that

$$\begin{aligned}\sum_{k=1}^{n} \frac{1}{k} &= \int_1^n \frac{1}{x}dx + \frac{1}{2}\left[\frac{1}{n}+1\right] + \frac{B_2}{2}\left[1 - \frac{1}{n^2}\right] \\ &\quad + \frac{B_4}{4!}\left[3! - \frac{3!}{n^4}\right] - \frac{1}{4!}\int_1^n b_4(x)\frac{4!}{x^5}dx \\ &= \ln n + \gamma + \frac{1}{2n} - \frac{1}{12n^2} + \frac{1}{120n^4} + \int_n^\infty b_4(x)\frac{1}{x^5}dx \\ &= \ln n + \gamma + \frac{1}{2n} - \frac{1}{12n^2} + O\Big(\frac{1}{n^4}\Big),\end{aligned}$$

where

$$\begin{aligned}\gamma &= \frac{1}{2} + \frac{1}{12} - \frac{1}{120} - \int_1^\infty b_4(x)\frac{1}{x^5}dx \qquad (12.11)\\ &\approx 0.5772\end{aligned}$$

is Euler's constant. ■

Example 12.4.2 *Stirling's Formula Again*

If we let $f(x)$ be the function $\ln x = \frac{d}{dx}[x\ln x - x]$ and $m = 2$ in Proposition 12.4.1, then we recover Stirling's formula

$$\begin{aligned}\ln n! &= \sum_{k=1}^{n} \ln k \\ &= \int_1^n \ln x\, dx + \frac{1}{2}\ln n + \frac{B_2}{2}\left[\frac{1}{n} - 1\right]\end{aligned}$$

$$
\begin{aligned}
& + \frac{B_4}{4!}\left[\frac{2!}{n^3} - 2!\right] + \frac{1}{4!}\int_1^n b_4(x)\frac{3!}{x^4}dx \\
= \;& n\ln n - n + \frac{1}{2}\ln n + s + \frac{1}{12n} - \frac{1}{360n^3} - \frac{1}{4}\int_n^\infty b_4(x)\frac{1}{x^4}dx \\
= \;& \Big(n + \frac{1}{2}\Big)\ln n - n + s + \frac{1}{12n} + O\Big(\frac{1}{n^3}\Big),
\end{aligned}
$$

where

$$
\begin{aligned}
s &= 1 - \frac{1}{12} + \frac{1}{360} + \frac{1}{4}\int_1^\infty b_4(x)\frac{1}{x^4}dx \qquad (12.12)\\
&= \ln\sqrt{2\pi}
\end{aligned}
$$

was determined in Section 12.3.1. ■

12.5 Asymptotics and Generating Functions

For the most part we have considered generating functions to be clotheslines on which to hang discrete probability densities and combinatorial sequences. In addition to this purely formal role, generating functions are also often analytic functions of a complex variable. This second perspective forges deep connections to classical mathematical analysis and helps in deriving asymptotic expansions and bounds.

Recall that an analytic function $f(x)$ can be expanded in a power series $f(x) = \sum_{n=0}^\infty a_n(x-y)^n$ around every point y in its domain, an open subset of the complex plane. For our purposes, the point $y = 0$ is the most pertinent. Convergence of the series expansion of $f(x)$ is absolute in the open disc around y of radius

$$
R = \frac{1}{\limsup_{n\to\infty}|a_n|^{1/n}};
$$

note that the value $R = \infty$ is possible. If R is finite, then one or more singularities must occur on the circle $\{x : |x-y| = R\}$. These are points where $f(x)$ lacks a derivative or blows up in some sense. It follows from the definition of R that whenever $r > R$, the strict inequality $|a_n|r^n > 1$ holds for infinitely many n. Hence, the series defining $f(x)$ locally around $y = 0$ cannot converge at the real number r. On the other hand, if $r < R$, then $|a_n|r^n < 1$ for all but finitely many n. In fact, if $r < s < R$, then $|a_n|r^n < (r/s)^n$ for all but finitely many n. It follows that $f(x)$ converges absolutely and uniformly within the disc $\{x : |x| \le r\}$ with derivative $f'(x) = \sum_{n=1}^\infty n a_n x^n$.

The most important asymptotic lesson learned from this exposition is that the radius of convergence around the origin bounds the rate of growth of the coefficients in the expansion of $f(x)$. To repeat, if $r < R$, then

$|a_n| \leq r^{-n}$ for all but finitely many n. If $r > R$, then $|a_n| \geq r^{-n}$ for infinitely many n.

Example 12.5.1 *Asymptotics of the Fibonacci Numbers*

The Fibonacci sequence f_n discussed in Example 4.2.4 is determined by the two initial values $f_0 = f_1 = 1$ and the recurrence $f_{n+1} = f_n + f_{n-1}$. Let $F(x) = \sum_{n=0}^{\infty} f_n x^n$ denote the corresponding generating function. Multiplying the recurrence by x^{n+1} and adding over the index set $n = 1, 2, \ldots$ produces the equation

$$F(x) - x - 1 \quad = \quad x[F(x) - 1] + x^2 F(x)$$

with solution

$$F(x) \quad = \quad \frac{1}{1 - x - x^2}.$$

The singularities of $F(x)$ occur at the roots

$$r_{\pm} \quad = \quad \frac{-1 \pm \sqrt{5}}{2}$$

of the quadratic $x^2 + x - 1 = 0$. The singularity r_+ is closer to the origin. Hence, the coefficient f_n is $O(r^{-n})$ for any $r < |r_+| = r_+$. In fact, the partial fraction decomposition

$$F(x) \quad = \quad \frac{1}{r_+ - r_-}\Big(\frac{1}{x - r_-} - \frac{1}{x - r_+}\Big)$$

implies

$$f_n \quad = \quad \frac{1}{\sqrt{5}}\Big(r_+^{-n-1} - r_-^{-n-1}\Big) \quad \asymp \quad \frac{1}{\sqrt{5}} r_+^{-n-1}.$$

Thus, the order of magnitude and the asymptotic expression agree. ■

Partial fraction decompositions can be generalized to more complicated rational functions. Let $f(x)$ equal the ratio of two polynomials $p(x)$ and $q(x)$, with $p(x)$ of lower degree than $q(x)$. If r is a root of $q(x)$ of order d, then there is a polynomial $t(x)$ such that $q(x) = (x - r)^d t(x)$. Furthermore for any constant a,

$$f(x) \quad = \quad \frac{p(x)}{q(x)} \quad = \quad \frac{a}{(x - r)^d} + \frac{p(x) - at(x)}{(x - r)^d t(x)}. \tag{12.13}$$

Choosing

$$a \quad = \quad \frac{p(r)}{t(r)} \quad = \quad \lim_{x \to r}(x - r)^d f(x) \quad = \quad \frac{d!\, p(r)}{q^{(d)}(r)} \tag{12.14}$$

makes the numerator $p(x) - at(x)$ of the second fraction on the right of equation (12.13) vanish at r. Hence, it can be written as $(x-r)p_1(x)$ for a polynomial $p_1(x)$ of lower degree than the degree of $(x-r)^{d-1}t(x)$. Nothing prevents us from attacking the remainder

$$f_1(x) \quad = \quad f(x) - \frac{a}{(x-r)^d} \quad = \quad \frac{p_1(x)}{(x-r)^{d-1}t(x)}$$

in exactly the same manner. The complete expansion

$$f(x) \quad = \quad \sum_j \sum_{k=1}^{d_j} \frac{a_{jk}}{(x-r_j)^k}$$

over all roots r_j of $q(x)$ exhausts the rational function because the degree of the denominator keeps dropping, and the degree of the numerator always trails the degree of the denominator. Here we assume that $q(x)$ resolves into a product $\prod_j (x-r_j)^{d_j}$. If $q(x)$ has leading coefficient 1, then the fundamental theorem of algebra guarantees such a factorization.

Example 12.5.2 *Making Change*

In a certain country there are m different types of coins with relative values $d_1, \ldots, d_m$. For the sake of simplicity, assume that integers $d_1, \ldots, d_m$ have greatest common divisor 1. For instance, the value d_1 could be 1. Let c_n denote the number of ways of paying a bill of amount n. Thus, c_n counts the number of vectors $(s_1, \ldots, s_m)$ with nonnegative integer entries such that $n = \sum_{j=1}^m s_j d_j$. Because $(1-x)^{-1} = \sum_{j=0}^\infty x^j$, the sequence c_n has generating function

$$C(x) \quad = \quad \sum_{n=0}^\infty c_n x^n \quad = \quad \frac{1}{(1-x^{d_1})(1-x^{d_2})\cdots(1-x^{d_m})}.$$

The identity

$$1 - x^d \quad = \quad (1-x)(1+x+\cdots+x^{d-1})$$

implies that 1 is a root of the denominator $q(x)$ of $C(x)$ of multiplicity m. It is clear that every other root r of $q(x)$ coincides with a root of unity $e^{2\pi ik/s}$, where $i = \sqrt{-1}$ and s and k are relatively prime positive integers. The factor $1 - x^{d_j}$ vanishes at r if and only if s divides $d_j k$. Because s and k are relatively prime, this condition holds if and only if s divides d_j. Because the $d_1, \ldots, d_m$ are relatively prime, the multiplicity of r is accordingly strictly less than m.

The dominant contribution to c_n in the partial fraction decomposition

$$C(x) \quad = \quad \sum_j \sum_{k=1}^{d_j} \frac{a_{jk}}{(x-r_j)^k}$$

$$\begin{aligned} &= \sum_j \sum_{k=1}^{d_j} \frac{(-1)^k a_{jk}}{r_j^k \left(1 - \frac{x}{r_j}\right)^k} \\ &= \sum_j \sum_{k=1}^{d_j} \frac{(-1)^k a_{jk}}{r_j^k} \sum_{n=0}^{\infty} \binom{n+k-1}{n} \left(\frac{x}{r_j}\right)^n \end{aligned}$$

is clearly identifiable in this problem. All of the roots r_j have absolute value 1, so everything hinges on the magnitude of

$$\binom{n+k-1}{n} \asymp \frac{n^{k-1}}{(k-1)!}.$$

The index k assumes its greatest value m for the choice $r = 1$. Hence, c_n satisfies

$$c_n \asymp a(-1)^m \frac{n^{k-1}}{(k-1)!},$$

where according to equation (12.14)

$$a = \lim_{x \to 1} (x-1)^m \frac{1}{(1-x^{d_1})(1-x^{d_2}) \cdots (1-x^{d_m})} = (-1)^m \prod_{j=1}^{m} \frac{1}{d_j}.$$

The lovely asymptotic formula

$$c_n \asymp \prod_{j=1}^{m} \frac{1}{d_j} \cdot \frac{n^{k-1}}{(k-1)!}$$

of Schur implies that it is possible to make change for every large bill even when $d_j = 1$ is not among the available coins. ■

12.6 Stochastic Forms of Convergence

Probabilists entertain many notions of convergence [24, 208]. The simplest of these is the pointwise convergence of a sequence of random variables X_n to a limit X. Because this form of convergence is allowed to fail on an event of probability 0, it is termed almost sure convergence. The usual calculus rules for dealing with limits apply to almost surely convergent sequences. The most celebrated almost sure convergence result is the strong law of large numbers. We refer readers to Example 10.3.1 for the statement and proof of one version of the strong law of large numbers.

More generally, a sequence X_n is said to converge to X in probability if for every $\epsilon > 0$

$$\lim_{n \to \infty} \Pr(|X_n - X| > \epsilon) = 0.$$

One can prove that X_n converges to X in probability if and only if every subsequence X_{n_m} of X_n possesses a subsubsequence $X_{n_{m_l}}$ converging to X almost surely. (See Problem 31.) This characterization makes it possible to generalize nearly all theorems involving almost surely convergent sequences to theorems involving sequences converging in probability. For instance, the dominated convergence theorem generalizes in this fashion. Convergence in probability can be deduced from the mean square convergence condition

$$\lim_{n\to\infty} \mathrm{E}(|X_n - X|^2) \;=\; 0$$

via Chebyshev's inequality, which is reviewed in Chapter 3.

Convergence in distribution is a very different matter. The underlying random variables need not even live on the same probability space. What is crucial is that the distribution functions $F_n(x)$ of the X_n converge to the distribution function $F(x)$ of X at every point of continuity x of $F(x)$. In other words, the way that the X_n attribute mass comes more and more to resemble the way X attributes mass. Besides this intuitive definition, there are other equivalent definitions useful in various contexts.

Proposition 12.6.1 *The following statements about the sequence X_n and its potential limit X are equivalent:*

(a) The sequence of distribution functions $F_n(x)$ converges to $F(x)$ at every continuity point x of $F(x)$.

(b) The sequence of characteristic functions $\mathrm{E}(e^{isX_n})$ converges to the characteristic function $\mathrm{E}(e^{isX})$ for every real number s.

(c) The sequence of expectations $\mathrm{E}[f(X_n)]$ converges to $\mathrm{E}[f(X)]$ for every bounded continuous function $f(x)$ defined on the real line.

(d) There exist random variables Y_n and Y defined on a common probability space such that Y_n has distribution function $F_n(x)$, Y has distribution function $F(x)$, and Y_n converges to Y almost surely.

The deep Skorokhod representation theorem mentioned in property (d) brings us full circle to almost sure convergence. It also enables us to prove some results with surprising ease. For instance, any continuous function $g(X_n)$ of a sequence converging in distribution also converges in distribution. This result follows trivially from applying $g(x)$ to the almost surely converging sequence Y_n. For another example, suppose $f(x)$ is a nonnegative continuous function. Then Fatou's lemma implies

$$\begin{aligned} \mathrm{E}[f(X)] &= \mathrm{E}[f(Y)] \\ &\le \liminf_{n\to\infty} \mathrm{E}[f(Y_n)] \\ &= \liminf_{n\to\infty} \mathrm{E}[f(X_n)]. \end{aligned}$$

The choices $f(x) = |x|$ and $f(x) = x^{2n}$ are the most important in practice.

Example 12.6.1 *Binary Expansions*

Let $X_1, X_2, \ldots$ be an i.i.d. sequence of Bernoulli random variables with success probability $\frac{1}{2}$. It is obvious that the finite sums $S_n = \sum_{i=1}^n 2^{-i} X_i$ converge almost surely to the infinite sum $S = \sum_{i=1}^\infty 2^{-i} X_i$. If we interpret X_i as the ith binary digit of the random number S [192], then S should be uniformly distributed on $[0, 1]$. To make this insight rigorous, let us prove that S_n converges in distribution to the uniform distribution. According to part (c) of Proposition 12.6.1, it suffices to prove that

$$\lim_{n\to\infty} \mathrm{E}[f(S_n)] \quad = \quad \int_0^1 f(x)dx$$

for every bounded continuous function $f(x)$. Now a little reflection shows that

$$E[f(S_n)] \quad = \quad \frac{1}{2^n} \sum_{m=0}^{2^n-1} f\left(\frac{m}{2^n}\right).$$

But the Riemann sum on the right of this equality converges to the integral $\int_0^1 f(x)dx$. ■

Doubtless the reader is already familiar with the central limit theorem. The simplest version deals with a sequence X_k of i.i.d. random variables with common mean μ and common variance σ^2. In this setting the normalized sums

$$Z_n \quad = \quad \frac{1}{\sqrt{n\sigma^2}} \sum_{k=1}^n (X_k - \mu)$$

converge in distribution to a standard normal deviate. These hypotheses can be weakened in various ways. The next proposition of Lindeberg represents the most general version of the central limit theorem retaining the assumption of independent summands.

Proposition 12.6.2 *Consider a sequence X_k of independent random variables with means $\mu_k = \mathrm{E}(X_k)$ and variances $\sigma_k^2 = \mathrm{Var}(X_k)$. Suppose that $s_n^2 = \sum_{k=1}^n \sigma_k^2$ tends to ∞ and that*

$$\lim_{n\to\infty} \frac{1}{s_n^2} \sum_{k=1}^n \mathrm{E}[1_{\{|X_k-\mu_k|>\epsilon s_n\}}(X_k - \mu_k)^2] \quad = \quad 0 \qquad (12.15)$$

for every $\epsilon > 0$. If $S_n = \sum_{k=1}^n X_k$ and $m_n = \sum_{k=1}^n \mu_k$, then the normalized sums

$$Z_n \quad = \quad \frac{1}{s_n}\Big(S_n - m_n\Big)$$

converge in distribution to a standard normal deviate.

Lindeberg's condition (12.15) is trivial to check in the classical central limit theorem. It is also obvious if the random variables X_n are uniformly bounded. Here are two combinatorial examples of the central limit theorem featured by Feller [59].

Example 12.6.2 *A Central Limit Theorem for Permutation Cycles*

Section 4.6 introduced Stirling numbers of the first kind. These count the permutations of $\{1, \ldots, n\}$ with various numbers of cycles. In discussing this material, we found there that the total number of cycles S_n in a random permutation could be represented as the sum $S_n = X_1 + \cdots + X_n$ of independent Bernoulli random variables X_k with success probabilities k^{-1}. It follows that $\mu_k = k^{-1}$, $\sigma_k^2 = k^{-1}(1 - k^{-1})$, and that

$$\begin{aligned} m_n &= \sum_{k=1}^{n} \frac{1}{k} \asymp \ln n + \gamma \\ s_n^2 &= \sum_{k=1}^{n} \frac{1}{k}\left(1 - \frac{1}{k}\right) \asymp \ln n + \gamma - \frac{\pi^2}{6}, \end{aligned}$$

where γ is Euler's constant. Because the random variables X_k are uniformly bounded, Proposition 12.6.2 applies. Thus, $(S_n - m_n)/s_n$ is approximately standard normal for large n. ■

Example 12.6.3 *A Central Limit Theorem for Permutation Inversions*

Two values π_i and π_j of a permutation π constitute an inversion if $i < j$ and $\pi_i > \pi_j$. Let S_n count the number of inversions of a random permutation of $\{1, \ldots, n\}$. For example, $S_6 = 8$ for the permutation $\pi = (3, 6, 1, 5, 2, 4)$ of $\{1, \ldots, 6\}$ because the numbers 3, 6, 1, 5, 2, and 4 induce two inversions, four inversions, no inversions, two inversions, no inversions, and no inversions, respectively. The inversion count X_k induced by the number k depends solely on the relative order of the numbers $1, \ldots, k$ in the permutation. Hence, X_k is uniformly distributed between 0 and $k - 1$ and independent of $X_1, \ldots, X_{k-1}$. The ingredients of the central limit theorem are $\mu_k = (k-1)/2$, $\sigma_k^2 = (k^2 - 1)/12$, and

$$\begin{aligned} m_n &= \frac{1}{2}\sum_{k=1}^{n}(k-1) = \frac{n(n-1)}{4} \asymp \frac{n^2}{4} \\ s_n^2 &= \frac{1}{12}\sum_{k=1}^{n}(k^2-1) = \frac{2n^3 + 3n^2 - 5n}{72} \asymp \frac{n^3}{36}. \end{aligned}$$

For any $\epsilon > 0$, the condition $|X_k - \mu_k| > \epsilon s_n$ fails for all $k \leq n$ when n is sufficiently large because $X_k - \mu_k$ is $O(n)$ and $s_n \asymp n^{3/2}/6$. Hence, Lindeberg's condition (12.15) is valid, and the central limit theorem holds for the sequence $(S_n - m_n)/s_n$. ■

12.7 Problems

1. Prove the following order relations:

   a) $1 - \cos^2 x \;=\; O(x^2)$ as $x \to 0$

   b) $\ln x \;=\; o(x^\alpha)$ as $x \to \infty$ for any $\alpha > 0$

   c) $\frac{x^2}{1+x^3} + \ln(1+x^2) \;=\; O(x^2)$ as $x \to 0$

   d) $\frac{x^2}{1+x^3} + \ln(1+x^2) \;=\; O(\ln x)$ as $x \to \infty$.

2. Show that $f(x) \;\asymp\; g(x)$ as $x \to x_0$ does not entail the stronger relation $e^{f(x)} \;\asymp\; e^{g(x)}$ as $x \to x_0$. However, argue that the condition $f(x) \;=\; g(x) + o(1)$ is sufficient to imply $e^{f(x)} \;\asymp\; e^{g(x)}$.

3. For two positive functions $f(x)$ and $g(x)$, prove that $f(x) \;\asymp\; g(x)$ as $x \to x_0$ implies $\ln f(x) \;=\; \ln g(x) + o(1)$ as $x \to x_0$. Hence, $\lim_{x\to x_0} \ln f(x) \;\neq\; 0$ entails $\ln f(x) \;\asymp\; \ln g(x)$ as $x \to x_0$.

4. Demonstrate that

$$\left(1 + \frac{1}{\sqrt{x}}\right)^x \;\asymp\; e^{\sqrt{x}-1/2}$$

   as $x \to \infty$.

5. Find an asymptotic expansion for $\int_x^\infty e^{-y^4} dy$ as $x \to \infty$.

6. Suppose that $0 < c < \infty$ and that $f(x)$ is bounded and continuous on $[0, c]$. If $f(c) \neq 0$, then show that

$$\int_0^c x^n f(x) dx \;\asymp\; \frac{c^{n+1}}{n} f(c)$$

   as $n \to \infty$.

7. Let $F(x)$ be a distribution function concentrated on $[0, \infty)$ with moments $m_k \;=\; \int_0^\infty y^k dF(y)$. For $x \geq 0$ define the Stieltjes function $f(x) \;=\; \int_0^\infty \frac{1}{1+xy} dF(y)$. Show that $\sum_{k=0}^\infty (-1)^k m_k x^k$ is an asymptotic expansion for $f(x)$ satisfying

$$f(x) - \sum_{k=0}^{n} (-1)^k m_k x^k \;=\; (-x)^{n+1} \int_0^\infty \frac{y^{n+1}}{1+xy} dF(y).$$

   Argue, therefore, that the remainders of the expansion alternate in sign and are bounded in absolute value by the first omitted term.

8. Show that $\int_0^\infty \frac{e^{-y}}{1+xy} dy \;\asymp\; \frac{\ln x}{x}$ as $x \to \infty$. (Hints: Write

$$\int_0^\infty \frac{e^{-y}}{1+xy} dy \;=\; \frac{1}{x} \int_0^\infty \frac{d}{dy} \ln(1+xy) e^{-y} dy,$$

and use integration by parts and the dominated convergence theorem.)

9. Prove that

$$\begin{aligned}\int_0^{\frac{\pi}{2}} e^{-x\tan y}dy &\asymp \frac{1}{x}\\ \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} (y+2)e^{-x\cos y}dy &\asymp \frac{4}{x}\end{aligned}$$

as $x \to \infty$.

10. For $0 < \lambda < 1$, demonstrate the asymptotic equivalence

$$\sum_{k=0}^{n} \binom{n}{k} k! n^{-k} \lambda^k \asymp \frac{1}{1-\lambda}$$

as $n \to \infty$. (Hint: Use the identity $k!n^{-k-1} = \int_0^\infty y^k e^{-ny} dy$.)

11. Demonstrate the asymptotic equivalence

$$\sum_{k=0}^{n} \binom{n}{k} k! n^{-k} \asymp \sqrt{\frac{\pi n}{2}}$$

as $n \to \infty$. (Hint: See Problem 10.)

12. The von Mises density

$$\frac{e^{\kappa\cos(y-\alpha)}}{2\pi I_0(\kappa)}, \qquad -\pi < y \le \pi,$$

is used to model random variation on a circle. Here α is a location parameter, $\kappa > 0$ is a concentration parameter, and the modified Bessel function $I_0(\kappa)$ is the normalizing constant

$$I_0(\kappa) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{\kappa\cos y} dy.$$

Verify that Laplace's method yields

$$I_0(\kappa) \asymp \frac{e^\kappa}{\sqrt{2\pi\kappa}}$$

as $\kappa \to \infty$. For large κ it is clear that the von Mises distribution is approximately normal.

13. Suppose the continuous function $f(t)$ on $[0,\infty)$ is $O(e^{ct})$ for some $c \geq 0$ as $t \to \infty$. For t positive use Laplace's method to prove Post's inversion formula [161]

$$f(t) = \lim_{k\to\infty} \frac{(-1)^k}{k!}\left(\frac{k}{t}\right)^{k+1} \tilde{f}^{(k)}\left(\frac{k}{t}\right)$$

for the Laplace transform $\tilde{f}(s)$ of $f(t)$. Here $\tilde{f}^{(k)}(s)$ is the kth derivative of the transform. (Hint: Consult Section 12.3.1.)

14. Let $\phi(x)$ and $\Phi(x)$ be the standard normal density and distribution functions. Demonstrate the bounds

$$\frac{x}{1+x^2}\phi(x) \leq 1-\Phi(x) \leq \frac{1}{x}\phi(x)$$

for $x > 0$. (Hints: Exploit the derivatives

$$\begin{aligned}\frac{d}{dx}e^{-x^2/2} &= -xe^{-x^2/2}\\ \frac{d}{dx}\left(x^{-1}e^{-x^2/2}\right) &= -\left(1+x^{-2}\right)e^{-x^2/2}\end{aligned}$$

and simple inequalities for the integral $1-\Phi(x)$.)

15. Prove the elementary inequalities

$$\ln n! - \frac{\ln n}{2} \leq \int_1^n \ln t\, dt = n\ln n - n + 1 \leq \ln n!$$

that point the way to Stirling's formula. (Hint: Using the concavity of $\ln t$, verify the inequality

$$\frac{\ln(m-1)+\ln m}{2} \leq \int_{m-1}^m \ln t\, dt\,.\ )$$

16. In the socks in the laundry problem, demonstrate that

$$\mathrm{E}(N_1) = \frac{(2^n n!)^2}{(2n)!}.$$

Conclude from this and Stirling's formula that $\mathrm{E}(N_1) \asymp \sqrt{\pi n}$. (Hint: Change variables in the first integral of equation (12.8).)

17. Let $f(x)$ be a periodic function on the real line whose kth derivative is piecewise continuous for some positive integer k. Show that the Fourier coefficients c_n of $f(x)$ satisfy

$$|c_n| \leq \frac{\int_0^1 |f^{(k)}(x)|dx}{|2\pi n|^k}$$

for $n \neq 0$.

18. Suppose that the periodic function $f(x)$ is square-integrable on $[0,1]$. Prove the assertions: (a) $f(x)$ is an even (respectively odd) function if and only if its Fourier coefficients c_n are even (respectively odd) functions of n, (b) $f(x)$ is real and even if and only if the c_n are real and even, and (c) $f(x)$ is even (odd) if and only if it is even (odd) around $1/2$. By even around $1/2$ we mean $f(1/2+x) = f(1/2-x)$.

19. Demonstrate that

$$\frac{\pi^2}{12} = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k^2}, \qquad \frac{\pi^4}{90} = \sum_{k=1}^{\infty} \frac{1}{k^4}.$$

20. Show that the even Bernoulli numbers can be expressed as

$$B_{2n} = (-1)^{n+1} \frac{2(2n)!}{(2\pi)^{2n}} \left[1 + \frac{1}{2^{2n}} + \frac{1}{3^{2n}} + \frac{1}{4^{2n}} + \cdots\right].$$

Apply Stirling's formula, and deduce the asymptotic relation

$$|B_{2n}| \asymp 4\sqrt{\pi n}\left(\frac{n}{\pi e}\right)^{2n}.$$

21. Show that the Bernoulli polynomials satisfy the identity

$$B_n(x) = (-1)^n B_n(1-x)$$

for all n and $x \in [0,1]$. Conclude from this identity that $B_n(1/2) = 0$ for n odd.

22. Continuing Problem 21, show inductively for $n \geq 1$ that $B_{2n}(x)$ has exactly one simple zero in $(0,1/2)$ and one in $(1/2,1)$, while $B_{2n+1}(x)$ has precisely the simple zeros 0, $1/2$, and 1.

23. Demonstrate that the Bernoulli polynomials satisfy the identity

$$B_n(x+1) - B_n(x) = nx^{n-1}.$$

Use this result to verify that the sum of the nth powers of the first m integers can be expressed as

$$\sum_{k=1}^{m} k^n = \frac{1}{n+1}\Big[B_{n+1}(m+1) - B_{n+1}(0)\Big].$$

(Hint: Prove the first assertion by induction or by expanding $B_n(x)$ in a Taylor series around the point 1.)

24. As an alternative definition of the Bernoulli polynomials, consider the bivariate exponential generating function

$$f(t,x) = \frac{te^{tx}}{e^t-1} = \sum_{n=0}^{\infty}\frac{B_n(x)}{n!}t^n.$$

Check the defining conditions (A.10) of the Bernoulli polynomials by evaluating $\lim_{t\to 0} f(t,x)$, $\int_0^1 f(t,x)\,dx$, and $\frac{\partial}{\partial x}f(t,x)$. Hence, the coefficients $B_n(x)$ are the Bernoulli polynomials. The special case

$$f(t,0) = \frac{t}{e^t-1} = \sum_{n=0}^{\infty}\frac{B_n}{n!}t^n \tag{12.16}$$

gives the exponential generating function of the Bernoulli numbers.

25. Verify the identities

$$t\coth t = \frac{2t}{e^{2t}-1}+t = \sum_{n=0}^{\infty}\frac{4^nB_{2n}}{(2n)!}t^{2n}$$

using equation (12.16). Show that this in turn implies

$$t\cot t = \sum_{n=0}^{\infty}\frac{(-1)^n4^nB_{2n}}{(2n)!}t^{2n}.$$

The Bernoulli numbers also figure in the Taylor expansion of $\tan t$ [78].

26. Verify the asymptotic expansion

$$\sum_{k=1}^{n}k^{\alpha} = C_\alpha + \frac{n^{\alpha+1}}{\alpha+1}+\frac{n^\alpha}{2}+\sum_{j=1}^{m}\frac{B_{2j}}{2j}\binom{\alpha}{2j-1}n^{\alpha-2j+1} + O(n^{\alpha-2m-1})$$

for a real number $\alpha \neq -1$ and some constant C_α, which you need not determine.

27. Find asymptotic expansions for the two sums $\sum_{k=1}^n (n^2+k^2)^{-1}$ and $\sum_{k=1}^n (-1)^k/k$ valid to $O(n^{-3})$.

28. Suppose $f(x)$ is continuously differentiable and monotone on the interval $[m,n]$. Prove that

$$\left|\sum_{k=m}^{n}f(k)-\int_m^n f(x)\,dx-\frac{1}{2}[f(m)+f(n)]\right| \leq \frac{1}{2}|f(n)-f(m)|$$

and that

$$\left| \sum_{k=m}^{n} f(k) - \int_m^n f(x)\,dx \right| \quad \leq \quad \max\{|f(m)|, |f(n)|\}.$$

If in addition $\int_m^\infty f(x)\,dx$ is finite, then show that

$$\left| \sum_{k=m}^{\infty} f(k) - \int_m^\infty f(x)\,dx - \frac{1}{2}f(m) \right| \quad \leq \quad \frac{1}{2}|f(m)|.$$

Use the second of these inequalities to prove that Euler's constant $\gamma = \lim_{n\to\infty}(\sum_{k=1}^{n} \frac{1}{k} - \ln n)$ exists. (Hint: In the spirit of Proposition 12.4.1, expand $\sum_{k=m}^{n} f(k)$ so that the remainder involves an integral of $f'(x)$.)

29. The function

$$f(x) \quad = \quad \frac{e^x}{1-x} \quad = \quad \sum_{n=0}^{\infty} a_n x^n$$

has a pole (singularity) at $x = 1$. Show that $a_n = e + O(r^n)$ for every $r > 0$ based on this fact. (Hint: Consider $f(x) - e/(1-x)$.)

30. Let q_n be the probability that in n tosses of a fair coin there are no occurrences of the pattern HHH [59]. Derive the recurrence relation

$$q_n \quad = \quad \frac{1}{2}q_{n-1} + \frac{1}{4}q_{n-2} + \frac{1}{8}q_{n-3}$$

for $n \geq 3$ and use it to calculate the generating function

$$Q(s) \quad = \quad \sum_{n=0}^{\infty} q_n s^n \quad = \quad \frac{2s^2 + 4s + 8}{8 - 4s - 2s^2 - s^3}.$$

Show numerically that the denominator has the real root

$$r_1 \quad = \quad 1.0873778$$

and two complex roots r_2 and r_3 satisfying $|r_2| > r_1$ and $|r_3| > r_1$. Deduce the asymptotic relation $q_n \asymp c r_1^{-n-1}$ for the constant

$$c \quad = \quad 1.236840.$$

(Hints: The n trials produce no HHH run only if they begin with T, HT, or HHT. The inequality

$$|4s + 2s^2 + s^3| \quad < \quad 4r_1 + 2r_1^2 + r_1^3 \quad = \quad 8$$

holds for all $|s| \leq r_1$ except $s = r_1$. Finally, apply equations (12.13) and (12.14) to find c.)

31. Suppose a sequence of random variables X_n converges to X in probability. Use the Borel-Cantelli lemma to show that some subsequence X_{n_m} converges to X almost surely. Now prove the full claim that X_n converges to X in probability if and only if every subsequence X_{n_m} of X_n possesses a subsubsequence $X_{n_{m_l}}$ converging to X almost surely.

32. Suppose X_n converges in probability to X and Y_n converges in probability to Y. Show that X_n+Y_n converges in probability to $X+Y$, that X_nY_n converges in probability to XY, and that X_n/Y_n converges in probability to X/Y when $Y \neq 0$ almost surely.

33. Consider two sequences X_n and Y_n of random variables. Suppose that X_n converges in distribution to the random variable X and the difference $X_n - Y_n$ converges in probability to the constant 0. Prove that Y_n converges in distribution to X. (Hints: For one possible proof, invoke part (b) of Proposition 12.6.1. The inequality

$$\begin{aligned} |e^{isb} - e^{isa}| &= \left| is \int_0^{b-a} e^{ist}\, dt \right| \\ &\leq |s(b-a)| \end{aligned}$$

will come in handy.)

34. Let $X_1, X_2, \ldots$ and $Y_1, Y_2, \ldots$ be two independent i.i.d. sequences of Bernoulli random variables with success probability $\frac{1}{2}$. Show that the random variable $Z_n = 2Y_n + X_n$ is uniformly distributed over the set $\{0, 1, 2, 3\}$, that $Z = \sum_{i=1}^{\infty} 4^{-i} Z_i$ is uniformly distributed over the interval $[0, 1]$, and that

$$\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{\infty} 2^{-i} X_i \\ \sum_{i=1}^{\infty} 2^{-i} Y_i \end{pmatrix}$$

is uniformly distributed over the square $[0, 1] \times [0, 1]$. Finally, demonstrate that the map $Z \mapsto (X, Y)^t$ preserves probability in the sense that every interval $[i4^{-n}, (i+1)4^{-n}]$ of length 4^{-n} with $0 \leq i < 4^n$ is sent into a square of area 4^{-n}. Note that the Borel-Cantelli lemma implies that Z almost surely determines its digits Z_n.

35. As a follow-up to Problems 10 and 11 of Chapter 5, let A_n be the number of ascents in a random permutation of $\{1, \ldots, n\}$. Demonstrate that A_n has the same distribution as $\lfloor U_1 + \cdots + U_n \rfloor$, where $\lfloor z \rfloor$ is the integer part of z and $U_1, \ldots, U_n$ are independent random variables with uniform distribution on $[0, 1]$. Use Problem 33 and the central limit theorem to show that $(A_n - n/2)/\sqrt{n/12}$ has an approximate standard normal distribution [195]. (Hints: The distribution of A_n is the same as the distribution of $D_n = n - 1 - A_n$, the number of descents. Now let $V_j = (U_1 + \cdots + U_j) \bmod 1$. Demonstrate that

$V_1, \ldots, V_n$ are independent random variables with uniform distribution on $[0, 1]$ and that $\lfloor U_1 + \cdots + U_n \rfloor$ provides the number of descents in the sequence $V_1, \ldots, V_n$. Finally, argue that the number of descents in the sequence $V_1, \ldots, V_n$ is distributed as D_n.)

13
Numerical Methods

13.1 Introduction

Stochastic modeling relies on a combination of exact solutions and numerical methods. As scientists and engineers tackle more realistic models, the symmetries supporting exact solutions fade. The current chapter sketches a few of the most promising numerical techniques. Further improvements in computing, statistics, and data management are bound to drive the rapidly growing and disorganized discipline of computational probability for decades to come.

Our first vignette stresses iterative methods for finding the equilibrium distribution π of a discrete-time Markov chain. The alternative algebraic methods such as Gaussian elimination solve the balance equation for π in a single step. However, Gaussian elimination requires on the order of m^3 arithmetic operations for m states. With a large, sparse transition probability matrix, a good iterative method can be much faster. The block versions of the Jacobi and Gauss-Seidel algorithms fall into this category.

We have already encountered the finite Fourier transform in our discussion of branching processes. We revisit the subject in the current chapter. Fresh examples from renewal theory and jump counting in continuous-time Markov chains reinforce the importance of Fourier analysis in dealing with discrete distributions. Although classical probability theory depends heavily on Fourier analysis, its emergence as a numerical tool has been slow.

Simulation is now thoroughly entrenched in statistics, particularly in Bayesian analyses. The development of MCMC techniques has gone hand

K. Lange, *Applied Probability,*
DOI 10.1007/978-1-4419-7165-4_13, © Springer Science+Business Media, LLC 2010

in hand with advances in Markov chain theory. However, simulation as a modeling tool rather than a statistical tool has lagged. The introduction of τ-leaping methods discussed in continuous-time Markov chains is therefore a welcome development. We summarize a step anticipation improvement that renders τ-leaping even more effective. Even if simulation is eventually replaced by exact methods in a given stochastic model, a good simulation tool is indispensable in rapidly implementing the model and suggesting mathematical conjectures.

Our final topic shows how one can bridge the gap between an exact discrete-state process and its diffusion approximation. Among other things, the methods applied lead to a better understanding of the phenomenon of extinction in the Wright-Fisher genetics model. How much of this analysis carries over to other models is unclear, but the problem is generic. Before moving on to specifics, let us recommend to readers the books [12, 149, 191] for further study.

13.2 Computation of Equilibrium Distributions

Suppose π is the equilibrium distribution of the ergodic transition probability matrix $P = (p_{ij})$. One way of using Gaussian elimination to solve for π is to rewrite the vector balance equation $\pi = \pi P$ as the system

$$\sum_{i=1}^{m} \pi_i (1_{\{i=j\}} - p_{ij}) \quad = \quad 0.$$

Adding the quantity $p_{mj} \sum_{i=1}^{m} \pi_i = p_{mj}$ to both sides produces the equivalent system

$$\sum_{i=1}^{m} \pi_i (1_{\{i=j\}} - p_{ij} + p_{mj}) \quad = \quad p_{mj},$$

whose first $m-1$ equations do not depend on π_m [153]. Hence, if the truncated $m-1 \times m-1$ matrix

$$R \quad = \quad (1_{\{i=j\}} - p_{ij} + p_{mj}) \tag{13.1}$$

is invertible, then one can solve for the first $m-1$ entries of π by Gaussian elimination and set $\pi_m = 1 - \sum_{i=1}^{m-1} \pi_i$. Problem 1 asks the reader to check that R^{-1} exists.

The power method is the simplest iterative method of computing the equilibrium distribution π of a discrete-time Markov chain. Under the hypothesis of ergodicity, the iterates $\pi^{(n+1)} = \pi^{(n)} P$ are guaranteed to converge to π from any starting distribution $\pi^{(0)}$. For a reversible ergodic chain, Proposition 7.5.1 shows that the total variation distance from $\pi^{(n)}$ to π is

$O(\rho^n)$, where $\rho < 1$ is the absolute value of the second-largest eigenvalue in magnitude of P. For a nonreversible chain, the corresponding bound is $O(n^k\rho^n)$ for some integer $k \geq 0$. The extra factor of n^k adjusts for the possible failure of P to be diagonalizable [169]. If ρ is close to 1, then the power method can be very slow to converge. One can accelerate convergence by computing P^2, P^4, P^8, and so forth by repeated squaring and waiting for the rows of the corresponding matrices to stabilize. The proof of Proposition 7.4.1 shows that

$$\min_i p_{ij}^{(2^n)} \quad \leq \quad \pi_j \quad \leq \quad \max_i p_{ij}^{(2^n)},$$

where $p_{ij}^{(2^n)}$ is a typical entry of P^{2^n}. Unfortunately, when there are m states, each squaring requires on the order of $O(m^3)$ arithmetic operations. Furthermore, any initial sparsity of P is lost in repeated squaring. This puts repeated squaring in the same league as Gaussian elimination. In its defense, repeated squaring suffers less from roundoff error.

The classical iterative schemes of Jacobi and Gauss and Seidel start from the rearrangement

$$\pi_j \quad = \quad \frac{1}{1-p_{jj}} \sum_{i \neq j} \pi_i p_{ij}$$

of the equilibrium equation. The Jacobi updates

$$\pi_j^{(n+1)} \quad = \quad \frac{1}{1-p_{jj}} \sum_{i \neq j} \pi_i^{(n)} p_{ij}$$

can be computed in parallel; in contrast, the Gauss-Seidel updates

$$\pi_j^{(n+1)} \quad = \quad \frac{1}{1-p_{jj}} \left(\sum_{i<j} \pi_i^{(n+1)} p_{ij} + \sum_{i>j} \pi_i^{(n)} p_{ij} \right)$$

must be computed sequentially from the first state 1 to the last state m. Because the Gauss-Seidel algorithm uses each update as soon as it is generated, it typically converges in fewer iterations than the Jacobi algorithm.

Table 13.1 compares the performance of the four iterative algorithms on a simple random walk problem with $m = 50$ states. Here all entries p_{ij} of the transition probability matrix P are 0 except for $p_{12} = p_{m,m-1} = 1$ and

$$p_{k,k-1} \quad = \quad \frac{1}{2\sqrt{k}}, \quad p_{kk} \quad = \quad \frac{1}{2}, \quad p_{k,k+1} \quad = \quad \frac{1}{2}\Big(1 - \frac{1}{\sqrt{k}}\Big)$$

for $2 \leq k \leq m$. Most of the probability mass piles up near state m. Starting from the uniform distribution, the table compares the largest entry of the nth iterate $\pi^{(n)}$ to the largest entry of π. In the case of squaring, we choose the largest entry in the first row of the appropriate power of P. The power

method is the slowest to converge, followed by the Jacobi algorithm, the Gauss-Seidel algorithm, and repeated squaring, in that order. The true value for the maximum coincides with the converged value under repeated squaring. As a safeguard for each of the first three methods, the vector $\pi^{(n)}$ is replaced by the normalized vector $(\pi^{(n)}\mathbf{1})^{-1}\pi^{(n)}$ as soon as all entries are computed. Likewise, the rows of P^2, P^4, and so forth are normalized under repeated squaring. The initial distribution $\pi^{(0)}$ is random.

TABLE 13.1. Comparison of Various Equilibrium-Seeking Algorithms

n	Power	Jacobi	Gauss-Seidel	Squaring
0	0.020	0.020	0.020	1.000
10	0.073	0.091	0.219	0.614
20	0.116	0.162	0.386	0.614
40	0.199	0.304	0.514	0.614
60	0.277	0.428	0.563	0.614
80	0.350	0.505	0.589	0.614
100	0.415	0.543	0.601	0.614
120	0.470	0.565	0.608	0.614
140	0.511	0.580	0.611	0.614
160	0.539	0.590	0.612	0.614
180	0.558	0.597	0.613	0.614
200	0.570	0.602	0.613	0.614

For very large systems, block Gauss-Seidel is one of the more competitive algorithms. Here we divide π and $Q = I - P$ into compatible contiguous blocks π_J and Q_{IJ}. The balance equation $\pi Q = \mathbf{0}^t$ then becomes the system of equations

$$\pi_J \;=\; -\left(\sum_{I \neq J} \pi_I Q_{IJ}\right) Q_{JJ}^{-1}. \tag{13.2}$$

The reasoning employed in Section 7.6 demonstrates that the inverse Q_{JJ}^{-1} exists. To turn equation (13.2) into an algorithm, we label the index sets $I_1, \ldots, I_b$ and cycle through the block updates

$$\pi_{I_k}^{(n+1)} \;=\; -\left(\sum_{j=1}^{k-1} \pi_{I_j}^{(n+1)} Q_{I_j I_k} + \sum_{j=k+1}^{b} \pi_{I_j}^{(n)} Q_{I_j I_k}\right) Q_{I_k I_k}^{-1} \tag{13.3}$$

in sequence. The block Gauss-Seidel algorithm (13.3) converges in fewer iterations than the univariate Gauss-Seidel algorithm because no approximations are made within a block. Precomputing and storing the inverse matrices $Q_{I_k I_k}^{-1}$ is advisable. As Problem 2 shows, all entries of the updated block π_{I_k} remain nonnegative. The article [193] highlights necessary

and sufficient conditions for convergence. The corresponding block Jacobi algorithm

$$\pi_{I_k}^{(n+1)} \;=\; -\Big(\sum_{j\neq k}\pi_{I_j}^{(n)}Q_{I_jI_k}\Big)Q_{I_kI_k}^{-1}$$

is inherently parallel but tends to take more iterations to converge. The book [191] discusses other methods for computing π.

Finding the equilibrium distribution of a continuous-time Markov chain reduces to finding the equilibrium distribution of an associated discrete-time Markov chain. Consider a continuous-time chain with transition intensities λ_{ij} and infinitesimal generator Λ. If we collect the off-diagonal entries of Λ into a matrix Ω and the negative diagonal entries into a diagonal matrix D, then equation (8.7) describing the balance conditions satisfied by the equilibrium distribution π can be recast as

$$\pi D \;=\; \pi\Omega.$$

Close examination of the matrix $P = D^{-1}\Omega$ shows that its entries are nonnegative, its row sums are 1, and its diagonal entries are 0. Furthermore, P is sparse whenever Ω is sparse, and all states communicate under P when all states communicate under Λ. Nothing prevents the transition probability matrix P from being periodic, but aperiodicity is irrelevant in deciding whether a unique equilibrium distribution exists. Indeed, for any fixed constant $\alpha \in (0,1)$, one can easily demonstrate that an equilibrium distribution of P is also an equilibrium distribution of the aperiodic transition probability matrix $Q = \alpha I + (1-\alpha)P$ and vice versa.

Suppose that we compute the equilibrium distribution ν of P by some method. Once ν is available, we set $\omega = \nu D^{-1}$. Because the two equations

$$\nu \;=\; \nu P, \qquad \omega D \;=\; \omega\Omega$$

are equivalent, ω coincides with π up to a multiplicative constant. In other words, Λ has equilibrium distribution $\pi = (\omega\mathbf{1})^{-1}\omega$. Hence, trivial adjustment of the equilibrium distribution for the associated discrete-time chain produces the equilibrium distribution of the original continuous-time chain.

13.3 Applications of the Finite Fourier Transform

As mentioned in Section 9.5, the finite Fourier transform can furnish approximations to the Fourier coefficients of a periodic function $f(x)$. We now explore these ideas more systematically. Appendices A.3 and A.5 provide a brief overview of the elementary theory. Here we focus on applications.

Our first application of the finite Fourier transform is to computing the convolution of two sequences c_j and d_j of finite length n. Close inspection of Proposition A.3.1 suggests the following procedure. Compute the transforms $\hat{c}_k$ and $\hat{d}_k$ via the fast Fourier transform, multiply pointwise to form the product transform $n\hat{c}_k\hat{d}_k$, and then invert the product transform via the fast inverse Fourier transform. This procedure requires on the order of $O(n \ln n)$ operations, whereas the naive evaluation of a convolution requires on the order of n^2 operations unless one of the sequences is sparse. Here is a prime example where fast convolution is useful.

Example 13.3.1 *Multiplication of Generating Functions*

One can write the generating function $R(s)$ of the sum $X + Y$ of two independent, nonnegative, integer-valued random variables X and Y as the product $R(s) = P(s)Q(s)$ of the generating function $P(s) = \sum_{j=0}^{\infty} p_j s^j$ of X and the generating function $Q(s) = \sum_{j=0}^{\infty} q_j s^j$ of Y. The coefficients of $R(s)$ are given by the convolution formula

$$r_k = \sum_{j=0}^{k} p_j q_{k-j}.$$

Assuming that the p_j and q_j are 0 or negligible for $j \geq m$, we can view the two sequences as having period $n = 2m$ provided we set $p_j = q_j = 0$ for $j = m, \ldots, n-1$. Introducing these extra 0's makes it possible to write

$$r_k = \sum_{j=0}^{n-1} p_j q_{k-j} \tag{13.4}$$

without embarrassment. The r_j returned by the suggested procedure are correct in the range $0 \leq j \leq m-1$. Clearly, the same process works if $P(s)$ and $Q(s)$ are arbitrary polynomials of degree $m-1$ or less. ■

We now turn to infinite sequences and Fourier series approximation. If $f(x)$ is a function defined on the real line with period 1, then its kth Fourier series coefficient c_k can be approximated by

$$c_k = \int_0^1 f(x)e^{-2\pi ikx}dx \approx \frac{1}{n}\sum_{j=0}^{n-1} f\Big(\frac{j}{n}\Big)e^{-2\pi i\frac{jk}{n}} = \hat{b}_k,$$

where $i = \sqrt{-1}$, $b_j = f(j/n)$, n is some large positive integer, and $\hat{b}_k$ is the finite Fourier transform of the sequence b_j evaluated at the integer k. Because the transformed values $\hat{b}_k$ are periodic, only n of them are distinct, say $\hat{b}_{-n/2}$ through $\hat{b}_{n/2-1}$ for n even.

An important question is how well $\hat{b}_k$ approximates c_k. To assess the error, suppose that $\sum_k |c_k| < \infty$ and that the Fourier series of $f(x)$ converges

to $f(x)$ at the points j/n for $j = 0, \ldots, n-1$. The calculation

$$\begin{aligned}
\hat{b}_k &= \frac{1}{n}\sum_{j=0}^{n-1} e^{-2\pi i \frac{jk}{n}} f\left(\frac{j}{n}\right) \\
&= \frac{1}{n}\sum_{j=0}^{n-1} e^{-2\pi i \frac{jk}{n}} \sum_m c_m e^{2\pi i \frac{jm}{n}} \\
&= \sum_m c_m \frac{1}{n}\sum_{j=0}^{n-1} e^{2\pi i \frac{j(m-k)}{n}} \\
&= \sum_m c_m \begin{cases} 1 & m = k \bmod n \\ 0 & m \neq k \bmod n \end{cases}
\end{aligned}$$

implies that

$$\hat{b}_k - c_k = \sum_{l \neq 0} c_{ln+k}. \tag{13.5}$$

If the Fourier coefficients c_j decline sufficiently rapidly to 0 as $|j|$ tends to ∞, then the error $\hat{b}_k - c_k$ will be small for $-n/2 \leq k \leq n/2 - 1$. Problems 11, 12, and 13 explore this question in more depth.

Example 13.3.2 *Fast Solution of a Renewal Equation*

The discrete renewal equation

$$u_n = a_n + \sum_{m=0}^{n} f_m u_{n-m} \tag{13.6}$$

arises in many applications of probability theory [59]. Here a_n and f_n are known bounded sequences with $f_0 = 0$. Beginning with the initial value $u_0 = a_0$, it takes on the order of n^2 operations to compute $u_0, \ldots, u_n$ recursively via the convolution equation (13.6). Alternatively, if we multiply both sides of (13.6) by s^n and sum on n, then we get the equation

$$U(s) = A(s) + F(s)U(s), \tag{13.7}$$

involving the generating functions

$$U(s) = \sum_{n=0}^{\infty} u_n s^n, \quad A(s) = \sum_{n=0}^{\infty} a_n s^n, \quad F(s) = \sum_{n=0}^{\infty} f_n s^n.$$

The explicit form of the solution

$$U(s) = \frac{A(s)}{1 - F(s)}$$

suggests that we can recover the coefficients u_n of $U(s)$ by the Fourier series method. For this tactic to work, the condition $\lim_{n\to\infty} u_n = 0$ must hold.

One of the aims of renewal theory is to clarify the nature of recurrent events. In Section 7.3.4 we introduced the first passage distribution f_n characterizing the number of epochs until the next occurrence of a recurrent event. If we let u_n be the probability of the recurrent event at epoch n, then we can write the renewal equation

$$u_n = 1_{\{n=0\}} + \sum_{m=0}^{n} f_m u_{n-m}.$$

Because the process starts with an occurrence, the choice $A(s) = 1$ is appropriate. In this setting the solution

$$U(s) = \frac{1}{1 - F(s)}$$

has a singularity at the point $s = 1$. If this is the only singularity on the unit circle, then we can adopt the strategy of Section 12.5 and try to show that the coefficients u_n tend to a nontrivial limit c. Fortunately, there is a simple necessary and sufficient condition for $s = 1$ to be the only root of the equation $F(s) = 1$ on the unit circle $|s| = 1$. Problem 15 asks the reader to prove that the relevant condition requires the set $\{n: f_n > 0\}$ to have greatest common divisor 1.

These observations suggest that it would be better to estimate the coefficients $v_n = u_n - c$ of the generating function

$$V(s) = U(s) - \frac{c}{(1-s)}$$

for the choice

$$c = \lim_{s\to 1}(1-s)U(s) = \lim_{s\to 1}\frac{1-s}{1-F(s)} = \frac{1}{\mu},$$

where μ is the mean recurrence time $F'(1)$. Provided $F(s)$ satisfies the greatest common divisor hypothesis, we can now recover the better-behaved coefficients v_n by the approximate Fourier series method. The advantage of this oblique attack on the problem is that it takes on the order of only $n \ln n$ operations to compute $v_0, \ldots, v_n$ and hence $u_0, \ldots, u_n$. Finally, note that Section 7.3.4 reaches the same conclusion $\lim_{n\to\infty} u_n = \mu^{-1}$ by invoking the ergodic theorem. To avoid such strong machinery here, we must assume that $F(s)$ has a radius of convergence around the origin strictly greater than 1.

As a concrete illustration of the proposed method, consider a classical coin tossing problem. Let u_n be the probability of a new run of r heads ending at toss n. In contrast to the pattern matching assumptions in Example

2.2.2, we adopt the simplifying convention that no segment of a current run of r heads can be counted as a segment of the next run of r heads. To calculate the generating function of the u_n, we take $u_0 = 1$ and set $u_1 = u_2 = \cdots = u_{r-1} = 0$. For $n \geq r$ and head probability $p = 1 - q$, the probability that all r trials $n - r + 1, \ldots, n - 1, n$ are heads satisfies

$$p^r = u_n + pu_{n-1} + \cdots + p^{r-1}u_{n-r+1}. \tag{13.8}$$

On the right-hand side of this identity, we segment the coin tossing history up to epoch n by the epoch of the last new run.

If we define

$$a_n = \begin{cases} p^n & 0 \leq n \leq r \\ p^r & n > r, \end{cases} \qquad g_n = \begin{cases} -p^n & 1 \leq n \leq r-1 \\ 0 & n = 0 \text{ or } n \geq r, \end{cases}$$

then equation (13.8) is a disguised form of the renewal equation (13.6) with g_n replacing f_n. The special cases $u_0 = p^0 = 1$ and $u_n = p^n - p^n = 0$ for $1 \leq n \leq r-1$ constitute the initial conditions. Straightforward calculations show that

$$1 - G(s) = 1 + ps + \cdots + (ps)^{r-1} = \frac{1 - (ps)^r}{1 - ps}$$

and

$$A(s) = \frac{1 - (ps)^{r+1}}{1 - ps} + \frac{p^r s^{r+1}}{1 - s} = \frac{1 - s + qp^r s^{r+1}}{(1 - s)(1 - ps)}.$$

It follows that

$$U(s) = \frac{A(s)}{1 - G(s)} = \frac{1 - s + qp^r s^{r+1}}{(1 - s)[1 - (ps)^r]}$$

and

$$c = \lim_{s \to 1}(1 - s)U(s) = \frac{qp^r}{1 - p^r}.$$

In this problem it is easier to calculate $U(s)$ directly and avoid the first passage probabilities altogether.

As a toy example, let us take $r = 2$ and $p = 1/2$. Fourier transforming $n = 32$ values of $V(s)$ on the boundary of the unit circle yields the renewal probabilities displayed in Table 13.2. In this example, convergence to the limiting value occurs so rapidly that the value of introducing the finite Fourier transform is debatable. Other renewal equations exhibit less rapid convergence. ■

TABLE 13.2. Renewal Probabilities in a Coin Tossing Example

n	u_n	n	u_n	n	u_n
0	1.0000	5	0.1563	10	0.1670
1	0.0000	6	0.1719	11	0.1665
2	0.2500	7	0.1641	12	0.1667
3	0.1250	8	0.1680	13	0.1666
4	0.1875	9	0.1660	∞	0.1667

13.4 Counting Jumps in a Markov Chain

Let X_t be a continuous-time Markov chain with n states, transition intensities λ_{ij}, and infinitesimal generator Λ. Recall that the ith diagonal entry of Λ is $-\lambda_i = -\sum_{j\neq i}\lambda_{ij}$. For a set of donor states A and a set of recipient states B, we wish to characterize the random number of jumps N_t from A to B during $[0,t]$ [121, 143]. As a form of shorthand, we write $k \to l$ whenever $k \in A$ and $l \in B$. Our attack on this problem proceeds via the expectations

$$f_{ij}(t,u) \;=\; \mathrm{E}(u^{N_t}1_{\{X_t=j\}} \mid X_0 = i)$$

for $|u| \leq 1$. Except for omission of the normalizing constant

$$\Pr(X_t = j \mid X_0 = i) \;=\; (e^{t\Lambda})_{ij},$$

the function $f_{ij}(t,u)$ serves as the generating function of N_t conditional on the events $X_0 = i$ and $X_t = j$. The sum $f_i(t,u) = \sum_j f_{ij}(t,u)$ represents the generating function of the number of jumps conditional only on the event $X_0 = i$. If we take $A = B = \{1,\ldots,n\}$, then we track all jumps. If we take $A = \{1,\ldots,n\}$ and $B = \{k\}$, then we track only jumps into state k. The opposite choice $A = \{k\}$ and $B = \{1,\ldots,n\}$ tracks jumps out of state k.

The convolution equation

$$f_{ij}(t,u) \;=\; e^{-\lambda_i t}1_{\{i=j\}} + \int_0^t e^{-\lambda_i s}\sum_{k\neq i}\lambda_{ik}u^{1_{\{i\to k\}}}f_{kj}(t-s,u)ds \quad (13.9)$$

is our departure point for calculating $f_{ij}(t,u)$. In equation (13.9) the term $e^{-\lambda_i t}1_{\{i=j\}}$ covers the event that the process never leaves state i, in which case $N_t = 0$ and $u^0 = 1$. The integral covers the possibility of the process leaving at some time $s \leq t$ and moving to the intermediate state k. The exponential factor $e^{-\lambda_i s}$ is the probability that the departure occurs after time s, and the factor $\lambda_{ik}ds$ is the infinitesimal probability that the chain jumps to state k during the time interval $(s, s+ds)$. Finally, the factor $u^{1_{\{i\to k\}}}f_{kj}(t-s,u)$ summarizes the expected value of $u^{N_t}1_{\{X_t=j\}}$ conditional on this jump.

Any convolution equation invites the Laplace transform. If we let

$$\hat{g}(\theta) = \int_0^\infty e^{-\theta t} g(t)\, dt$$

denote the Laplace transform of $g(t)$, then taking transforms of equation (13.9) produces

$$\hat{f}_{ij}(\theta, u) = \frac{1}{\theta + \lambda_i} 1_{\{i=j\}} + \frac{1}{\theta + \lambda_i} \sum_{k \neq i} \lambda_{ik} u^{1_{\{i \to k\}}} \hat{f}_{kj}(\theta, u). \quad (13.10)$$

To make further progress, we collect the transforms $\hat{f}_{ij}(\theta, u)$ into a matrix $\hat{F}(\theta, u)$ and write the system of equations (13.10) as the single matrix equation

$$\hat{F}(\theta, u) = D(\theta) + D(\theta) C(u) \hat{F}(\theta, u),$$

where $D(\theta)$ is the diagonal matrix with ith diagonal entry $(\theta + \lambda_i)^{-1}$ and $C(u)$ is the matrix with diagonal entries 0 and off-diagonal entries $\lambda_{ik} u^{1_{\{i \to k\}}}$.

Arranging things in this manner allows us to solve for $\hat{F}(\theta, u)$ in the form

$$\begin{aligned} \hat{F}(\theta, u) &= \{I - D(\theta) C(u)\}^{-1} D(\theta) \\ &= \{D^{-1}(\theta) - C(u)\}^{-1} \\ &= \{\theta I - \text{diag}(\Lambda) - C(u)\}^{-1} \\ &= (\theta I - M)^{-1}, \end{aligned} \quad (13.11)$$

where $M = \text{diag}(\Lambda) + C(u)$. The matrix $\theta I - M$ is invertible because it is strictly diagonally dominant. (See Problem 18.) Fortunately, we can invert the Laplace transform (13.11). Indeed, the fundamental theorem of calculus implies that

$$\begin{aligned} \int_0^\infty e^{-\theta t} e^{tM} dt &= \int_0^\infty e^{-t(\theta I - M)} dt \\ &= -(\theta I - M)^{-1} e^{-t(\theta I - M)} \Big|_0^\infty \\ &= (\theta I - M)^{-1}. \end{aligned}$$

It follows that

$$F(t, u) = e^{tM} = e^{t[\text{diag}(\Lambda) + C(u)]}. \quad (13.12)$$

Note that $F(t, u)$ satisfies the necessary initial condition $F(0, u) = I$.

If all $\lambda_i = \lambda$ are equal, then $\text{diag}(\Lambda)$ is a multiple of the identity matrix, and

$$F(t, u) = e^{-\lambda t} e^{tC(u)} = \sum_{l=0}^\infty e^{-\lambda t} \frac{t^l}{l!} C(u)^l.$$

If in addition $A = B = \{1, \ldots, n\}$, then $C(u) = u[\Lambda - \text{diag}(\Lambda)]$, and we can equate the joint probability $\Pr(N_t = l,\ X_t = j)$ to the entry in row i and column j of the matrix $e^{-\lambda t}\frac{t^l}{l!}[\Lambda - \text{diag}(\Lambda)]^l$.

In general, we can always compute the joint probabilities by extending the matrix-valued function $F(t,u) = \sum_{l=0}^{\infty} u^l F_l(t)$ to the unit circle and extracting its Fourier coefficient

$$F_l(t) \;=\; \int_0^1 F(t, e^{2\pi i\theta}) e^{-2\pi i l\theta} d\theta.$$

As usual, we approximate the integral in question by a Riemann sum and calculate the Riemann sum entry by entry by applying the fast Fourier transform. This procedure is obviously dependent on accurate evaluation of the matrix exponential (13.12).

We can also recover the unnormalized moments of the random variables N_t by evaluating the derivatives of the solution (13.12) with respect to u at $u = 1$. For example, $E_t = \frac{\partial}{\partial u}F(t,u)|_{u=1}$ equals the matrix of expected jumps up to a normalizing constant. The function E_t satisfies the differential equation

$$\begin{aligned}
\frac{d}{dt}E_t &= \frac{\partial}{\partial u}\frac{\partial}{\partial t}F(t,u)|_{u=1} \\
&= \frac{\partial}{\partial u}\Big\{[\text{diag}(\Lambda) + C(u)]e^{t[\text{diag}(\Lambda)+C(u)]}\Big\}\Big|_{u=1} \\
&= [C(1) - C(0)]e^{t\Lambda} + \Lambda E_t
\end{aligned}$$

because $\frac{d}{du}C(u) = C(1) - C(0)$ and $\text{diag}(\Lambda) + C(1) = \Lambda$. The solution of this system is

$$\begin{aligned}
E_t &= e^{t\Lambda}\int_0^t e^{-s\Lambda}[C(1) - C(0)]e^{s\Lambda} ds \\
&= \int_0^t e^{(t-s)\Lambda}[C(1) - C(0)]e^{s\Lambda} ds
\end{aligned}$$

subject to the initial condition $E(0) = \mathbf{0}$. For the mean number of jumps of any type, $C(1) - C(0) = \Lambda - \text{diag}(\Lambda)$. If in addition all λ_i coincide, then Λ and $\text{diag}(\Lambda)$ commute, and E_t reduces to $te^{t\Lambda}[\Lambda - \text{diag}(\Lambda)]$.

More generally, suppose SDS^{-1} is a diagonalization of Λ. If we define $R = S^{-1}[C(1) - C(0)]S$, then

$$\begin{aligned}
E_t &= S\int_0^1 e^{(t-s)D} R e^{sD} ds\, S^{-1} \\
&= S\int_0^t \Big[e^{(t-s)d_i} r_{ij} e^{sd_j}\Big] ds\, S^{-1} \\
&= S(q_{ij})S^{-1},
\end{aligned}$$

where r_{ij} is a typical entry of R, d_i is a typical diagonal entry of D, and

$$q_{ij} = \begin{cases} \frac{r_{ij}}{d_j - d_i}(e^{td_j} - e^{td_i}) & d_i \neq d_j \\ te^{td_i} r_{ij} & d_i = d_j. \end{cases}$$

This result is a special case of a generic formula for the derivative of a matrix exponential with respect to a parameter [176].

As an illustration of these methods in action, consider Kimura's nucleotide evolution chain as sketched in Example 8.5.2 and Section 8.6. If we set $\alpha = 1$, $\beta = 2$, and let the chain progress to time $t = \frac{1}{2}$, then Table 13.3 shows the expected number of jumps conditional on all possible starting states i and ending states j. Each value in the table represents the ratio of an entry of E_t to the corresponding entry of $e^{t\Lambda}$. Here we have computed E_t exactly by the formulas $te^{t\Lambda}[\Lambda - \text{diag}(\Lambda)]$ and $S(q_{ij})S^{-1}$ and numerically by $\sum_l l \Pr(N_t = l,\ X_t = j \mid X_0 = i)$. All three methods give the same result.

TABLE 13.3. Expected Number of Nucleotide Substitutions in Kimura's Chain

	$j = 1$	$j = 2$	$j = 3$	$j = 4$
$i = 1$	2.1672	2.7454	2.5746	2.5746
$i = 2$	2.7454	2.1672	2.5746	2.5746
$i = 3$	2.5746	2.5746	2.1672	2.7454
$i = 4$	2.5746	2.5746	2.7454	2.1672

13.5 Stochastic Simulation and Intensity Leaping

Many chemical and biological models depend on continuous-time Markov chains with a finite number of particle types [91]. The particles interact via a finite number of reaction channels, and each reaction destroys and/or creates particles in a predictable way. In this section, we consider the problem of simulating the behavior of such chains. Before we launch into simulation specifics, it is helpful to carefully define a typical process. If d denotes the number of types, then the chain follows the count vector X_t whose ith component X_{ti} is the number of particles of type i at time $t \geq 0$. We typically start the system at time 0 and let it evolve via a succession of random reactions. Let c denote the number of reaction channels. Channel j is characterized by an intensity function $r_j(x)$ depending on the current vector of counts x. In a small time interval of length s, we expect $r_j(x)s + o(s)$ reactions of type j to occur. Reaction j changes the count vector by a fixed integer vector v^j. Some components v^j_k of v^j may be positive, some 0, and

some negative. From the wait and jump perspective of Markov chain theory, we wait an exponential length of time until the next reaction. If the chain is currently in state $X_t = x$, then the intensity of the waiting time is $r_0(x) = \sum_{j=1}^{c} r_j(x)$. Once the decision to jump is made, we jump to the neighboring state $x + v^j$ with probability $r_j(x)/r_0(x)$.

Table 13.4 lists typical reactions, their intensities $r(x)$, and increment vectors v. In the table, S_i denotes a single particle of type i. Only the nonzero increments v_i are shown. The reaction intensities invoke the law of mass action and depend on rate constants a_i. Each discipline has its own vocabulary. Chemists use the name propensity instead of the name intensity and call the increment vector a stoichiometric vector. Physicists prefer creation to immigration. Biologists speak of death and mutation rather than of decay and isomerization. Despite the variety of processes covered, the allowed chains form a subset of all continuous-time Markov chains. Chains with an infinite number of reaction channels or random increments are not allowed. For instance, many continuous-time branching processes do not qualify. Branching processes that grow by budding serve as useful substitutes for more general branching processes.

TABLE 13.4. Some Examples of Reaction Channels

Name	**Reaction**	$r(x)$	v
Immigration	$0 \to S_1$	a_1	$v_1 = 1$
Decay	$S_1 \to 0$	$a_2 x_1$	$v_1 = -1$
Dimerization	$S_1 + S_1 \to S_2$	$a_3 \binom{x_1}{2}$	$v_1 = -2$, $v_2 = 1$
Isomerization	$S_1 \to S_2$	$a_4 x_1$	$v_1 = -1$, $v_2 = 1$
Dissociation	$S_2 \to S_1 + S_1$	$a_5 x_2$	$v_1 = 2$, $v_2 = -1$
Budding	$S_1 \to S_1 + S_2$	$a_6 x_1$	$v_2 = 1$
Replacement	$S_1 + S_2 \to S_2 + S_2$	$a_7 x_1 x_2$	$v_1 = -1$, $v_2 = 1$
Complex Reaction	$S_1 + S_2 \to S_3 + S_4$	$a_8 x_1 x_2$	$v_1 = v_2 = -1$ $v_3 = v_4 = 1$

The wait and jump mechanism constitutes a perfectly valid method of simulating one of these chains. Gillespie first recognized the practicality of this approach in chemical kinetics [74]. Although his stochastic simulation algorithm works well in some contexts, it can be excruciatingly slow in others. Unfortunately, reaction rates can vary by orders of magnitude, and the fastest reactions dominate computational expense in stochastic simulation. For the fast reactions, stochastic simulation takes far too many small steps. Our goal is to describe an alternative approximate algorithm that takes larger, less frequent steps. The alternative is predicated on the observation that reaction intensities change rather slowly in many models. Before describing how we can take advantage of this feature, it is worth

mentioning the chemical master equations, which is just another name for the forward equations of Markov chain theory.

Let $p_{xy}(t)$ denote the finite-time transition probability of going from state x at time 0 to state y at time t. The usual reasoning leads to the expansions

$$p_{xy}(t+s) = p_{xy}(t)\left[1-\sum_{j=1}^{c} r_j(y)s\right]+\sum_{j=1}^{c} p_{x,y-v^j}(t)r_j(y-v^j)s+o(s).$$

Forming the corresponding difference quotient and sending s to 0 produce the master equations

$$\frac{d}{dt}p_{xy}(t) = \sum_{j=1}^{c}\left[p_{x,y-v^j}(t)r_j(y-v^j)-p_{xy}(t)r_j(y)\right]$$

with initial conditions $p_{xy}(0) = 1_{\{x=y\}}$. Only in special cases can the master equations be solved. In deterministic models where particle counts are high, one is usually content to follow mean particle counts. The mean behavior $\mu(t) = \mathrm{E}(X_t)$ is then roughly modeled by the system of ordinary differential equations

$$\frac{d}{dt}\mu(t) = \sum_{j=1}^{c} r_j[\mu(t)]v^j.$$

This approximation becomes more accurate as mean particle counts increase.

For the sake of argument, suppose all reaction intensities are constant. In the time interval $(t, t+s)$, reaction j occurs a Poisson number of times with mean $r_j s$. If we can sample from the Poisson distribution with an arbitrary mean, then we can run stochastic simulation accurately with s of any duration. If we start the process at $X_t = x$, then at time $t+s$ we have $X_{t+s} = x + \sum_{j=1}^{c} N_j v^j$, where the N_j are independent Poisson variates with means $r_j s$. The catch, of course, is that reaction intensities change as the particle count vector X_t changes. In the τ-leaping method of simulation, we restrict the time increment $\tau > 0$ to sufficiently small values such that each intensity $r_j(x)$ suffers little change over $(t, t+\tau)$ [33, 75].

Before we discuss exactly how to achieve this, let us pass to a more sophisticated update that anticipates how intensities change [178]. Assume that X_t is a deterministic process with a well-defined derivative. Over a short time interval $(t, t+\tau)$, the intensity $r_j(X_t)$ should then change by the approximate amount $\frac{d}{dt}r_j(X_t)\tau$. Reactions of type j now occur according to an inhomogeneous Poisson process with a linear intensity. Thus, we anticipate a Poisson number of reactions of type j with mean

$$\omega_j(t, t+\tau) = \int_0^\tau \left[r_j(X_t)+\frac{d}{dt}r_j(X_t)s\right]ds$$

$$= \quad r_j(X_t)\tau + \frac{d}{dt}r_j(X_t)\frac{1}{2}\tau^2.$$

At time $t+\tau$, we put $X_{t+\tau} = X_t + \sum_{j=1}^{c} N_j v^j$, where the N_j are independent Poisson variates with means $\omega_j(t, t+\tau)$. This is all to the good, but how do we compute the time derivatives of $r_j(X_t)$? The most natural approach is to invoke the chain rule

$$\frac{d}{dt}r_j(x) \quad = \quad \sum_{k=1}^{d} \frac{\partial}{\partial x_k} r_j(x) \frac{d}{dt} x_k$$

and set

$$\frac{d}{dt}x_k \quad = \quad \sum_{j=1}^{c} r_j(x) v_k^j$$

as dictated by the approximate mean growth of the system. In most models the matrix $dr(x) = [\frac{\partial}{\partial x_k} r_j(x)]$ is sparse, with nontrivial entries that are constant or linear in x.

This exposition gives some insight into how we choose the increment τ in the τ-leaping method. It seems reasonable to take the largest τ such that

$$\left|\frac{d}{dt}r_j(x)\right|\tau \quad \leq \quad \epsilon r_j(x)$$

holds for all j, where $\epsilon > 0$ is a small constant. If $r_j(x) = 0$ is possible, then we might amend this to

$$\left|\frac{d}{dt}r_j(x)\right|\tau \quad \leq \quad \epsilon \max\{r_j(x), a_j\},$$

where a_j is the rate constant for reaction j. In each instance in Table 13.4, a_j is the smallest possible change in $r_j(x)$. In common with other τ-leaping strategies, we revert to the stochastic simulation update whenever the intensity $r_0(x)$ for leaving a state x falls below a certain threshold δ.

As a test case, we apply the above version of τ-leaping to Kendall's birth, death, and immigration process. In the time-homogeneous case, this Markov chain is governed by the birth rate α per particle, the death rate μ per particle, and the overall immigration rate ν. Equation (8.17) provides the mean number of particles $m_i(t)$ at time t starting with i particles at time 0. This exact expression permits us to evaluate the accuracy of τ-leaping. For the sake of illustration, we consider $t = 4$, $i = 5$, and average particle counts over 10,000 simulations. Table 13.5 lists the exact value of $m_5(4)$ and the average particle counts from τ-leaping for two methods. Method 1 ignores the derivative correction, and method 2 incorporates it. The table also gives for method 2 the time in seconds over all 10,000 runs and the fraction of steps attributable to stochastic simulation (SSA) when

TABLE 13.5. Mean Counts for $\alpha = 2$, $\mu = 1$, and $\nu = \frac{1}{2}$ in Kendall's Process

ϵ	Average 1	Average 2	Predicted	Time	SSA Fraction
1.0000	153.155	251.129	299.790	0.781	0.971
0.5000	195.280	279.999	299.790	0.875	0.950
0.2500	232.495	292.790	299.790	1.141	0.909
0.1250	261.197	297.003	299.790	1.625	0.839
0.0625	279.176	301.671	299.790	2.328	0.726
0.0313	286.901	298.565	299.790	3.422	0.575
0.0156	297.321	301.560	299.790	4.922	0.405
0.0078	294.487	300.818	299.790	7.484	0.256

the threshold constant $\delta = 100$. Although the table makes a clear case for the more accurate method 2, more testing is necessary. This is an active area of research, and given its practical importance, even more research is warranted.

13.6 A Numerical Method for Diffusion Processes

It is straightforward to simulate a diffusion process X_t. The definition tells us to extend X_t to X_{t+s} by setting the increment $X_{t+s} - X_t$ equal to a normal deviate with mean $\mu(t,x)s$ and variance $\sigma^2(t,x)s$. The time increment s should be small, and each sampled normal variate should be independent. Techniques for generating random normal deviates are covered in standard texts on computational statistics and will not be discussed here [112, 122]. Of more concern is how to cope with a diffusion process with finite range I. Because a normally distributed random variable has infinite range, it is possible in principle to generate an increment that takes the simulated process outside I. One remedy for this problem is to take s extremely small. It also helps if the infinitesimal variance $\sigma^2(t,x)$ tends to 0 as x approaches the boundary of I. This is the case with the neutral Wright-Fisher process.

Simulation offers a crude method of finding the distribution of X_t. Simply conduct multiple independent simulations and compute a histogram of the recorded values of X_t. Although this method is neither particularly accurate nor efficient, it has the virtue of yielding simultaneously the distributions of all of the X_t involved in the simulation process. Thus, if 1000 times are sampled per simulation, then the method yields all 1000 distributions, assuming that enough computer memory is available. Much greater accuracy can be achieved by solving Kolmogorov's forward equation. The ideal of an exact solution is seldom attained in practice, even for time-homogeneous problems. However, Kolmogorov's forward equation can be solved numerically by standard techniques for partial differential equations. Here we would like to discuss a nonstandard method for finding

the distribution of X_t that directly exploits the definition of a diffusion process.

This method recursively computes the distribution of X_{t_i} at n time points labeled $0 < t_1 < \cdots < t_n = t$. In the diffusion approximation to Markov chain models such as the Wright-Fisher model, it is reasonable to let $\delta t_i = t_{i+1} - t_i$ be one generation. It is also convenient to supplement these points with the initial point $t_0 = 0$. For each t_i, we would like to compute the probability that $X_{t_i} \in [a_{ij}, a_{i,j+1}]$ for r_i+1 points $a_{i0} < \cdots < a_{i,r_i}$. We will say more about these mesh points later. In the meantime let p_{ij} denote the probability $\Pr(X_{t_i} \in [a_{ij}, a_{i,j+1}])$ and c_{ij} the center of probability $\mathrm{E}(X_{t_i} \mid X_{t_i} \in [a_{ij}, a_{i,j+1}])$. Our method carries forward approximations to both of these sequences starting from an arbitrary initial distribution for X_0.

In passing from time t_i to time t_{i+1}, the diffusion process redistributes a certain amount of probability mass from the interval $[a_{ij}, a_{i,j+1}]$ to the interval $[a_{i+1,k}, a_{i+1,k+1}]$. Given the definition of a diffusion process and the notation $m(i,x) = x + \mu(t_i,x)\delta t_i$ and $s^2(i,x) = \sigma^2(t_i,x)\delta t_i$, the amount redistributed is approximately

$$\begin{aligned} & p_{ij\to i+1,k} \\ &= \int_{a_{ij}}^{a_{i,j+1}} \frac{1}{\sqrt{2\pi s^2(i,x)}} \int_{a_{i+1,k}}^{a_{i+1,k+1}} e^{-\frac{[y-m(i,x)]^2}{2s^2(i,x)}} dy f(t_i,x)\, dx \\ &= \int_{a_{ij}}^{a_{i,j+1}} \frac{1}{\sqrt{2\pi}} \int_{\frac{a_{i+1,k}-m(i,x)}{s(i,x)}}^{\frac{a_{i+1,k+1}-m(i,x)}{s(i,x)}} e^{-\frac{z^2}{2}} dz f(t_i,x)\, dx. \end{aligned} \tag{13.13}$$

(Here and in the remainder of this section, the equality sign indicates approximate equality.) Similarly, the center of probability $c_{ij\to i+1,k}$ of the redistributed probability approximately satisfies

$$\begin{aligned} & c_{ij\to i+1,k} p_{ij\to i+1,k} \\ &= \int_{a_{ij}}^{a_{i,j+1}} \frac{1}{\sqrt{2\pi s^2(i,x)}} \int_{a_{i+1,k}}^{a_{i+1,k+1}} y e^{-\frac{[y-m(i,x)]^2}{2s^2(i,x)}} dy f(t_i,x)\, dx \qquad (13.14) \\ &= \int_{a_{ij}}^{a_{i,j+1}} \frac{1}{\sqrt{2\pi}} \int_{\frac{a_{i+1,k}-m(i,x)}{s(i,x)}}^{\frac{a_{i+1,k+1}-m(i,x)}{s(i,x)}} [m(i,x) + s(i,x)z] e^{-\frac{z^2}{2}} dz f(t_i,x)\, dx. \end{aligned}$$

Given these quantities, we calculate

$$\begin{aligned} p_{i+1,k} &= \sum_{j=0}^{r_i-1} p_{ij\to i+1,k} \\ c_{i+1,k} &= \frac{1}{p_{i+1,k}} \sum_{j=0}^{r_i-1} c_{ij\to i+1,k} p_{ij\to i+1,k}\,, \end{aligned} \tag{13.15}$$

assuming that X_{t_i} is certain to belong to one of the intervals $[a_{ij}, a_{i,j+1}]$.

To carry out this updating scheme, we must approximate the integrals $p_{ij\to i+1,k}$ and $c_{ij\to i+1,k}p_{ij\to i+1,k}$. If the interval $[a_{ij}, a_{i,j+1}]$ is fairly narrow, then the linear approximations

$$\begin{aligned} m(i,x) &= \mu_{ij0} + \mu_{ij1}x \\ s^2(i,x) &= \sigma_{ij}^2 \\ f(t_i,x) &= f_{ij0} + f_{ij1}x \end{aligned} \tag{13.16}$$

should suffice for all x in the interval. The first two of these linear approximations follow directly from the diffusion model. The constants involved in the third approximation are determined by the equations

$$\begin{aligned} p_{ij} &= \int_{a_{ij}}^{a_{i,j+1}} (f_{ij0} + f_{ij1}x)\,dx \\ &= f_{ij0}(a_{i,j+1} - a_{ij}) + \frac{1}{2}f_{ij1}(a_{i,j+1} + a_{ij})(a_{i,j+1} - a_{ij}) \\ c_{ij}p_{ij} &= \int_{a_{ij}}^{a_{i,j+1}} x(f_{ij0} + f_{ij1}x)\,dx \\ &= \frac{1}{2}f_{ij0}(a_{i,j+1} + a_{ij})(a_{i,j+1} - a_{ij}) \\ &\quad + \frac{1}{3}f_{ij1}(a_{i,j+1}^2 + a_{i,j+1}a_{ij} + a_{ij}^2)(a_{i,j+1} - a_{ij}) \end{aligned}$$

with inverses

$$\begin{aligned} f_{ij0} &= \frac{2p_{ij}(2a_{ij}^2 + 2a_{ij}a_{i,j+1} + 2a_{i,j+1}^2 - 3a_{ij}c_{ij} - 3a_{i,j+1}c_{ij})}{(a_{i,j+1} - a_{ij})^3} \\ f_{ij1} &= \frac{6p_{ij}(2c_{ij} - a_{ij} - a_{i,j+1})}{(a_{i,j+1} - a_{ij})^3}. \end{aligned} \tag{13.17}$$

Problem 22 asks the reader to check that the linear density $f_{ij0} + f_{ij1}x$ is nonnegative throughout the interval $(a_{ij}, a_{i,j+1})$ if and only if its center of mass c_{ij} lies in the middle third of the interval.

Given the linear approximations (13.16), we now show that the double integrals (13.13) and (13.14) reduce to expressions involving elementary functions and the standard normal distribution function

$$\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-y^2/2}dy.$$

The latter can be rapidly evaluated by either a power series or a continued fraction expansion [122]. It also furnishes the key to evaluating the hierarchy of special functions

$$\Phi_k(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} y^k e^{-y^2/2}dy$$

through the integration-by-parts recurrence

$$\Phi_k(x) = -\frac{1}{\sqrt{2\pi}} x^{k-1} e^{-x^2/2} + (k-1)\Phi_{k-2}(x) \tag{13.18}$$

beginning with $\Phi_0(x) = \Phi(x)$. We can likewise evaluate the related integrals

$$\Psi_{jk}(x) = \int_{-\infty}^{x} y^j \Phi_k(y)\, dy$$

via the integration-by-parts reduction

$$\Psi_{jk}(x) = \frac{1}{j+1} x^{j+1} \Phi_k(x) - \frac{1}{j+1} \Phi_{j+k+1}(y). \tag{13.19}$$

Based on the definition of $\Phi(x)$, the integral (13.13) becomes

$$p_{ij\to i+1,k} = \int_{a_{ij}}^{a_{i,j+1}} \Phi\Big(\frac{z-\mu_{ij0}-\mu_{ij1}x}{\sigma_{ij}}\Big)\Big|_{a_{i+1,k}}^{a_{i+1,k+1}} (f_{ij0}+f_{ij1}x)\, dx,$$

and based on the recurrence (13.18), the integral (13.14) becomes

$$\begin{aligned} & c_{ij\to i+1,k} p_{ij\to i+1,k} \\ = & \int_{a_{ij}}^{a_{i,j+1}} (\mu_{ij0}+\mu_{ij1}x)\Phi\Big(\frac{z-\mu_{ij0}-\mu_{ij1}x}{\sigma_{ij}}\Big)\Big|_{a_{i+1,k}}^{a_{i+1,k+1}} (f_{ij0}+f_{ij1}x)\, dx \\ & - \frac{\sigma_{ij}}{\sqrt{2\pi}} \int_{a_{ij}}^{a_{i,j+1}} e^{-(z-\mu_{ij0}-\mu_{ij1}x)^2/(2\sigma_{ij}^2)}\Big|_{a_{i+1,k}}^{a_{i+1,k+1}} (f_{ij0}+f_{ij1}x)\, dx. \end{aligned}$$

To evaluate the one-dimensional integrals in these expressions for $p_{ij\to i+1,k}$ and $c_{ij\to i+1,k}p_{ij\to i+1,k}$, we make appropriate linear changes of variables so that $e^{-x^2/2}$ and $\Phi(x)$ appear in the integrands and then apply formulas (13.18) and (13.19) as needed. Although the details are messy, it is clear that these maneuvers reduce everything to combinations of elementary functions and the standard normal distribution function.

To summarize, the algorithm presented approximates the probability p_{ij} and center of probability c_{ij} of each interval $[a_{ij}, a_{i,j+1}]$ of a subdivision of the range I of X_t. Equation (13.17) converts these parameters into a piecewise linear approximation to the density of the process in preparation for propagation to the next subdivision. The actual propagation of probability from an interval of the current subdivision to another interval of the next subdivision is accomplished by computing $p_{ij\to i+1,k}$ and $c_{ij\to i+1,k}p_{ij\to i+1,k}$ based on elementary functions and the standard normal distribution function. The pieces $p_{ij\to i+1,k}$ and $c_{ij\to i+1,k}$ are then reassembled into probabilities and centers of probabilities using equations (13.15).

The choice of the mesh points $a_{i0} < \cdots < a_{i,r_i}$ at time t_i is governed by several considerations. First, the probability $\Pr(X_{t_i} \notin [a_{i0}, a_{i,r_i}])$ should be

negligible. Second, $\sigma^2(t_i, x)$ should be well approximated by a constant and $\mu(t_i, x)$ by a linear function on each interval $[a_{ij}, a_{i,j+1}]$. Third, the density $f(t_i, x)$ should be well approximated by a linear function on $[a_{ij}, a_{i,j+1}]$ as well. This last requirement is the hardest to satisfy in advance, but nothing prevents one from choosing the next subdivision adaptively based on the distribution of probability within the current subdivision. Adding more mesh points will improve accuracy at the expense of efficiency. Mesh points need not be uniformly spaced. It makes sense to cluster them in regions of high probability and rapid fluctuations of $f(t, x)$. Given the smoothness expected of $f(t, x)$, rapid fluctuations are unlikely.

Many of the probabilities $p_{ij\to i+1,k}$ are negligible. We can accelerate the algorithm by computing $p_{ij\to i+1,k}$ and $c_{ij\to i+1,k}$ only for $[a_{i+1,k}, a_{i+1,k+1}]$ close to $[a_{ij}, a_{i,j+1}]$. Because the conditional increment $X_{t_{i+1}} - X_{t_i}$ is normally distributed, it is very unlikely to extend beyond a few standard deviations σ_{ij} given X_{t_i} is in $[a_{ij}, a_{i,j+1}]$. Thus, the most sensible strategy is to visit each interval $[a_{ij}, a_{i,j+1}]$ in turn and propagate probability only to those intervals $[a_{i+1,k}, a_{i+1,k+1}]$ that lie a few standard deviations to the left or right of $[a_{ij}, a_{i,j+1}]$.

13.7 Application to the Wright-Fisher Process

We now apply the numerical method just described to the Wright-Fisher process. Our application confronts the general issue of how to deal with a diffusion approximation when it breaks down. In the case of the Wright-Fisher Markov chain, the diffusion approximation degrades for very low allele frequencies. Because of the interest in gene extinction, this is regrettable. However, in the regime of low allele frequencies, we can always fall back on the Wright-Fisher Markov chain. As population size grows, the Markov chain updates become more computationally demanding. The most pressing concern thus becomes how to merge the Markov chain and diffusion approaches seamlessly into a single algorithm for following the evolution of an allele. Here we present one possible algorithm and apply it to understanding disease-gene dynamics in a population isolate.

The algorithm outlined in the previous section has the virtue of being posed in terms of distribution functions rather than density functions. For low allele frequencies, discreteness is inevitable, and density functions are unrealistic. In adapting the algorithm to the regime of low allele frequencies, it is useful to let

$$a_{ij} = \frac{j - \frac{1}{2}}{2N_i}$$

for $0 \le j \le q$ and some positive integer q. The remaining a_{ij} are distributed over the interval $[a_{iq}, 1]$ less uniformly. This tactic separates the possibility

of exactly j alleles at time t_i, $0 \le j \le q$, from other possibilities. For $0 \le j \le q$, binomial sampling dictates that

$$\begin{aligned} p_{ij \to i+1,k} &= \sum_l \binom{2N_{i+1}}{l} p^l (1-p)^{2N_{i+1}-l} \\ c_{ij \to i+1,k} &= \frac{1}{p_{ij \to i+1,k}} \sum_l \binom{2N_{i+1}}{l} \frac{l}{2N_{i+1}} p^l (1-p)^{2N_{i+1}-l} \end{aligned}$$

where $p = m(i,x)$ is the gamete pool probability at frequency $x = j/(2N_i)$ and the sums occur over all l such that $l/(2N_{i+1}) \in [a_{i+1,k}, a_{i+1,k+1})$. When $0 \le k \le q$, it is sensible to set $c_{ij \to i+1,k} = k/(2N_{i+1})$. For $j > q$, we revert to the updates based on the normal approximation.

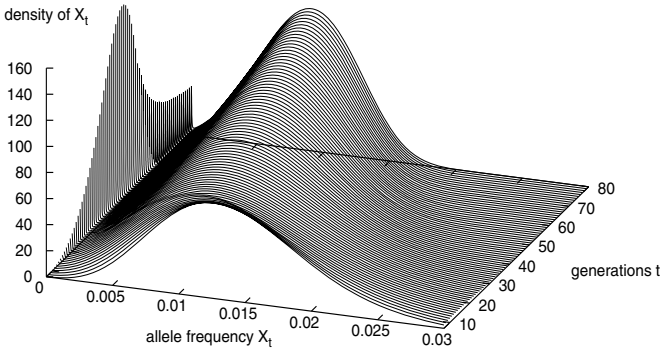


FIGURE 13.1. Density of the Frequency of a Recessive Gene

To illustrate this strategy for a recessive disease, we turn to Finland, a relatively isolated population of northern Europe. We assume that the Finnish population has grown exponentially from 1000 founders to 5,000,000 contemporary people over a span of 80 generations. Our hypothetical recessive disease has mutation rate $\eta = 10^{-6}$, fitness $f = 0.5$, and a high initial gene frequency of $X_0 = 0.015$. The slow deterministic decay to the equilibrium gene frequency of $\sqrt{\eta/(1-f)} = 0.0014$ extends well beyond the present. Figure 13.1 plots the density of the frequency of the recessive gene from generation 7 to generation 80. The figure omits the first seven generations because the densities in that time range are too concentrated for the remaining densities to scale well. The left ridge of the gene density sur-

face represents a moderate probability mass collecting in the narrow region where the gene is either extinct or in danger of going extinct.

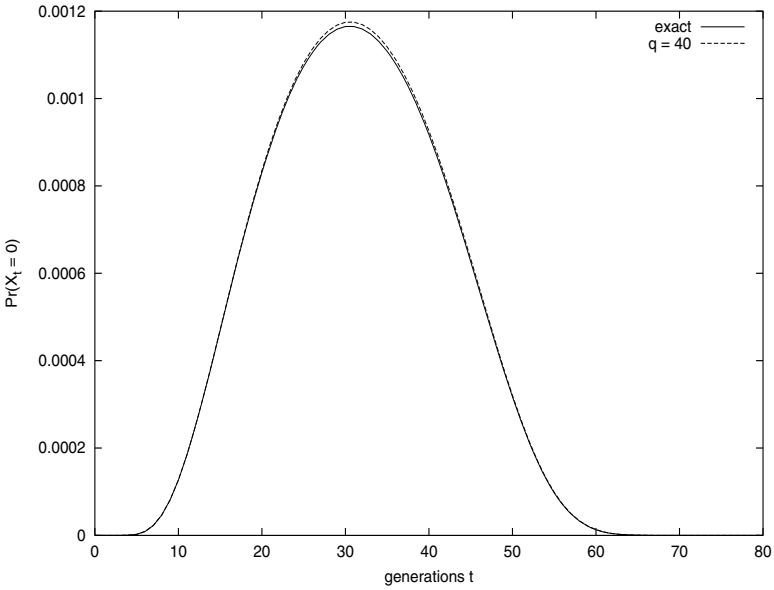


FIGURE 13.2. Extinction Probability of a Recessive Gene

As a technical aside, it is interesting to compare two versions of the algorithm. Version one carries forward probabilities but not centers of probabilities. Version two carries both forward. Version one is about twice as fast as version two, given the same mesh points at each generation. In Figure 13.1, version two relies on 175 intervals in the continuous region. With 2000 intervals in the continuous region, version one takes 25 times more computing cpu time and still fails to achieve the same accuracy at generation 80 as version two. Needless to say, version one is not recommended.

Gene extinction is naturally of great interest. Figure 13.2 depicts the probability that the recessive gene is entirely absent from the population. This focuses our attention squarely on the discrete domain where we would expect the diffusion approximation to deteriorate. The solid curve of the graph shows the outcome of computing directly with the exact Wright-Fisher chain. At about generation 60, the matrix times vector multiplications implicit in the Markov chain updates start to slow the computations drastically. In this example, it took 14 minutes of computing time on a desktop PC to reach 80 generations. The hybrid algorithm with $q = 40$ intervals covering the discrete region and 500 intervals covering the continuous region takes only 11 seconds to reach generation 80. The resulting dashed curve is quite close to the solid curve in Figure 13.2, and setting $q = 50$ makes it practically identical.

13.8 Problems

1. Prove that the matrix R defined by equation (13.1) is invertible. (Hints: Apply Proposition 7.6.1 and the Sherman-Morrison formula

$$(A+uv^T)^{-1} = A^{-1} - \frac{A^{-1}uv^TA^{-1}}{1+v^TA^{-1}u}$$

for the inverse of a rank-one perturbation of an invertible matrix.)

2. Show that the entries of the block Gauss-Seidel update (13.3) are nonnegative.

3. Suppose you are given a transition probability matrix P and desire the n-step transition probability matrix P^n for a large value of n. Devise a method of computing P^n based on the binary expansion of n that requires far fewer than $n-1$ matrix multiplications. Do not assume that P is diagonalizable.

4. Assume X and Y are independent random variables whose ranges are the nonnegative integers. Specify a finite Fourier transform method for computing the discrete density of the difference $X-Y$. Implement the method in computer code and apply it to Poisson deviates with means λ and μ. For several choices of λ and μ, check that your code yields nonnegative probabilities that sum to 1 and the correct mean and variance.

5. Explicitly calculate the finite Fourier transforms of the four sequences $c_j = 1$, $c_j = 1_{\{0\}}$, $c_j = (-1)^j$, and $c_j = 1_{\{0,1,\ldots,n/2-1\}}$ defined on $\{0,1,\ldots,n-1\}$. For the last two sequences assume that n is even.

6. Show that the sequence $c_j = j$ on $\{0,1,\ldots,n-1\}$ has finite Fourier transform

$$\hat{c}_k = \begin{cases} \frac{n-1}{2} & k=0 \\ -\frac{1}{2}+\frac{i}{2}\cot\frac{k\pi}{n} & k\neq 0. \end{cases}$$

7. For $0 \leq r < n/2$, define the rectangular and triangular smoothing sequences

$$\begin{aligned} c_j &= \frac{1}{2r+1}1_{\{-r\leq j\leq r\}} \\ d_j &= \frac{1}{r}1_{\{-r\leq j\leq r\}}\left(1-\frac{|j|}{r}\right) \end{aligned}$$

and extend them to have period n. Show that

$$\hat{c}_k = \frac{1}{n(2r+1)}\frac{\sin\frac{(2r+1)k\pi}{n}}{\sin\frac{k\pi}{n}}$$

$$\hat{d}_k = \frac{1}{nr^2}\Big(\frac{\sin\frac{rk\pi}{n}}{\sin\frac{k\pi}{n}}\Big)^2.$$

8. Prove parts (a) through (c) of Proposition A.3.1.

9. Consider a power series $f(x) = \sum_{m=0}^{\infty} c_m x^m$ with radius of convergence $r > 0$. Prove that

$$\sum_{m=k \bmod n}^{\infty} c_m x^m = \frac{1}{n}\sum_{j=0}^{n-1} u_n^{-jk} f(u_n^j x)$$

for $u_n = e^{2\pi i/n}$ and any x with $|x| < r$. As a special case, verify the identity

$$\sum_{m=k \bmod n}^{\infty} \binom{p}{m} = \frac{2^p}{n}\sum_{j=0}^{n-1} \cos\Big[\frac{(p-2k)j\pi}{n}\Big]\cos^p\Big[\frac{j\pi}{n}\Big]$$

for any positive integer p.

10. From a periodic sequence c_k with period n, form the circulant matrix

$$C = \begin{pmatrix} c_0 & c_{n-1} & c_{n-2} & \cdots & c_1 \\ c_1 & c_0 & c_{n-1} & \cdots & c_2 \\ \vdots & \vdots & \vdots & & \vdots \\ c_{n-1} & c_{n-2} & c_{n-3} & \cdots & c_0 \end{pmatrix}.$$

For $u_n = e^{2\pi i/n}$ and m satisfying $0 \le m \le n-1$, show that the vector $(u_n^{0m}, u_n^{1m}, \ldots, u_n^{(n-1)m})^t$ is an eigenvector of C with eigenvalue $n\hat{c}_m$. From this fact deduce that the circulant matrix C can be written in the diagonal form $C = UDU^*$, where D is the diagonal matrix with kth diagonal entry $n\hat{c}_{k-1}$, U is the unitary matrix with entry $u_n^{(j-1)(k-1)}/\sqrt{n}$ in row j and column k, and U^* is the conjugate transpose of U.

11. For $0 \le m \le n-1$ and a periodic function $f(x)$ on [0,1], define the sequence $b_m = f(m/n)$. If $\hat{b}_k$ is the finite Fourier transform of the sequence b_m, then we can approximate $f(x)$ by $\sum_{k=-\lfloor n/2\rfloor}^{\lfloor n/2\rfloor} \hat{b}_k e^{2\pi ikx}$. Show that this approximation is exact when $f(x)$ is equal to $e^{2\pi ijx}$, $\cos(2\pi jx)$, or $\sin(2\pi jx)$ for j satisfying $0 \le |j| < \lfloor n/2\rfloor$.

12. Continuing Problem 11, let c_k be the kth Fourier series coefficient of a general periodic function $f(x)$. If $|c_k| \le ar^{|k|}$ for constants $a \ge 0$ and $0 \le r < 1$, then verify using equation (13.5) that

$$|\hat{b}_k - c_k| \le ar^n \frac{r^k + r^{-k}}{1 - r^n}$$

for $|k| < n$. Functions analytic around 0 automatically possess Fourier coefficients satisfying the bound $|c_k| \leq ar^{|k|}$.

13. Continuing Problems 11 and 12, suppose a constant $a \geq 0$ and positive integer p exist such that

$$|c_k| \leq \frac{a}{|k|^{p+1}}$$

for all $k \neq 0$. Integration by parts shows that this criterion holds if $f^{(p+1)}(x)$ is piecewise continuous. Verify the inequality

$$|\hat{b}_k - c_k| \leq \frac{a}{n^{p+1}} \sum_{j=1}^{\infty} \left[\frac{1}{\left(j + \frac{k}{n}\right)^{p+1}} + \frac{1}{\left(j - \frac{k}{n}\right)^{p+1}} \right]$$

when $|k| < n/2$. To simplify this inequality, demonstrate that

$$\begin{aligned} \sum_{j=1}^{\infty} \frac{1}{(j+\alpha)^{p+1}} &< \int_{\frac{1}{2}}^{\infty} (x+\alpha)^{-p-1} dx \\ &= \frac{1}{p\left(\frac{1}{2}+\alpha\right)^p} \end{aligned}$$

for $\alpha > -1/2$. Finally, conclude that

$$|\hat{b}_k - c_k| < \frac{a}{pn^{p+1}} \left[\frac{1}{\left(\frac{1}{2} + \frac{k}{n}\right)^p} + \frac{1}{\left(\frac{1}{2} - \frac{k}{n}\right)^p} \right].$$

14. For a complex number c with $|c| > 1$, show that the periodic function $f(x) = (c - e^{2\pi i x})^{-1}$ has the simple Fourier series coefficients $c_k = c^{-k-1} 1_{\{k \geq 0\}}$. Argue from equation (13.5) that the finite Fourier transform approximation $\hat{b}_k$ to c_k is

$$\hat{b}_k = \begin{cases} c^{-k-1} \frac{1}{1-c^{-n}} & 0 \leq k \leq \frac{n}{2} - 1 \\ c^{-n-k-1} \frac{1}{1-c^{-n}} & -\frac{n}{2} \leq k \leq 0. \end{cases}$$

15. Let $F(s) = \sum_{n=1}^{\infty} f_n s^n$ be a probability generating function. Show that the equation $F(s) = 1$ has only the solution $s = 1$ on $|s| = 1$ if and only if the set $\{n: f_n > 0\}$ has greatest common divisor 1.

16. In the coin tossing example, prove that the probabilities u_n satisfy

$$u_n = \frac{qp^r}{1-p^r} + O(r^{-n})$$

for any $r < p^{-1}$. (Hint: Identify the singularity of $U(s) - c/(1-s)$ closest to the origin.)

17. Consider n equally spaced points on the boundary of a circle. Turing suggested a simple model for the diffusion of a morphogen, a chemical important in development, that involves the migration of the morphogen from point to point. In the stochastic version of his model, we follow a single morphogen particle. At any time the particle has the same intensity λ of jumping to the neighboring points on its right and left. Let $p_j(t)$ be the probability that the morphogen occupies point j at time t, where $j = 0, \ldots, n-1$. One can solve for these n probabilities using the finite Fourier transform on periodic sequences c_j of period n.

    (a) Show that $p_j'(t) = \lambda[p_{j-1}(t) - 2p_j(t) + p_{j+1}(t)]$.

    (b) If $a * b_k = \sum_{j=0}^{n-1} a_{k-j} b_j$ denotes the convolution of two periodic sequences of period n, then express the differential equation in part (a) as $p_j'(t) = p(t) * d_j$ for $p(t) = [p_j(t)]$ and an appropriate sequence $d = (d_j)$ of period n.

    (c) Prove that the finite Fourier transform maps the convolution $a * b_j$ into the pointwise product $n\hat{a}_k\hat{b}_k$.

    (d) Solve the transformed equations $\hat{p}_k'(t) = n\hat{p}_k(t)\hat{d}_k$.

    (e) Inverse transform to find the solutions $p_j(t)$.

    (f) Compute $\hat{d}_k$, and show that $\hat{d}_0 = 0$ and all other $\hat{d}_k$ are negative.

    (g) Deduce that $\lim_{t\to\infty} p_j(t) = \hat{p}_0(0) = 1/n$ for all j.

18. A square matrix $M = (m_{ij})$ is said to be diagonally dominant if it satisfies $|m_{ii}| > \sum_{j\neq i} |m_{ij}|$ for all i. Demonstrate that a diagonally dominant matrix is invertible. (Hint: Suppose $Mx = \mathbf{0}$. Consider the largest entry of x in magnitude.)

19. In counting jumps in a Markov chain, it is possible to explicitly calculate the matrix exponential (13.12) for a two-state chain. Show that the eigenvalues of the matrix $\text{diag}(\Lambda) + C(u)$ are

$$\omega_\pm = \frac{-(\lambda_{12} + \lambda_{21}) \pm \sqrt{(\lambda_{12} - \lambda_{21})^2 + 4u^2\lambda_{12}\lambda_{21}}}{2}$$

    when the donor and recipient subsets A and B equal $\{1, 2\}$. Argue that both eigenvalues are real and negative when $u \in [0, 1)$. One is 0 and the other is negative when $u = 1$. Find the corresponding eigenvectors of the matrix $\text{diag}(\Lambda) + C(u)$ for all $u \in [0, 1]$. What happens if $A = \{1\}$ and $B = \{2\}$ or vice versa?

20. In Moran's population genetics model, n genes evolve by substitution and mutation. Suppose each gene can be classified as one of d alleles, and let X_{ti} denote the number of alleles of type i at time t. The

count process X_t moves from state to state by randomly selecting two genes, which may coincide. The first gene dies, and the second gene reproduces a replacement. If the second gene is of type i, then its daughter gene is of type j with probability p_{ij}. The replacement times are independent and exponentially distributed with intensity λ. Reformulate Moran's model to proceed by reaction channels. What are the intensity and the increment of each channel?

21. Consider a continuous-time branching process in which the possible number of daughter particles is bounded above a common integer b for all particle types. Show how the process can be identified with a continuous-time Markov chain with a finite number of reaction channels. No approximation is necessary.

22. Prove that the linear density $f_{ij0} + f_{ij1}x$ is nonnegative throughout the interval $(a_{ij}, a_{i,j+1})$ if and only if its center of mass c_{ij} lies in the middle third of the interval. (Hint: Without loss of generality, take $a_{ij} = 0$.)

14
Poisson Approximation

14.1 Introduction

In the past few years, mathematicians have developed a powerful technique known as the Chen-Stein method for approximating the distribution of a sum of weakly dependent Bernoulli random variables [11, 18, 187]. In contrast to many asymptotic methods, this approximation carries with it explicit error bounds. Let X_α be a Bernoulli random variable with success probability p_α, where α ranges over some finite index set I. As a generalization of the law of rare events discussed in Example 14.3.1, it is natural to speculate that the sum $S = \sum_{\alpha \in I} X_\alpha$ is approximately Poisson with mean $\lambda = \sum_{\alpha \in I} p_\alpha$. The Chen-Stein method estimates the error in this approximation using the total variation distance introduced in equation (7.6) of Chapter 7.

The coupling method is one technique for explicitly bounding the total variation distance between $S = \sum_{\alpha \in I} X_\alpha$ and a Poisson random variable Z with the same mean λ [18, 136]. In many concrete examples, it is possible to construct for each α a random variable V_α on a common probability space with S such that V_α is distributed as $S - 1$ conditional on the event $X_\alpha = 1$. The bound

$$\|\pi_S - \pi_Z\|_{\text{TV}} \quad \leq \quad \frac{1 - e^{-\lambda}}{\lambda} \sum_{\alpha \in I} p_\alpha \, \text{E}(|S - V_\alpha|) \tag{14.1}$$

then applies, where π_S and π_Z denote the distributions of S and Z. The size of this bound depends on how tightly S and each V_α are coupled. If

K. Lange, *Applied Probability,*
DOI 10.1007/978-1-4419-7165-4_14, © Springer Science+Business Media, LLC 2010

$S \geq V_\alpha$ for all α, then the simplified bound

$$\|\pi_S - \pi_Z\|_{\text{TV}} \quad \leq \quad \frac{1 - e^{-\lambda}}{\lambda}[\lambda - \text{Var}(S)] \tag{14.2}$$

holds. Not only is the upper bound (14.2) easier to evaluate than the upper bound (14.1), but it also makes it clear that the approximate equality $\text{Var}(S) \approx \text{E}(S)$ is nearly a sufficient as well as a necessary condition for S to be approximately Poisson.

The neighborhood method of bounding the total variation distance exploits certain neighborhoods of dependency N_α associated with each $\alpha \in I$ [10]. Here N_α is a subset of I containing α such that X_α is independent of those X_β with $\beta \notin N_\alpha$. In this situation of short-range dependency, the total variation distance between S and its Poisson approximate Z satisfies

$$\|\pi_S - \pi_Z\|_{\text{TV}} \quad \leq \quad \frac{1 - e^{-\lambda}}{\lambda}\left(\sum_{\alpha \in I}\sum_{\beta \in N_\alpha} p_\alpha p_\beta + \sum_{\alpha \in I}\sum_{\beta \in N_\alpha \setminus \{\alpha\}} p_{\alpha\beta}\right), \tag{14.3}$$

where again $\lambda = \text{E}(S) = \text{E}(Z)$ and

$$p_{\alpha\beta} \quad = \quad \text{E}(X_\alpha X_\beta) \quad = \quad \Pr(X_\alpha = 1, X_\beta = 1).$$

The neighborhood method works best when each N_α is taken as small as possible.

Both Chen-Stein methods are well adapted to solving a myriad of practical problems. The next few sections present a few typical examples. The chapter ends with a mathematical proof of the Chen-Stein bounds. Readers primarily interested in applications can skip this theoretical section. A more comprehensive development of theory and further examples can be found in the references [11, 18, 136].

14.2 Applications of the Coupling Method

Example 14.2.1 *Ménage Problem*

In the classical ménage problem of combinatorics, n married couples are seated around a circular table [22]. If men and women alternate, but husbands and wives are randomly scrambled, then the number of married couples S seated next to each other is approximately Poisson distributed. Given that X_α is the indicator of the event that seats α and $\alpha + 1$ contain a married couple, we can write

$$S \quad = \quad \sum_{\alpha=1}^{2n} X_\alpha,$$

where $X_{2n+1} = X_1$. Symmetry dictates that $p_\alpha = \frac{1}{n}$ and $\lambda = \mathrm{E}(S) = 2$. The total variation distances between S and a Poisson random variable Z with mean λ can be estimated by the coupling method.

To construct the coupled random variable V_α, we exchange the person in seat $\alpha + 1$ with the spouse of the person in seat α and then count the number of adjacent spouse pairs, excluding the pair now occupying seats α and $\alpha + 1$. It is clear that V_α so constructed is distributed as $S - 1$ conditional on $X_\alpha = 1$. One can also show that this construction entails $|S - V_\alpha| \leq 1$. Indeed, suppose the spouse of the person in seat α occupies seat β. If $\beta = \alpha + 1$, then $X_\alpha = 1$ and $V_\alpha = S - 1$. If $\beta \neq \alpha + 1$, then the gain of a matched couple in the pair $\{\alpha, \alpha + 1\}$ does not contribute to V_α. The other possible gains and losses of matched couples occur in the three pairs $\{\alpha + 1, \alpha + 2\}$, $\{\beta - 1, \beta\}$, and $\{\beta, \beta + 1\}$. Although some of these pairs may coincide, it is not hard to see that at most one of the three pairs can suffer a loss and at most one of the three pairs can reap a gain.

We now appeal to the Chen-Stein bound (14.1). To avoid a messy consideration of special cases in calculating $\mathrm{E}(|S - V_\alpha|)$, we will bound the probability $\Pr(V_\alpha = S)$. The dominant contribution to the event $\{V_\alpha = S\}$ arises when $\beta \notin \{\alpha - 1, \alpha + 1, \alpha + 3\}$ and the person in seat $\alpha + 1$ is not the spouse of any of the people in seats $\alpha + 2$, $\beta - 1$, and $\beta + 1$. Careful consideration of this special case leads to the inequality

$$\Pr(V_\alpha = S) \geq \frac{n-3}{n}\Big(1 - \frac{3}{n-1}\Big)$$

and therefore to the further inequality

$$\begin{aligned} \mathrm{E}(|S - V_\alpha|) &= \Pr(S \neq V_\alpha) \\ &\leq 1 - \frac{n-3}{n}\Big(1 - \frac{3}{n-1}\Big) \\ &= \frac{6n-12}{n(n-1)}. \end{aligned}$$

The Chen-Stein bound (14.1) now reduces to

$$\|\pi_S - \pi_Z\|_{\mathrm{TV}} \leq \frac{2(1 - e^{-\lambda})(6n-12)}{\lambda n(n-1)},$$

which decreases in n for $n \geq 3$. ■

Example 14.2.2 *Birthday Problem*

Consider a multinomial experiment with m categories. The statistic W_d denotes the number of categories with d or more successes after n trials. For example, each category might be a day of the year, and each trial might record the birthday of another random person. If we let q_α be the success

rate per trial for category α, then this category accumulates d or more successes with probability

$$p_\alpha = \sum_{k=d}^{n} \binom{n}{k} q_\alpha^k (1-q_\alpha)^{n-k}.$$

The coupling method provides a bound on the total variation distance between W_d and a Poisson random variable with mean $\lambda = \sum_{\alpha=1}^m p_\alpha$.

To validate the coupling bound (14.2) with $S = W_d$, we must construct the coupled random variable V_α. If the number of outcomes Y_α falling in category α satisfies $Y_\alpha \geq d$, then $X_\alpha = 1$, and we set $V_\alpha = \sum_{\beta \neq \alpha} X_\beta$. If $Y_\alpha < d$, then we resample from the conditional distribution of Y_α given the event $Y_\alpha \geq d$. This produces a random variable $Y_\alpha^* > Y_\alpha$, and we redefine the outcomes of the first $Y_\alpha^* - Y_\alpha$ trials falling outside category α so that they now fall in category α. If we let V_α be the number of categories other than α that now exceed their quota d, it is obvious that V_α is distributed as $S - 1$ conditional on the event $X_\alpha = 1$. Because of the redirection of outcomes, it is also clear that $S \geq V_\alpha$. Thus, the conditions for the Chen-Stein bound (14.2) apply. Unfortunately, the variance $\mathrm{Var}(W_d)$ is not entirely trivial to calculate. In the special case $d = 1$, we have

$$\begin{aligned}
\mathrm{E}(W_1) &= \sum_{\alpha=1}^m [1 - (1-q_\alpha)^n] \\
\mathrm{Var}(W_1) &= \sum_{\alpha=1}^m \mathrm{Var}(1_{\{Y_\alpha \geq 1\}}) + \sum_{\alpha=1}^m \sum_{\beta \neq \alpha} \mathrm{Cov}(1_{\{Y_\alpha \geq 1\}}, 1_{\{Y_\beta \geq 1\}}) \\
&= \sum_{\alpha=1}^m \mathrm{Var}(1_{\{Y_\alpha = 0\}}) + \sum_{\alpha=1}^m \sum_{\beta \neq \alpha} \mathrm{Cov}(1_{\{Y_\alpha = 0\}}, 1_{\{Y_\beta = 0\}}) \\
&= \sum_{\alpha=1}^m (1-q_\alpha)^n [1 - (1-q_\alpha)^n] \\
&\quad + \sum_{\alpha=1}^m \sum_{\beta \neq \alpha} [(1-q_\alpha-q_\beta)^n - (1-q_\alpha)^n (1-q_\beta)^n],
\end{aligned}$$

which are certainly easy to evaluate numerically. Readers can consult reference [18] for various approximations to $\mathrm{E}(W_d)$ and $\mathrm{Var}(W_d)$. ■

Example 14.2.3 *Biggest Random Gap*

Questions about the spacings of uniformly distributed points crop up in many application areas [124, 184]. If we scatter n points randomly on the unit interval [0,1], then it is natural to ask for the distribution of the largest gap between two adjacent points or between either endpoint and its nearest adjacent point. We can attack this problem by the coupling method of

Chen-Stein approximation. Corresponding to the order statistics $Y_1, \ldots, Y_n$ of the n points, define indicator random variables $X_1, \ldots, X_{n+1}$ such that $X_\alpha = 1$ when $Y_\alpha - Y_{\alpha-1} \geq d$. At the ends we take $Y_0 = 0$ and $Y_{n+1} = 1$. The sum $S = \sum_{\alpha=1}^{n+1} X_\alpha$ gives the number of gaps of length d or greater.

Because we can circularize the interval, all gaps, including the first and the last, behave symmetrically. Just think of scattering $n+1$ points on the unit circle and then breaking the circle into an interval at the first random point. It therefore suffices in the coupling method to consider the first Bernoulli variable $X_1 = 1_{\{Y_1 \geq d\}}$. If $Y_1 \geq d$, then define V_1 to be the number of gaps other than Y_1 that exceed d. If, on the other hand, $Y_1 < d$, then resample Y_1 conditional on the event $Y_1 \geq d$ to get Y_1^*. For $\alpha > 1$, replace the gap $Y_\alpha - Y_{\alpha-1}$ by the gap $(Y_\alpha - Y_{\alpha-1})(1 - Y_1^*)/(1 - Y_1)$ so that the points to the right of Y_1 are uniformly chosen from the interval $[Y_1^*, 1]$ rather than from $[Y_1, 1]$. This procedure narrows all remaining gaps but leaves them in the same proportion. If we now define V_1 as the number of remaining gaps that exceed d in length, it is clear that V_1 has the same distribution as $S-1$ conditional on $X_1 = 1$. Because $S \geq V_1$, the Chen-Stein inequality (14.2) applies.

To calculate the mean $\lambda = \mathrm{E}(S)$, we again focus on the first interval. Clearly, $\Pr(X_1 = 1) = \Pr(Y_1 \geq d) = (1-d)^n$ implies that

$$\lambda \;=\; (n+1)(1-d)^n.$$

In similar fashion, we calculate

$$\begin{aligned}
\mathrm{Var}(S) &= (n+1)\,\mathrm{Var}(X_1) + (n+1)n\,\mathrm{Cov}(X_1, X_{n+1}) \\
&= (n+1)(1-d)^n - (n+1)(1-d)^{2n} \\
&\quad + (n+1)n\,\mathrm{E}(X_1 X_{n+1}) - (n+1)n(1-d)^{2n}.
\end{aligned}$$

To calculate $\mathrm{E}(X_1 X_{n+1}) = \Pr(X_1 = 1,\ X_{n+1} = 1)$ when $2d < 1$, we simply observe that $X_1 = X_{n+1} = 1$ if and only if all n random points are confined to the interval $[d, 1-d]$. It follows that $\mathrm{E}(X_1 X_{n+1}) = (1-2d)^n$ and therefore that

$$\begin{aligned}
\mathrm{Var}(S) &= (n+1)(1-d)^n - (n+1)(1-d)^{2n} \\
&\quad + (n+1)n(1-2d)^n - (n+1)n(1-d)^{2n}.
\end{aligned}$$

If d is small and n is large, then one can demonstrate that $\mathrm{Var}(S) \approx \mathrm{E}(S)$, and the Poisson approximation is good [18].

It is of some interest to estimate the average number of points required to reduce the largest gap below d. From the Poisson approximation, the median n should satisfy $e^{-(n+1)(1-d)^n} \approx \frac{1}{2}$. This approximate equality can be rewritten as

$$n \;\approx\; \frac{-\ln(n+1) + \ln\ln 2}{\ln(1-d)} \tag{14.4}$$

and used iteratively to approximate the median. If one chooses evenly spaced points, it takes only $\frac{1}{d}$ random points to saturate the interval $[0,1]$. For the crude guess $n = \frac{1}{d}$, substitution in (14.4) leads to the improved approximation

$$\begin{aligned} n &\approx \frac{-\ln(\frac{1}{d}+1)+\ln\ln 2}{\ln(1-d)} \\ &\approx \frac{1}{d}\ln\frac{1}{d}. \end{aligned}$$

In fact, a detailed analysis shows that the average required number of points is asymptotically similar to $\frac{1}{d}\ln\frac{1}{d}$ for d small [64, 184]. The factor $\ln\frac{1}{d}$ summarizes the penalty exacted for selecting random points rather than evenly spaced points. ■

14.3 Applications of the Neighborhood Method

Example 14.3.1 *The Law of Rare Events*

Suppose that $X_1, \ldots, X_n$ are independent Bernoulli random variables with success probabilities $p_1, \ldots, p_n$. If the p_α are small and $\lambda = \sum_{\alpha=1}^n p_\alpha$ is moderate in size, then the law of rare events declares that the random sum $S = \sum_{\alpha=1}^n X_\alpha$ is approximately Poisson distributed. The neighborhood method provides an easy verification of this result. If we let N_α be the singleton set $\{\alpha\}$ and Z be a Poisson random variable with mean λ, then inequality (14.3) reduces to

$$\|\pi_S - \pi_Z\|_{\text{TV}} \leq \frac{1-e^{-\lambda}}{\lambda}\sum_{\alpha=1}^n p_\alpha^2$$

because the sum $\sum_{\beta\in N_\alpha\setminus\{\alpha\}} p_{\alpha\beta}$ is empty. ■

Example 14.3.2 *Construction of Somatic Cell Hybrid Panels*

Prior to the sequencing of the human genome, somatic cell hybrids were routinely used to assign particular human genes to particular human chromosomes [48, 202]. In brief outline, somatic cell hybrids are constructed by fusing normal human cells with permanently transformed rodent cells. The resulting hybrid cells retain all of the rodent chromosomes while losing random subsets of the human chromosomes. A few generations after cell fusion, clones of cells can be identified with stable subsets of the human chromosomes. All chromosomes, human and rodent, normally remain functional. With a broad enough collection of different hybrid clones, it is possible to establish a correspondence between the presence or absence of a given human gene and the presence or absence of each of the 24 distinct

0 1 0 1 0 0 0 1 0 0 0 0 0 0 1 0 1 1 0 1 1 1 1
1 0 1 0 1 1 0 0 1 0 0 0 0 1 0 0 1 0 1 0 1 1 1
0 1 1 1 1 0 1 0 0 0 0 0 1 0 0 1 1 0 1 1 0 1 1
1 1 1 0 0 1 1 0 0 1 0 1 0 0 0 0 1 1 1 0 0 1 0 1
0 0 0 1 1 1 1 0 0 0 1 1 1 1 1 0 1 0 0 0 1 1 0
0 1 1 1 1 1 1 1 1 1 1 1 0 0 0 0 0 1 0 0 0 0 0 0
0 0 1 0 1 0 1 1 0 1 1 1 0 0 0 0 1 1 1 1 1 0 0
0 0 0 1 0 1 1 1 0 0 0 1 0 1 1 1 1 0 1 0 1 0 1
1 0 0 0 1 1 0 0 0 1 0 1 1 0 1 0 1 0 1 1 0 0 1

FIGURE 14.1. A Somatic Cell Hybrid Panel

human chromosomes in each clone. From this pattern one can assign the gene to a particular chromosome.

For this program of gene assignment to be successful, certain critical assumptions must be satisfied. First, the human gene should be present on a single human chromosome or on a single pair of homologous human chromosomes. Second, the human gene should be detectable when present in a clone and should be distinguishable from any rodent analog of the human gene in the clone. Genes are usually detected by electrophoresis of their protein products or by annealing an appropriate DNA probe directly to part of the gene. Third, each of the 24 distinct human chromosomes should be either absent from a clone or cytologically or biochemically detectable in the clone. Chromosomes can be differentiated cytologically by size, by the position of their centromeres, and by their distinctive banding patterns under appropriate stains. It is also possible to distinguish chromosomes by in situ hybridization of large, fluorescent DNA probes or by isozyme assays that detect unique proteins produced by genes on the chromosomes.

In this application of the Chen-Stein method, we consider the information content of a panel of somatic cell hybrids [77]. Let n denote the number of hybrid clones in a panel. Since the Y chromosome bears few genes of interest, hybrids are usually created from human female cells. This gives a total of 23 different chromosome types—22 autosomes and the X chromosome. Figure 14.1 depicts a hybrid panel with $n = 9$ clones. Each row of this panel corresponds to a particular clone. Each of the 23 columns corresponds to a particular chromosome. A 1 in row i and column j of the panel indicates the presence of chromosome j in clone i. A 0 indicates the absence of a chromosome in a clone. An additional test column of 0's and 1's is constructed when each clone is assayed for the presence of a given human gene. Barring assay errors or failure of one of the critical assumptions, the test column will uniquely match one of the columns of the panel. In this case the gene is assigned to the corresponding chromosome.

If two columns of a panel are identical, then gene assignment becomes ambiguous for any gene residing on one of the two corresponding chromo-

somes. Fortunately, the columns of the panel in Figure 14.1 are unique. This panel has the unusual property that every pair of columns differs in at least three entries. This level of redundancy is useful. If a single assay error is made in creating a test column for a human gene, then the gene can still be successfully assigned to a particular human chromosome because it will differ from one column of the panel in one entry and from all other columns of the panel in at least two entries. This consideration suggests that built-in redundancy of a panel is desirable. In practice, the chromosome constitution of a clone cannot be predicted in advance, and the level of redundancy is random. Minimum Hamming distance is a natural measure of the redundancy of a panel. The Hamming distance $\rho(c_s, c_t)$ between two columns c_s and c_t is just the number of entries in which they differ. The minimum Hamming distance of a panel is defined as $\min_{\{s,t\}} \rho(c_s, c_t)$, where $\{s, t\}$ ranges over all pairs of columns from the panel.

When somatic cell hybrid panels are randomly created, it is reasonable to make three assumptions. First, each human chromosome is lost or retained independently during the formation of a stable clone. Second, there is a common retention probability p applying to all chromosome pairs. This means that at least one member of each pair of homologous chromosomes is retained with probability p. Rushton [175] estimates a range of p from .07 to .75. The value $p = \frac{1}{2}$ simplifies our theory considerably. Third, different clones behave independently in their retention patterns.

Now denote column s of a random panel of n clones by C_s^n. For any two distinct columns C_s^n and C_t^n, define $X_{\{s,t\}}^n$ to be the indicator of the event $\rho(C_s^n, C_t^n) < d$, where d is some fixed Hamming distance. The random variable $Y_d^n = \sum_{\{s,t\}} X_{\{s,t\}}^n$ is 0 precisely when the minimum Hamming distance equals or exceeds d. There are $\binom{23}{2}$ pairs $\alpha = \{s, t\}$ in the index set I, and each of the associated X_α^n has the same mean

$$p_\alpha = \sum_{i=0}^{d-1} \binom{n}{i} q^i (1-q)^{n-i},$$

where $q = 2p(1-p)$ is the probability that C_s^n and C_t^n differ in any entry. This gives the mean of Y_d^n as $\lambda = \binom{23}{2} p_\alpha$.

The Chen-Stein heuristic suggests estimating $\Pr(Y_d^n > 0)$ by the Poisson tail probability $1 - e^{-\lambda}$. The error bound (14.3) on this approximation can be computed by defining the neighborhoods $N_\alpha = \{\beta : |\beta| = 2,\ \beta \cap \alpha \neq \emptyset\}$, where vertical bars enclosing a set indicate the number of elements in the set. It is clear that X_α^n is independent of those X_β^n with β outside N_α. Straightforward counting arguments give

$$\sum_{\alpha \in I} \sum_{\beta \in N_\alpha} p_\alpha p_\beta = \binom{23}{2} |N_\alpha| p_\alpha^2$$

TABLE 14.1. Chen-Stein Estimate of $\Pr(Y_d^n > 0)$

d	n	**Estimate**	**Lower Bound**	**Upper Bound**
1	10	0.2189	0.1999	0.2379
1	15	0.0077	0.0077	0.0077
1	20	0.0002	0.0002	0.0002
1	25	0.0000	0.0000	0.0000
2	10	0.9340	0.0410	1.0000
2	15	0.1162	0.1112	0.1213
2	20	0.0051	0.0050	0.0051
2	25	0.0002	0.0002	0.0002
3	10	1.0000	0.0410	1.0000
3	15	0.6071	0.4076	0.8066
3	20	0.0496	0.0487	0.0505
3	25	0.0025	0.0025	0.0025

and

$$|N_\alpha| \;=\; \binom{23}{2} - \binom{21}{2} \;=\; 43.$$

Since the joint probability $p_{\alpha\beta}$ does not depend on the particular column pair $\beta \in N_\alpha \backslash \{\alpha\}$ chosen, we also deduce that

$$\sum_{\alpha \in I} \sum_{\beta \in N_\alpha \backslash \{\alpha\}} p_{\alpha\beta} \;=\; \binom{23}{2} (|N_\alpha| - 1) p_{\alpha\beta}.$$

Fortunately, $p_{\alpha\beta} = p_\alpha^2$ when $p = 1/2$. Indeed, upon conditioning on the value of the common column shared by α and β, it is obvious in this special case that the events $X_\alpha^n = 1$ and $X_\beta^n = 1$ are independent and occur with constant probability p_α. The case $p \neq 1/2$ is more subtle, and we defer the details of computing $p_{\alpha\beta}$ to Problem 10. Table 14.1 provides some representative estimates of the probabilities $\Pr(Y_d^n > 0)$ for $p = 1/2$. Because the Chen-Stein method also provides upper and lower bounds on the estimates, we can be confident that the estimates are accurate for large n. In two cases in Table 14.1, the Chen-Stein upper bound is truncated to the more realistic value 1. ■

14.4 Proof of the Chen-Stein Estimates

Verification of the Chen-Stein estimates depends on forging a subtle connection between Chen's lemma in Example 2.7.3 and the definition of the total variation norm in equation (7.6). If the sum $S = \sum_\alpha X_\alpha$ were actually

Poisson, then Chen's lemma of Example 2.7.3 would entail the identity

$$\lambda \operatorname{E}[g(S+1)] - \operatorname{E}[Sg(S)] \;=\; 0$$

for every bounded function $g(s)$. To the extent that S is approximately Poisson, the left-hand side of this equality should be approximately 0. The total variation distance between S and a Poisson random variable Z with the same mean λ is given by

$$\|\pi_S - \pi_Z\|_{\text{TV}} \;=\; \sup_{A \subset \mathcal{Z}} \left| \Pr(S \in A) - \sum_{j \in A} e^{-\lambda} \frac{\lambda^j}{j!} \right|.$$

The key to proving the Chen-Stein estimates is to concoct a particular bounded function $g(s)$ satisfying

$$\lambda \operatorname{E}[g(S+1)] - \operatorname{E}[Sg(S)] \;=\; \Pr(S \in A) - \sum_{j \in A} e^{-\lambda} \frac{\lambda^j}{j!} \tag{14.5}$$

and then to bound the difference in expectations on the left-hand side of this equality.

The easiest way of securing equality (14.5) is to force $g(s)$ to satisfy the identity

$$\lambda g(s+1) - sg(s) \;=\; 1_A(s) - \sum_{j \in A} e^{-\lambda} \frac{\lambda^j}{j!} \tag{14.6}$$

for all nonnegative integers s. Indeed, if equation (14.6) holds, then we simply substitute the random variable S for the integer s and take expectations. Fortunately, equation (14.6) can be viewed as a recurrence relation for calculating $g(s+1)$ from $g(s)$. The value $g(0)$ is irrelevant in determining $g(1)$, so we adopt the usual convention $g(0) = 0$. One can explicitly solve the difference equation (14.6) by multiplying it by $e^{-\lambda}\lambda^{s-1}/s!$ and defining the new function $f(s) = e^{-\lambda}\lambda^s g(s+1)/s!$. These maneuvers yield the difference equation

$$f(s) - f(s-1) \;=\; e^{-\lambda} \frac{\lambda^{s-1}}{s!} 1_A(s) - e^{-\lambda} \frac{\lambda^{s-1}}{s!} \sum_{j \in A} e^{-\lambda} \frac{\lambda^j}{j!}$$

with the initial condition $f(-1) = 0$. One can now find $f(s)$ via the telescoping sum

$$\begin{aligned} f(s) &= \sum_{k=0}^{s} [f(k) - f(k-1)] \\ &= \lambda^{-1} \left[\sum_{k=0}^{s} e^{-\lambda} \frac{\lambda^k}{k!} 1_A(k) - \sum_{k=0}^{s} e^{-\lambda} \frac{\lambda^k}{k!} \sum_{j \in A} e^{-\lambda} \frac{\lambda^j}{j!} \right]. \end{aligned}$$

This translates into the solution

$$g(s+1) = \frac{e^\lambda s!}{\lambda^{s+1}} \left[\sum_{k=0}^{s} e^{-\lambda} \frac{\lambda^k}{k!} 1_A(k) - \sum_{k=0}^{s} e^{-\lambda} \frac{\lambda^k}{k!} \sum_{j \in A} e^{-\lambda} \frac{\lambda^j}{j!} \right] \quad (14.7)$$

for $g(s+1)$. Although it is not immediately evident from formula (14.7), we will demonstrate that $g(s)$ is bounded and satisfies the Lipschitz inequality $|g(s+1) - g(s)| \leq (1 - e^{-\lambda})/\lambda$ for all s.

Before attending to these important details, let us return to the main line of argument. We first note that for any random variable T

$$\begin{aligned} \mathrm{E}(ST) &= \sum_\alpha \mathrm{E}(X_\alpha T) \\ &= \sum_\alpha p_\alpha \, \mathrm{E}(X_\alpha T \mid X_\alpha = 1) \\ &= \sum_\alpha p_\alpha \, \mathrm{E}(T \mid X_\alpha = 1). \end{aligned} \quad (14.8)$$

We now apply this identity to $T = g(S)$ and invoke the coupling-method premise that $V_\alpha + 1$ has the same distribution as S conditional on the event $X_\alpha = 1$. These considerations imply that

$$\begin{aligned} \mathrm{E}[Sg(S)] &= \sum_\alpha p_\alpha \, \mathrm{E}[g(S) \mid X_\alpha = 1] \\ &= \sum_\alpha p_\alpha \, \mathrm{E}[g(V_\alpha + 1)]. \end{aligned} \quad (14.9)$$

We also observe that the Lipschitz condition on $g(s)$ can be extended to

$$\begin{aligned} |g(t) - g(s)| &\leq \sum_{j=s}^{t-1} |g(j+1) - g(j)| \\ &\leq \frac{1 - e^{-\lambda}}{\lambda} |t - s| \end{aligned}$$

for any $t > s$. Mindful of these facts, we infer from equality (14.5) that

$$\begin{aligned} \left| \Pr(S \in A) - \sum_{j \in A} e^{-\lambda} \frac{\lambda^j}{j!} \right| &= |\lambda \, \mathrm{E}[g(S+1)] - \mathrm{E}[Sg(S)]| \\ &= \left| \sum_\alpha p_\alpha \{ \mathrm{E}[g(S+1)] - \mathrm{E}[g(V_\alpha + 1)] \} \right| \\ &\leq \sum_\alpha p_\alpha \, \mathrm{E}[|g(S+1) - g(V_\alpha + 1)|] \\ &\leq \frac{1 - e^{-\lambda}}{\lambda} \sum_\alpha p_\alpha \, \mathrm{E}(|S - V_\alpha|). \end{aligned}$$

Taking the supremum over A now yields the Chen-Stein bound (14.1).

If $S \geq V_\alpha$ for all α, then equation (14.8) with $T = S$ implies

$$\begin{aligned}\sum_\alpha p_\alpha \operatorname{E}(|S - V_\alpha|) &= \sum_\alpha p_\alpha \operatorname{E}(S) + \lambda - \sum_\alpha p_\alpha \operatorname{E}(V_\alpha + 1) \\ &= \lambda^2 + \lambda - \sum_\alpha p_\alpha \operatorname{E}(S \mid X_\alpha = 1) \\ &= \lambda^2 + \lambda - \operatorname{E}(S^2) \\ &= \lambda - \operatorname{Var}(S).\end{aligned}$$

This establishes the Chen-Stein bound (14.2).

In the neighborhood method, it is convenient to define the random variable $U_\alpha = \sum_{\beta \notin N_\alpha} X_\beta$, which is independent of X_α. Because

$$g(S - X_\alpha + 1) = \begin{cases} g(S+1), & X_\alpha = 0 \\ g(S), & X_\alpha = 1, \end{cases}$$

we have

$$\begin{aligned}&\lambda \operatorname{E}[g(S+1)] - \operatorname{E}[Sg(S)] \\ = \;& \sum_\alpha \operatorname{E}[p_\alpha g(S+1) - X_\alpha g(S)] \\ = \;& \sum_\alpha \operatorname{E}[p_\alpha g(S+1) - p_\alpha g(S - X_\alpha + 1)] \\ &+ \sum_\alpha \operatorname{E}[p_\alpha g(S - X_\alpha + 1) - X_\alpha g(S)] \qquad (14.10) \\ = \;& \sum_\alpha \operatorname{E}\{p_\alpha X_\alpha [g(S+1) - g(S)]\} \\ &+ \sum_\alpha \operatorname{E}\{(p_\alpha - X_\alpha)[g(S - X_\alpha + 1) - g(U_\alpha + 1)]\} \\ &+ \sum_\alpha \operatorname{E}[(p_\alpha - X_\alpha) g(U_\alpha + 1)].\end{aligned}$$

We will bound each of the sums defining the final quantity in the string of equalities (14.10). The third sum equals 0 since $\operatorname{E}[(p_\alpha - X_\alpha) g(U_\alpha + 1)] = 0$ by independence. The first sum is bounded in absolute value by

$$\begin{aligned}\sum_\alpha \operatorname{E}[p_\alpha X_\alpha |g(S+1) - g(S)|] &\leq \frac{1 - e^{-\lambda}}{\lambda} \sum_\alpha p_\alpha \operatorname{E}(X_\alpha) \\ &= \frac{1 - e^{-\lambda}}{\lambda} \sum_\alpha p_\alpha^2 \qquad (14.11)\end{aligned}$$

owing to the Lipschitz property of $g(s)$. The middle sum is bounded in absolute value by

$$\sum_\alpha \operatorname{E}[(p_\alpha + X_\alpha) |g(S - X_\alpha + 1) - g(U_\alpha + 1)|]$$

$$\begin{aligned} &\leq \frac{1-e^{-\lambda}}{\lambda} \sum_{\alpha} \sum_{\beta \in N_\alpha \setminus \{\alpha\}} \mathrm{E}[(p_\alpha + X_\alpha) X_\beta] \qquad (14.12) \\ &= \frac{1-e^{-\lambda}}{\lambda} \sum_{\alpha} \sum_{\beta \in N_\alpha \setminus \{\alpha\}} (p_\alpha p_\beta + p_{\alpha\beta}) \end{aligned}$$

based on the extended Lipschitz property. Combining inequalities (14.11) and (14.12) with equalities (14.5) and (14.10) now produces the Chen-Stein bound (14.3).

Returning now to the question of whether $g(s)$ is a bounded function, we rearrange equation (14.7) by subtracting and adding the same quantity $e^\lambda s! \lambda^{-s-1} \sum_{k=0}^s e^{-\lambda} \frac{\lambda^k}{k!} 1_A(k) \sum_{j=0}^s e^{-\lambda} \frac{\lambda^j}{j!}$. This tactic gives

$$\begin{aligned} g(s+1) &= e^\lambda s! \lambda^{-s-1} \sum_{k=0}^{s} e^{-\lambda} \frac{\lambda^k}{k!} 1_A(k) \sum_{j=s+1}^{\infty} e^{-\lambda} \frac{\lambda^j}{j!} \\ &\quad - e^\lambda s! \lambda^{-s-1} \sum_{k=s+1}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} 1_A(k) \sum_{j=0}^{s} e^{-\lambda} \frac{\lambda^j}{j!}. \end{aligned}$$

From this representation and Proposition 4.3, we deduce the integral bound

$$\begin{aligned} |g(s+1)| &\leq e^\lambda s! \lambda^{-s-1} \sum_{k=s+1}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} \\ &= e^\lambda s! \lambda^{-s-1} \int_0^\lambda e^{-r} \frac{r^s}{s!} \, dr \\ &= \lambda^{-1} \int_0^\lambda e^{\lambda - r} \left(\frac{r}{\lambda}\right)^s dr \\ &\leq \lambda^{-1} \int_0^\lambda e^{\lambda - r} dr. \end{aligned}$$

Finally, let us tackle the delicate issue of proving that $g(s)$ is Lipschitz. We first note that $g(s)$ can be decomposed as the sum $g(s) = \sum_{j \in A} g_j(s)$ of solutions $g_j(s)$ to the difference equation (14.6) with the singletons $\{j\}$ substituting for the set A. This fact is immediately obvious from formula (14.7). Furthermore, the function $g_j(s)$ can be expressed as

$$g_j(s+1) = \begin{cases} 0, & s = -1 \\ -\frac{s!}{\lambda^{s+1-j} j!} \sum_{k=0}^{s} e^{-\lambda} \frac{\lambda^k}{k!}, & 0 \leq s < j \\ \frac{s!}{\lambda^{s+1-j} j!} \sum_{k=s+1}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!}, & s \geq j \, . \end{cases} \qquad (14.13)$$

Problem 14 asks the reader to check this solution. For $s < j$ the difference $g_j(s+1) - g_j(s) \leq 0$ because

$$\frac{s}{\lambda} \sum_{k=0}^{s} e^{-\lambda} \frac{\lambda^k}{k!} \geq \sum_{k=0}^{s-1} e^{-\lambda} \frac{\lambda^k}{k!}$$

owing to the inequality $\lambda^k/k! \geq \lambda^k/[(k-1)!s]$ for $k = 1, \ldots, s$. For $s > j$ again the difference $g_j(s+1) - g_j(s) \leq 0$ because

$$\frac{s}{\lambda} \sum_{k=s+1}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} \quad \leq \quad \sum_{k=s}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!}$$

owing to the opposite inequality $\lambda^k/k! \leq \lambda^k/[(k-1)!s]$ for $k \geq s+1$. Only the difference $g_j(j+1) - g_j(j) \geq 0$, and this difference is bounded above by

$$\begin{aligned} g_j(j+1) - g_j(j) &= \frac{1}{\lambda} \sum_{k=j+1}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} + \frac{1}{j} \sum_{k=0}^{j-1} e^{-\lambda} \frac{\lambda^k}{k!} \\ &= \frac{e^{-\lambda}}{\lambda} \left[\sum_{k=j+1}^{\infty} \frac{\lambda^k}{k!} + \sum_{k=1}^{j} \frac{\lambda^k}{k!} \frac{k}{j} \right] \\ &\leq \frac{e^{-\lambda}}{\lambda} (e^{\lambda} - 1) \\ &= \frac{1 - e^{-\lambda}}{\lambda}. \end{aligned}$$

This upper-bound inequality carries over to

$$g(s+1) - g(s) \quad = \quad \sum_{j \in A} [g_j(s+1) - g_j(s)]$$

since only one difference on its right-hand sum is nonnegative for any given s. Finally, inspection of the solution (14.7) makes it evident that the function $h(s) = -g(s)$ solves the difference equation (14.6) with the complement A^c replacing A. It follows that

$$g(s) - g(s+1) \quad = \quad h(s+1) - h(s) \quad \leq \quad \frac{1 - e^{-\lambda}}{\lambda},$$

and this completes the proof that $g(s)$ satisfies the Lipschitz condition.

14.5 Problems

1. Verify that $\lambda^{-1}(1 - e^{-\lambda}) \leq \min\{1, \lambda^{-1}\}$ for all $\lambda > 0$.
2. For a random permutation $\sigma_1, \ldots, \sigma_n$ of $\{1, \ldots, n\}$, let $X_\alpha = 1_{\{\sigma_\alpha = \alpha\}}$ be the indicator of a match at position α. Show that the total number of matches $S = \sum_{\alpha=1}^{n} X_\alpha$ satisfies the coupling bound

$$\|\pi_S - \pi_Z\|_{\text{TV}} \quad \leq \quad \frac{2(1 - e^{-1})}{n},$$

where Z follows a Poisson distribution with mean 1.

3. In Problem 2 prove that the total variation inequality can be improved to

$$\|\pi_S - \pi_Z\|_{\text{TV}} \quad \leq \quad \frac{2^{n+1} + e^{-1}}{(n+1)!}.$$

This obviously represents much faster convergence. (Hints: Use the exact probability of k matches $p_{[k]}$ from Example 4.3.1 and the second definition of the total variation norm in (7.6). Combine these with the binomial expansion of $(1+1)^{n+1}$ and the bound

$$\sum_{j=k}^{\infty} \frac{1}{j!} \quad \leq \quad \frac{2}{k!}$$

for $k \geq 1$.)

4. In the ménage problem, prove that $\text{Var}(S) = 2 - 2/(n-1)$.

5. In certain situations the hypergeometric distribution can be approximated by a Poisson distribution. Suppose that w white balls and b black balls occupy a box. If you extract $n < w + b$ balls at random, then the number of white balls S extracted follows a hypergeometric distribution. Note that if we label the white balls $1, \ldots, w$, and let X_α be the random variable indicating whether white ball α is chosen, then $S = \sum_{\alpha=1}^{w} X_\alpha$. One can construct a coupling between S and V_α by the following device. If white ball α does not show up, then randomly take one of the balls extracted and exchange it for white ball α. Calculate an explicit Chen-Stein bound, and give conditions under which the Poisson approximation to S will be good.

6. In the context of Example 14.3.1 on the law of rare events, prove the less stringent bound

$$\|\pi_S - \pi_Z\|_{\text{TV}} \quad \leq \quad \sum_{\alpha=1}^{n} p_\alpha^2$$

by invoking Problems 29 and 30 of Chapter 7.

7. Consider the n-dimensional unit cube $[0,1]^n$. Suppose that each of its $n2^{n-1}$ edges is independently assigned one of two equally likely orientations. Let S be the number of vertices at which all neighboring edges point toward the vertex. The Chen-Stein method implies that S has an approximate Poisson distribution Z with mean 1. Use the neighborhood method to verify the estimate

$$\|\pi_S - \pi_Z\|_{\text{TV}} \leq (n+1)2^{-n}(1 - e^{-1}).$$

(Hints: Let I be the set of all 2^n vertices, X_α the indicator that vertex α has all of its edges directed toward α, and $N_\alpha = \{\beta : \|\beta - \alpha\| \leq 1\}$. Note that X_α is independent of those X_β with $\|\beta - \alpha\| > 1$. Also, $p_{\alpha\beta} = 0$ for $\|\beta - \alpha\| = 1$.)

8. A graph with n nodes is created by randomly connecting some pairs of nodes by edges. If the connection probability per pair is p, then all pairs from a triple of nodes are connected with probability p^3. For p small and $\lambda = \binom{n}{3}p^3$ moderate in size, the number of such triangles in the random graph is approximately Poisson with mean λ. Use the neighborhood method to estimate the total variation error in this approximation.

9. Suppose n balls (people) are uniformly and independently distributed into m boxes (days of the year). The birthday problem involves finding the approximate distribution of the number of boxes that receive d or more balls for some fixed positive integer d. This is a special case of the Poisson approximation treated in Example 14.2.2 by the coupling method. In this exercise we attack the birthday problem by the neighborhood method. To get started, let the index set I be the collection of all sets of trials $\alpha \subset \{1, \ldots, n\}$ having $|\alpha| = d$ elements. Let X_α be the indicator of the event that the balls indexed by α all fall into the same box. If $S = \sum_\alpha X_\alpha$, then argue that the approximation $\Pr(S = 0) \approx e^{-\lambda}$ is plausible when

$$\lambda = \binom{n}{d}\frac{1}{m^{d-1}}.$$

Now define the neighborhoods N_α so that X_α is independent of those X_β with β outside N_α. Demonstrate that

$$\begin{aligned}
\sum_{\alpha\in I}\sum_{\beta\in N_\alpha} p_\alpha p_\beta &= \binom{n}{d}\left[\binom{n}{d} - \binom{n-d}{d}\right]\left(\frac{1}{m}\right)^{2d-2} \\
\sum_{\alpha\in I}\sum_{\beta\in N_\alpha\setminus\{\alpha\}} p_{\alpha\beta} &= \binom{n}{d}\sum_{i=1}^{d-1}\binom{d}{i}\binom{n-d}{d-i}\left(\frac{1}{m}\right)^{2d-i-1}.
\end{aligned}$$

When $d = 2$, calculate the total variation bound

$$\|\pi_S - \pi_Z\|_{\text{TV}} \leq \frac{1 - e^{-\lambda}}{\lambda}\frac{\binom{n}{2}(4n-7)}{m^2}.$$

10. In the somatic cell hybrid model, suppose that the retention probability $p \neq \frac{1}{2}$. Define $w_{n,d_{12},d_{13}} = \Pr[\rho(C_1^n, C_2^n) = d_{12},\ \rho(C_1^n, C_3^n) = d_{13}]$ for a random panel with n clones. Show that

$$p_{\alpha\beta} = \sum_{d_{12}=0}^{d-1}\sum_{d_{13}=0}^{d-1} w_{n,d_{12},d_{13}},$$

regardless of which $\beta \in N_\alpha \backslash \{\alpha\}$ is chosen [76]. Setting $r = p(1-p)$, verify the recurrence relation

$$\begin{aligned} w_{n+1,d_{12},d_{13}} &= r(w_{n,d_{12}-1,d_{13}} + w_{n,d_{12},d_{13}-1} + w_{n,d_{12}-1,d_{13}-1}) \\ &\quad + (1-3r)w_{n,d_{12},d_{13}}. \end{aligned}$$

Under the natural initial conditions, $w_{0,d_{12},d_{13}}$ is 1 when $d_{12} = d_{13} = 0$ and 0 otherwise.

11. In the somatic cell hybrid model, suppose that one knows a priori that the number of assay errors does not exceed some positive integer d. Prove that assay error can be detected if the minimum Hamming distance of the panel is strictly greater than d. Prove that the locus can still be correctly assigned to a single chromosome if the minimum Hamming distance is strictly greater than $2d$.

12. Consider an infinite sequence $W_1, W_2, \ldots$ of independent, Bernoulli random variables with common success probability p. Let X_α be the indicator of the event that a success run of length t or longer begins at position α. Note that $X_1 = \prod_{k=1}^t W_k$ and

$$X_j = (1 - W_{j-1}) \prod_{k=j}^{j+t-1} W_k$$

for $j > 1$. The number of such success runs starting in the first n positions is given by $S = \sum_{\alpha \in I} X_\alpha$, where the index set $I = \{1, \ldots, n\}$. The Poisson heuristic suggests the S is approximately Poisson with mean $\lambda = p^t[(n-1)(1-p)+1]$. Let $N_\alpha = \{\beta \in I : |\beta - \alpha| \leq t\}$. Show that X_α is independent of those X_β with β outside N_α. In the Chen-Stein bound (14.3), prove that $\sum_{\alpha \in I} \sum_{\beta \in N_\alpha \backslash \{\alpha\}} p_{\alpha\beta} = 0$. Finally, show that $b_1 = \sum_{\alpha \in I} \sum_{\beta \in N_\alpha} p_\alpha p_\beta \leq \lambda^2(2t+1)/n + 2\lambda p^t$. (Hint:

$$\begin{aligned} b_1 &= p^{2t} + 2tp^{2t}(1-p) \\ &\quad + [2nt - t^2 + n - 3t - 1]p^{2t}(1-p)^2 \end{aligned}$$

exactly. Note that the pairs α and β entering into the double sum for b_1 are drawn from the integer lattice points $\{(i,j) : 1 \leq i, j \leq n\}$. An upper-left triangle and a lower-right triangle of lattice points from this square do not qualify for the double sum defining b_1. The term p^{2t} in b_1 corresponds to the lattice point $(1,1)$.)

13. In the coupling method demonstrate the bound

$$\Pr(S > 0) \geq \sum_{\alpha \in I} \frac{p_\alpha}{1 + \mathrm{E}(V_\alpha)}.$$

See reference [170] for some numerical examples. (Hints: Choose T appropriately in equality (14.8) and apply Jensen's inequality.)

14. Verify formula (14.13) by induction on s.

15
Number Theory

15.1 Introduction

Number theory is one of the richest and oldest branches of mathematics. It is notable for its many unsolved but easily stated conjectures. The current chapter touches on issues surrounding prime numbers and their density. In particular, the chapter and book culminate with a proof of the prime number theorem. This highlight of 19th century mathematics was surmised by Legendre and Gauss, attacked by Riemann and Chebyshev, and finally proved by Hadamard and de la Valléé Poussin. These mathematicians created a large part of analytic function theory in the process. In the mid-20th century, Erdös and Selberg succeeded in crafting a proof that avoids analytic functions. Even so, their elementary proof is longer and harder to comprehend than the classical proofs. Our treatment follows the recent trail blazed by Newman [150] and Zagier [211] that uses a minimum of analytic function theory. We particularly stress the connections and insight provided by probability.

In our exposition, we will take several mathematical facts for granted. For example provided no $a_n = -1$, the absolute convergence of the infinite product $\prod_{n=1}^{\infty}(1 + a_n)$ to a nonzero number is equivalent to the absolute convergence of the infinite series $\sum_{n=1}^{\infty} a_n$. Absolute convergence of either an infinite series or an infinite product implies convergence of the corresponding series or product [6, 97].

The necessary background in number theory is even more slender [8, 83, 104, 151]; Appendix A.1 covers the bare essentials. Multiplicative number

K. Lange, *Applied Probability,*
DOI 10.1007/978-1-4419-7165-4_15, © Springer Science+Business Media, LLC 2010

theory deals with the set of positive integers (or natural numbers). The integer a divides the integer b, written $a \mid b$, if $b = ac$ for some integer c. A natural number $p > 1$ is said to be prime if its only positive divisors are 1 and itself. Thus, 2, 3, 5, 7, 11, 13, 17, 19, and so forth are primes; 1 is not prime. The number of primes is infinite. The classical proof of Euclid proceeds by contradiction. Suppose $p_1, \ldots, p_n$ is a complete list of the primes. Then the number $1 + \prod_{i=1}^{n} p_i$ is not divisible by any of the primes in the list and consequently must itself be prime. Equally important is the fundamental theorem of arithmetic. This theorem says that every natural number can be factored into a product of primes. Such a representation is unique except for the order of the factors. For example, the composite number $60 = 2^2 3^1 5^1$. Two integers are said to be relatively prime if they possess no common factors.

15.2 Zipf's Distribution and Euler's Theorem

Riemann's zeta function $\zeta(s) = \sum_{n=1}^{\infty} n^{-s}$ converges for $s > 1$. Thus, Zipf's probability measure

$$\omega_s(A) \quad = \quad \frac{1}{\zeta(s)} \sum_{n \in A} n^{-s}$$

on the natural numbers makes sense. It obviously puts more weight on small numbers than on large numbers. If $p \neq q$ are prime numbers, then one can also show by direct calculation that the sets $D_p = \{kp : k \geq 1\}$ and $D_q = \{kq : k \geq 1\}$ of integers divisible by p and q are independent under ω_s.

It is more illuminating to reverse the process, start from independence, and construct ω_s indirectly. Toward this end, consider a sequence of independent, geometrically distributed random variables X_p indexed by the prime numbers p. Here X_p counts the number of failures until success in a sequence of Bernoulli trials with failure probability p^{-s}. Straightforward calculations demonstrate that X_p has mean

$$\frac{p^{-s}}{1 - p^{-s}} \quad = \quad \frac{1}{p^s - 1}$$

and that the event $A_p = \{X_p > 0\}$ has probability p^{-s}. Because

$$\sum_p \Pr(A_p) \quad = \quad \sum_p p^{-s} \quad < \quad \sum_{n=1}^{\infty} n^{-s} \quad < \quad \infty,$$

the Borel-Cantelli lemma implies that only finitely many of the A_p occur. This means that all but a finite number of the factors of the infinite product $N = \prod_p p^{X_p}$ reduce to 1.

We now claim that N has Zipf's distribution. Indeed, if the integer n has unique prime factorization $n = \prod_p p^{x_p}$, then the continuity of probability on a decreasing sequence of events implies

$$\begin{aligned}\Pr(N = n) &= \prod_p \Pr(X_p = x_p) \\ &= \prod_p \left(1 - p^{-s}\right) p^{-x_p s} \\ &= n^{-s} \prod_p \left(1 - p^{-s}\right).\end{aligned}$$

The equation

$$\prod_p \left(1 - p^{-s}\right) = \frac{1}{\zeta(s)} \tag{15.1}$$

identifying the normalizing constant figures in the proof of Proposition 15.2.1. More to the point, independence of the events D_p and D_q is now trivial because in this setting D_p reduces to A_p and D_q to A_q.

Number theory, like many branches of mathematics, has its own jargon. A real or complex-valued function defined on the natural numbers is called an arithmetic function. From our perspective, an arithmetic function is a random variable $Y = f(N)$ with value $f(n)$ at the sample point n. To avoid confusion, we will subscript our expectation signs by the parameter s so that

$$\mathrm{E}_s(Y) = \int f(n)\, d\omega_s(n) = \zeta(s)^{-1} \sum_{n=1}^{\infty} f(n) n^{-s}.$$

The best behaved arithmetic functions Y are completely multiplicative in the sense that $f(mn) = f(m)f(n)$ for all m and n. The arithmetic function $f(n) = n^r$ furnishes an example. A multiplicative function $f(n)$ is only required to satisfy $f(mn) = f(m)f(n)$ when m and n are relatively prime. Excluding the trivial case $f(n) \equiv 0$, both definitions require $f(1) = 1$. Indeed, if $f(n) \neq 0$ for some n, then the equation $f(n) = f(n)f(1)$ is only possible if $f(1) = 1$. The sets of completely multiplicative and multiplicative functions are closed under the formation of pointwise products and, provided division by 0 is not involved, pointwise quotients. A random variable Y defined by a multiplicative function $f(n)$ splits into a product $Y = \prod_p f(p^{X_p})$ of independent random variables depending on the prime powers X_p. With probability 1 only a finite number of factors of the infinite product $\prod_p f(p^{X_p})$ differ from 1. The importance of multiplicative functions stems from the following result.

Proposition 15.2.1 (Euler) *Suppose the multiplicative arithmetic function $Y = f(N)$ has finite expectation. Then*

$$\mathrm{E}_s(Y) \;=\; \prod_p \mathrm{E}_s\left[f\left(p^{X_p}\right)\right], \tag{15.2}$$

where p extends over all primes. If $f(n)$ is completely multiplicative, then

$$\mathrm{E}_s\left[f\left(p^{X_p}\right)\right] \;=\; \mathrm{E}_s\left[f(p)^{X_p}\right] \;=\; \frac{1-p^{-s}}{1-f(p)p^{-s}}. \tag{15.3}$$

Proof: In view of equation (15.1), it suffices to prove

$$\sum_{n=1}^{\infty} \frac{f(n)}{n^s} \;=\; \prod_p \left[1+\sum_{m=1}^{\infty} \frac{f(p^m)}{p^{ms}}\right].$$

If we let $g(n) = f(n)n^{-s}$, then $g(n)$ is multiplicative whenever $f(n)$ is multiplicative. In other words, it is enough to prove that

$$\sum_{n=1}^{\infty} g(n) \;=\; \prod_p \left[1+\sum_{m=1}^{\infty} g(p^m)\right] \tag{15.4}$$

for $g(n)$ multiplicative. The infinite sum on the left of equation (15.4) converges absolutely by assumption. The infinite product on the right is well defined and converges absolutely because

$$\begin{aligned}\sum_{p\le q}\left|\sum_{m=1}^{\infty} g(p^m)\right| &\le \sum_{p\le q}\sum_{m=1}^{\infty} |g(p^m)| \\ &\le \sum_{n=1}^{\infty} |g(n)|\end{aligned}$$

for all primes q. Therefore, both sides of the proposed equality (15.4) make sense.

We can rearrange the terms in the finite product

$$\pi_q \;=\; \prod_{p\le q}\left[1+\sum_{m=1}^{\infty} g(p^m)\right]$$

of absolutely convergent series without altering the value of the product. Exploiting the multiplicative nature of $g(n)$, we find that

$$\pi_q \;=\; \sum_{n\in B_q} g(n),$$

where B_q consists of all natural numbers having no prime factor strictly greater than q. It follows that

$$\begin{aligned}\left|\sum_{n=1}^{\infty} g(n) - \pi_q\right| &\leq \sum_{n \notin B_q} |g(n)| \\ &\leq \sum_{n>q} |g(n)|.\end{aligned}$$

This last sum can be made arbitrarily small by taking q large enough. This proves identity (15.2).

To verify identity (15.3), we calculate

$$\begin{aligned}\mathrm{E}_s\left[f\left(p^{X_p}\right)\right] &= \mathrm{E}_s[f(p)^{X_p}] \\ &= \frac{\sum_{m=0}^{\infty} f(p)^m p^{-ms}}{\sum_{m=0}^{\infty} p^{-ms}} \\ &= \frac{1-p^{-s}}{1-f(p)p^{-s}}\end{aligned}$$

by summing the indicated geometric series. ■

Example 15.2.1 *A Scaled Version of Euler's Totient*

Euler's totient function $\varphi(n)$, mentioned in Example 4.3.2, counts the numbers between 1 and n that are relatively prime to n. Because $\varphi(n)$ satisfies

$$\frac{\varphi(n)}{n} = \prod_p \left(1-\frac{1}{p}\right)^{1_{\{x_p>0\}}}, \tag{15.5}$$

we have

$$\begin{aligned}E_s\left[\frac{\varphi(N)}{N}\right] &= \prod_p \mathrm{E}_s\left[\left(1-p^{-1}\right)^{1_{\{X_p>0\}}}\right] \\ &= \prod_p \left[\left(1-p^{-s}\right)\left(1-p^{-1}\right)^0 + p^{-s}\left(1-p^{-1}\right)^1\right] \\ &= \prod_p \left(1-p^{-s-1}\right) \\ &= \frac{1}{\zeta(s+1)}.\end{aligned}$$

It is noteworthy that this expectation tends to the limit $\zeta(2)^{-1} = 6/\pi^2$ as s tends to 1. Section 15.5 and Problem 17 explore this phenomenon in more detail. ■

Example 15.2.2 *Expected Number of Divisors*

A moment's reflection shows that the number of divisors of a random natural number N can be expressed as $\tau(N) = \prod_p (1 + X_p)$. Euler's formula (15.2) and equation (15.1) give

$$\begin{aligned} \mathrm{E}_s[\tau(N)] &= \prod_p \mathrm{E}_s(1 + X_p) \\ &= \prod_p \left(1 + \frac{1}{p^s - 1}\right) \\ &= \prod_p \frac{1}{1 - p^{-s}} \\ &= \zeta(s). \end{aligned}$$

This pleasant surprise should not lull the reader into thinking that other relevant expectations yield so easily. ■

Example 15.2.3 *Evaluation of* $\mathrm{E}_s(\ln N)$

To avoid the impression that Euler's formula is the only method of calculating expectations, consider

$$\begin{aligned} -\frac{d}{ds} \ln \zeta(s) &= -\zeta(s)^{-1} \zeta'(s) \\ &= \zeta(s)^{-1} \sum_{n=1}^{\infty} \frac{\ln n}{n^s} \\ &= \mathrm{E}_s(\ln N). \end{aligned}$$

Symbolic algebra programs such as Maple are capable of numerically evaluating and differentiating $\zeta(s)$. The accurate bounds

$$\frac{\ln 2}{2^{s-1} - 1} \left[1 - \frac{1}{2\zeta(s)}\right] \leq \mathrm{E}_s(\ln N) \leq \frac{\ln 2}{2^{s-1} - 1} \tag{15.6}$$

derived in Problem 8 are more enlightening because they clarify the order of magnitude of $\mathrm{E}_s(\ln N)$. ■

15.3 Dirichlet Products and Möbius Inversion

Many of the arithmetic functions pursued in analytic number theory are multiplicative. A few examples are

$$\delta(n) = \begin{cases} 1 & n = 1 \\ 0 & n > 1 \end{cases}$$

$$\begin{aligned}
\mathbf{1}(n) &= 1 \\
\mathrm{id}(n) &= n \\
\varphi(n) &= n \prod_{p|n} \left(1 - \frac{1}{p}\right) \\
\sigma_k(n) &= \sum_{d|n} d^k.
\end{aligned}$$

In this list, n is any natural number, p any prime, and k any nonnegative integer. We have already met the two arithmetic functions $\tau(n) = \sigma_0(n)$ and $\varphi(n)$. The sum of the divisors of n has the alias $\sigma(n) = \sigma_1(n)$. One of the goals of this section is to prove that all of the arithmetic functions $\sigma_k(n)$ are multiplicative. The von Mangoldt function

$$\Lambda(n) = \begin{cases} \ln p & \text{if } n = p^m \text{ for some prime } p \\ 0 & \text{otherwise} \end{cases}$$

is a prominent arithmetic function that fails to be multiplicative. Its relevance arises from the representation

$$\ln n = \sum_{d|n} \Lambda(d). \tag{15.7}$$

To recognize multiplicative functions, it helps to define a convolution operation sending a pair of arithmetic functions $f(n)$ and $g(n)$ into the new arithmetic function

$$f * g(n) = \sum_{d|n} f(d)g(n/d)$$

termed their Dirichlet product. Among the virtues of this definition are the formulas

$$\begin{aligned}
f * \delta(n) &= f(n) \\
f * g(n) &= \sum_{ab=n} f(a)g(b) \\
&= g * f(n) \\
f * (g * h)(n) &= \sum_{abc=n} f(a)g(b)h(c) \\
&= (f * g) * h(n).
\end{aligned}$$

Except for the existence of an inverse, these are precisely the axioms characterizing a commutative group [8]. Because δ serves as an identity element, the inverse $f^{[-1]}$ of an arithmetic function f must satisfy the equation $f * f^{[-1]} = \delta$. In particular, $1 = f * f^{[-1]}(1) = f(1)f^{[-1]}(1)$. Thus, we must restrict our attention to arithmetic functions with $f(1) \neq 0$

to achieve a group structure. With this proviso, it remains to specify the inverse $f^{[-1]}$ of f. Fortunately, this is accomplished inductively through the formula

$$f^{[-1]}(n) = -f(1)^{-1} \sum_{d|n;\, d>1} f(d) f^{[-1]}(n/d),$$

beginning with $f^{[-1]}(1) = f(1)^{-1}$.

The multiplicative functions form a subgroup of this group. To prove part of this assertion, consider two such functions f and g and two relatively prime natural numbers m and n. In the expression

$$f * g(mn) = \sum_{d|mn} f(d) g\left(\frac{mn}{d}\right),$$

simply observe that every divisor d of mn can be expressed as a product $d = ab$ of two relatively prime numbers a and b such that $a \mid m$, $b \mid n$, and m/a and n/b are relatively prime. It follows that

$$\begin{aligned} \sum_{d|mn} f(d) g\left(\frac{mn}{d}\right) &= \sum_{a|m;\, b|n} f(a) f(b) g(m/a) g(n/b) \\ &= \sum_{a|m} f(a) g(m/a) \sum_{b|n} f(b) g(n/b) \\ &= f * g(m) f * g(n), \end{aligned}$$

completing the proof that $f * g$ is multiplicative. Problem 9 asks the reader to check that $f^{[-1]}$ is multiplicative whenever f is multiplicative.

The forgoing is summarized by the next proposition.

Proposition 15.3.1 *The set of arithmetic functions $f(n)$ with $f(1) \neq 0$ forms a commutative group. The subset of multiplicative functions constitutes a subgroup of this group.*

Proof: See the above arguments. ■

Example 15.3.1 *The Möbius Function $\mu(n)$*

The Möbius function $\mu(n)$ is the inverse of $\mathbf{1}(n)$. We claim that

$$\mu(n) = \begin{cases} 1, & n = 1 \\ (-1)^k, & n = p_1 \cdots p_k \\ 0, & \text{otherwise,} \end{cases}$$

where $p_1, \ldots, p_k$ are distinct primes. It suffices to verify that $\mu * \mathbf{1}(n) = \delta(n)$. This is clear for $n = 1$, and for $n = p_1^{e_1} \cdots p_k^{e_k} > 1$ we have

$$\sum_{d|n} \mu(d) = \mu(1) + \sum_i \mu(p_i) + \sum_{i<j} \mu(p_i p_j) + \cdots + \mu(p_1 \cdots p_k)$$

$$\begin{aligned} &= \sum_{i=0}^{k} \binom{k}{i} (-1)^i \\ &= (1-1)^k \\ &= 0. \end{aligned}$$

The logical equivalence of the Möbius relations

$$\begin{aligned} g(n) &= \sum_{d|n} f(d) \\ f(n) &= \sum_{d|n} \mu(d) g(n/d) \end{aligned}$$

just restates the equivalence of the relations $g = \mathbf{1} * f$ and $f = \mu * g$. ■

Example 15.3.2 *Examples of Möbius Inversion*

Applying the identity $\mu * \mathbf{1}(n) \ln n = 0$, we find that equation (15.7) has inverse

$$\begin{aligned} \Lambda(n) &= \sum_{d|n} \mu(d) \ln(n/d) \\ &= -\sum_{d|n} \mu(d) \ln d. \end{aligned}$$

Euler's totient satisfies the pair of relations

$$\begin{aligned} n &= \sum_{d|n} \varphi(d) \\ \varphi(n) &= \sum_{d|n} \mu(d) \frac{n}{d}. \end{aligned}$$

To prove the first of these, note that n is obviously multiplicative and that $\mathbf{1} * \varphi(n) = \sum_{d|n} \varphi(d)$ is multiplicative by virtue of Proposition 15.3.1. If $n = p^k$ is a power of a prime, then equation (15.5) yields

$$\begin{aligned} \sum_{d|n} \varphi(d) &= \sum_{l=0}^{k} \varphi(p^l) \\ &= 1 + \sum_{l=1}^{k} p^l (1 - p^{-1}) \\ &= 1 + \sum_{l=1}^{k} (p^l - p^{l-1}) \\ &= p^k. \end{aligned}$$

Since n and $\sum_{d|n} \varphi(d)$ agree for every power of a prime, they agree for every natural number n. ■

If $f(n)$ and $g(n)$ are two arithmetic functions for which $\mathrm{E}_s[f(N)]$ and $\mathrm{E}_s[g(N)]$ exist, then

$$\left[\sum_{l=1}^{\infty} \frac{f(l)}{l^s}\right]\left[\sum_{m=1}^{\infty} \frac{g(m)}{m^s}\right] = \sum_{n=1}^{\infty} \frac{1}{n^s} \sum_{lm=n} f(l)g(m) = \sum_{n=1}^{\infty} \frac{f * g(n)}{n^s}.$$

Here all series converge absolutely. This result can be restated as

$$\mathrm{E}_s[f * g(N)] = \zeta(s)\, \mathrm{E}_s[f(N)]\, \mathrm{E}_s[g(N)].$$

For instance, Example 15.2.3 and equation (15.7) together imply

$$\begin{aligned} -\frac{d}{ds} \ln \zeta(s) &= \mathrm{E}_s(\ln N) \\ &= \zeta(s)\, \mathrm{E}_s[\Lambda(N)]\, \mathrm{E}_s(1) \\ &= \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s} \\ &= \sum_p \sum_{k=1}^{\infty} \frac{\ln p}{(p^k)^s} \\ &= \sum_p \frac{\ln p}{p^s - 1}, \end{aligned} \tag{15.8}$$

where the index p represents a generic prime.

15.4 Averages of Arithmetic Functions

Most questions in analytic number theory involve number density rather than Zipf probability. This is certainly the case for the prime number theorem [105, 197]. The Zipf distribution is easier to work with than number density because it is countably additive. Fortunately, the two notions are intimately connected. The connections can best be exposed by defining the long-run average

$$A(f) = \lim_{n \to \infty} \frac{1}{n} \sum_{m=1}^{n} f(m)$$

of an arithmetic function $f(n)$ and contrasting it to $\lim_{s \to 1} \mathrm{E}_s[f(N)]$. It turns out that in many cases these two limits coincide. When $f(n)$ is the indicator function of a set of natural numbers, this surprising insight helps in computing the number density of the set.

Before tackling this issue and how it bears on the prime number theorem, some preliminary spadework is needed. Our first result is Abel's summation formula.

Proposition 15.4.1 (Abel) *Suppose $f(n)$ is an arithmetic function and $g(t)$ is a continuously differentiable function. If $F(t) = \sum_{n=1}^{[t]} f(n)$, then*

$$\sum_{m=1}^{n} f(m)g(m) \quad = \quad F(n)g(n) - \int_1^n F(t)g'(t)\,dt.$$

Proof: With the convention $F(0) = 0$, the fundamental theorem of calculus implies

$$\begin{aligned}
\sum_{m=1}^{n} f(m)g(m) &= \sum_{m=1}^{n} [F(m) - F(m-1)]g(m) \\
&= \sum_{m=1}^{n} F(m)g(m) - \sum_{m=1}^{n-1} F(m)g(m+1) \\
&= F(n)g(n) - \sum_{m=1}^{n-1} F(m)[g(m+1) - g(m)] \\
&= F(n)g(n) - \sum_{m=1}^{n-1} F(m) \int_m^{m+1} g'(t)\,dt \\
&= F(n)g(n) - \sum_{m=1}^{n-1} \int_m^{m+1} F(t)g'(t)\,dt \\
&= F(n)g(n) - \int_1^n F(t)g'(t)\,dt.
\end{aligned}$$

This completes the proof. ∎

It is fairly obvious that Riemann's zeta function $\zeta(s)$ can be extended to the complex domain $\mathsf{Re}(s) > 1$. However, as s tends to 1 along the real axis, $\zeta(s)$ tends to ∞. Our next result says that this singularity is removable and the domain can be enlarged to $\mathsf{Re}(s) > 0$.

Proposition 15.4.2 *The difference $\zeta(s) - (s-1)^{-1}$ can be analytically continued to the domain $\mathsf{Re}(s) > 0$. Consequently, $\lim_{s\to 1}(s-1)\zeta(s) = 1$.*

Proof: Let us write the difference as

$$\begin{aligned}
\zeta(s) - \frac{1}{s-1} &= \sum_{n=1}^{\infty} \frac{1}{n^s} - \int_1^{\infty} \frac{1}{t^s} dt \\
&= \sum_{n=1}^{\infty} \int_n^{n+1} \left(\frac{1}{n^s} - \frac{1}{t^s} \right) dt.
\end{aligned}$$

For each n the integral

$$g_n(s) \;=\; \int_n^{n+1} \left(\frac{1}{n^s} - \frac{1}{t^s}\right) dt \;=\; s\int_n^{n+1}\int_n^t \frac{1}{u^{s+1}} du\, dt$$

is clearly analytic in s. The estimate

$$|g_n(s)| \;\le\; \frac{|s|}{2} \sup_{n\le u\le n+1} \left|\frac{1}{u^{s+1}}\right| \;=\; \frac{|s|}{2}\frac{1}{n^{\mathrm{Re}(s)+1}}$$

shows that the series $\sum_{n=1}^{\infty} g_n(s)$ converges uniformly and absolutely on every compact set of the domain $\mathsf{Re}(s) > 0$. Therefore, its limit is an analytic function throughout this domain. ■

We are now in position to establish one connection between $A(f)$ and the $\mathrm{E}_s[f(N)]$.

Proposition 15.4.3 *Let $f(n)$ be an arithmetic function such that $A(f)$ and all $\mathrm{E}_s[f(N)]$ exist. Then $\lim_{s\to 1} E_s[f(N)]$ exists and equals $A(f)$.*

Proof: Let $F(t) = \sum_{n=1}^{[t]} f(n)$ and $G(t) = t^{-1}F(t)$. Abel's summation formula gives

$$\begin{aligned}\sum_{m=1}^{n} \frac{f(m)}{m^s} &= \frac{1}{n^{s-1}}\frac{F(n)}{n} + \int_1^n \frac{F(t)}{t}\frac{s}{t^s}\, dt.\\ &= \frac{G(n)}{n^{s-1}} + s\int_0^n \frac{G(t)}{t^s}\, dt.\end{aligned}$$

Taking limits on n and invoking the existence of $A(f)$ therefore imply

$$\sum_{m=1}^{\infty} \frac{f(m)}{m^s} \;=\; s\int_0^{\infty} \frac{G(t)}{t^s}\, dt$$

for each $s > 1$. In view of Proposition 15.4.2, it now suffices to prove that

$$\lim_{s\to 1}(s-1)\sum_{m=1}^{\infty} \frac{f(m)}{m^s} \;=\; \lim_{s\to 1}(s-1)s\int_0^{\infty} \frac{G(t)}{t^s}\, dt \;=\; A(f).$$

To achieve this end, we exploit the integral $(s-1)\int_1^{\infty} t^{-s}dt = 1$ and the inequality

$$\begin{aligned}\left|(s-1)s\int_0^{\infty} \frac{G(t)}{t^s}dt - sA(f)\right| &= \left|(s-1)s\int_0^{\infty} \frac{G(t)-A(f)}{t^s}dt\right|\\ &\le (s-1)s\int_1^{\infty} \left|\frac{G(t)-A(f)}{t^s}\right| dt.\end{aligned}$$

For $\epsilon > 0$ small, choose $\delta \geq 1$ so that $|G(t) - A(f)| < \epsilon$ for $t \geq \delta$. This permits us to form the bound

$$\begin{aligned}(s-1)s\int_1^\infty \left|\frac{G(t)-A(f)}{t^s}\right| dt &\leq (s-1)s\int_1^\delta \left|\frac{G(t)-A(f)}{t^s}\right| dt \\ &\quad +\epsilon(s-1)s\int_\delta^\infty \frac{1}{t^s}\,dt \\ &= (s-1)s\int_1^\delta \left|\frac{G(t)-A(f)}{t^s}\right| dt + \epsilon s\delta^{-s+1}.\end{aligned}$$

For s sufficiently close to 1, this bound can be made less than 2ϵ. To finish the proof, we merely note that $\lim_{s\to 1}(s-1)A(f) = 0$. ■

Example 15.4.1 *Periodic Arithmetic Functions*

A periodic arithmetic function $f(n)$ satisfies $f(n+r) = f(n)$ for all n and some fixed r. A brief calculation shows that

$$A(f) = \frac{1}{r}\sum_{n=1}^{r} f(n).$$

Proposition 15.4.3 guarantees that $\lim_{s\to 1} \mathrm{E}_s[f(N)] = A(f)$. We can also deduce this result by bringing in Hurwitz's zeta function

$$\zeta(s,a) = \sum_{n=0}^{\infty} \frac{1}{(n+a)^s}$$

for $a > 0$. The techniques of Proposition 15.4.2 demonstrate that

$$\zeta(s,a) - \frac{1}{(s-1)a^{s-1}} \tag{15.9}$$

is analytic throughout $\mathsf{Re}(s) > 0$ and satisfies $\lim_{s\to 1}(s-1)\zeta(s,a) = 1$. It follows that

$$\begin{aligned}\frac{1}{\zeta(s)}\sum_{n=1}^{\infty}\frac{f(n)}{n^s} &= \frac{1}{\zeta(s)}\sum_{n=1}^{r} f(n)\sum_{m=0}^{\infty}\frac{1}{(mr+n)^s} \\ &= \frac{1}{(s-1)\zeta(s)}\sum_{n=1}^{r} f(n)\frac{s-1}{r^s}\zeta(s,n/r).\end{aligned}$$

Setting $a = n/r$ and sending s to 1 now produce the desired limit $A(f)$. ■

The Riemann hypothesis, arguably the most famous unsolved problem in mathematics, says all zeros of $\zeta(s)$ lie on the line $\mathsf{Re}(s) = \frac{1}{2}$. For our purposes, the following simpler result suffices.

Proposition 15.4.4 *The line* $\mathsf{Re}(s) = 1$ *contains no zeros of* $\zeta(s)$.

Proof: We present the clever proof of Mertens, which hinges on the trigonometric identity

$$3 + 4\cos\theta + \cos 2\theta \;=\; 2(1+\cos\theta)^2 \;\geq\; 0.$$

Equation (15.1) yields

$$\zeta(s) \;=\; e^{-\sum_p \ln(1-p^{-s})} \;=\; e^{\sum_p \sum_{k=1}^{\infty} k^{-1}p^{-ks}},$$

so if $s = a + bi$, then

$$|\zeta(s)| \;=\; e^{\sum_p \sum_{k=1}^{\infty} k^{-1}p^{-ka}\cos(kb\ln p)}.$$

Hence,

$$\begin{aligned} & |\zeta(a)|^3|\zeta(a+bi)|^4|\zeta(a+2bi)| \\ = \;& e^{\sum_p \sum_{k=1}^{\infty} k^{-1}p^{-ka}[3+4\cos(kb\ln p)+\cos(2kb\ln p)]} \\ \geq \;& 1. \end{aligned}$$

Dividing this inequality by $a-1$ produces

$$|(a-1)\zeta(a)|^3 \left|\frac{\zeta(a+bi)}{a-1}\right|^4 |\zeta(a+2bi)| \;\geq\; \frac{1}{a-1}.$$

If we suppose that $\zeta(1+bi) = 0$ with $b \neq 0$, then the left-hand side of this inequality approaches $1 \cdot |\zeta'(1+bi)|^4|\zeta(1+2bi)|$ as a tends to 1 while the right-hand side approaches ∞. This contradiction proves that $\zeta(1+bi) = 0$ is impossible. ■

15.5 The Prime Number Theorem

Let $\pi(n)$ be the number of primes less than or equal to n. The prime number theorem says

$$\lim_{n\to\infty} \frac{\pi(n)\ln n}{n} \;=\; 1.$$

We would like to rephrase this celebrated result as a long-run arithmetic average involving the summatory function $\vartheta(t) = \sum_{p\leq t} \ln p$. Here p denotes a generic prime, and noninteger values of t are permitted. On one hand, it is clear that

$$\vartheta(t) \;\leq\; \sum_{p\leq t} \ln t \;=\; \pi(t)\ln t.$$

On the other hand for any $\epsilon \in (0,1)$,

$$\begin{aligned}\vartheta(t) &\geq \sum_{t^{1-\epsilon} \leq p \leq t} \ln p \\ &\geq \sum_{t^{1-\epsilon} \leq p \leq t} (1-\epsilon) \ln t \\ &= (1-\epsilon) \ln t [\pi(t) + O(t^{1-\epsilon})].\end{aligned}$$

These two estimates make it evident that the prime number theorem is equivalent to the equality $A(\Lambda) = \lim_{n\to\infty} n^{-1}\vartheta(n) = 1$.

The proof of the prime number theorem depends on several technical propositions.

Proposition 15.5.1 *The function $\vartheta(t) = \sum_{p\leq t} \ln p$ satisfies the inequality $0 \leq \vartheta(t) \leq ct$ for $c = 4 \ln 2$.*

Proof: For each positive integer n, the binomial coefficient $\binom{2n}{n}$ is divisible by all primes p on the interval $(n, 2n]$. It follows that

$$\prod_{n<p\leq 2n} p \leq \binom{2n}{n} < 2^{2n}.$$

For the choice $2n = 2^k$, this implies

$$\sum_{2^{k-1}<p\leq 2^k} \ln p \leq 2^k \ln 2.$$

Summing this inequality on k produces

$$\sum_{p\leq 2^k} \ln p \leq (2^k + 2^{k-1} + \cdots + 1) \ln 2 < 2^{k+1} \ln 2.$$

If $2^{k-1} < t \leq 2^k$, then the inequalities

$$\frac{1}{t} \sum_{p\leq t} \ln p \leq \frac{1}{2^{k-1}} \sum_{p\leq 2^k} \ln p \leq \frac{1}{2^{k-1}} 2^{k+1} \ln 2$$

complete the proof. ∎

The prime number theorem can be deduced from a more general proposition giving a sufficient condition for the existence of an arithmetic average $A(f)$. In deriving this result, we will rely on an expanded definition of the integral on the real line known as the gauge integral or generalized Riemann integral [141, 210]. The gauge integral subsumes both the Lebesgue integral and the improper integrals met in advanced calculus. The integrands of the gauge integral are not necessarily absolutely integrable. An integral

of a function $g(t)$ over an infinite interval such as $[0,\infty)$ exists if and only if $\int_0^u g(t)\,dt$ exists for all finite $u \geq 0$ and

$$\lim_{u\to\infty} \int_0^u g(t)\,dt \quad = \quad \int_0^\infty g(t)\,dt.$$

We will employ the following Tauberian result on Laplace transforms.

Proposition 15.5.2 *Let the bounded function $g(t)$ be integrable over every finite interval $[0,u]$. If its Laplace transform*

$$\hat{g}(s) \quad = \quad \int_0^\infty g(t)e^{-st}dt$$

exists for all $s > 0$, is analytic throughout the domain $\mathsf{Re}(s) > 0$, and can be continued analytically to a neighborhood of every point on the imaginary axis $\mathsf{Re}(s) = 0$, then the integral $\int_0^\infty g(t)\,dt$ also exists and equals $\hat{g}(0)$.

Proof: Appendix A.7 reproduces the proof given in the references [150, 211]. ■

With these preliminaries under our belt, we state our general proposition.

Proposition 15.5.3 *Suppose $f(n)$ is a nonnegative arithmetic function such that $\mathrm{E}_s[f(N)]$ exists for all $s > 1$, the averages $\frac{1}{n}\sum_{m=1}^n f(m)$ are bounded in n, and for some constant a*

$$h(s) \quad = \quad \frac{1}{s}\sum_{n=1}^\infty \frac{f(n)}{n^s} - \frac{a}{s-1} \tag{15.10}$$

is analytic throughout the domain $\mathsf{Re}(s) > 1$ and can be continued analytically to a neighborhood of every point on the vertical line $\mathsf{Re}(s) = 1$. Then both $\lim_{s\to 1}\mathrm{E}_s[f(N)]$ and $A(f)$ exist and equal a.

Proof: The identity

$$h(s) \quad = \quad \frac{1}{s(s-1)}\left[(s-1)\sum_{n=1}^\infty \frac{f(n)}{n^s} - a\right] - \frac{a}{s}$$

and the continuity of $h(s)$ at $s = 1$ jointly show that

$$\lim_{s\to 1}(s-1)\sum_{n=1}^\infty \frac{f(n)}{n^s} \quad = \quad a.$$

Hence, Proposition 15.4.2 implies

$$\begin{aligned}\lim_{s\to 1}\mathrm{E}_s[f(N)] &= \lim_{s\to 1}\frac{s-1}{(s-1)\zeta(s)}\sum_{n=1}^\infty \frac{f(n)}{n^s} \\ &= \lim_{s\to 1}\frac{a}{(s-1)\zeta(s)} \\ &= a.\end{aligned}$$

Now let $F(t) = \sum_{m=1}^{[t]} f(m)$. For $s > 1$ Abel's summation formula implies

$$\sum_{m=1}^{n} \frac{f(m)}{m^s} = \frac{F(n)}{n^s} + s\int_1^n \frac{F(t)}{t^{s+1}}\,dt.$$

Because the averages $\frac{1}{n}F(n)$ are bounded, this gives in the limit

$$\sum_{m=1}^{\infty} \frac{f(m)}{m^s} = s\int_1^\infty \frac{F(t)}{t^{s+1}}\,dt.$$

Dividing by s and subtracting $a\int_1^\infty t^{-s}dt = a(s-1)^{-1}$ therefore yield

$$\begin{aligned}\frac{1}{s}\sum_{n=1}^{\infty}\frac{f(n)}{n^s} - \frac{a}{s-1} &= \int_1^\infty \left[\frac{1}{t}F(t) - a\right]\frac{1}{t^s}\,dt \\ &= \int_0^\infty \left[e^{-u}F(e^u) - a\right]e^{-(s-1)u}\,du\end{aligned}$$

We now apply Proposition 15.5.2 to the function $g(u) = e^{-u}F(e^u) - a$, which is bounded by assumption. The proposition guarantees the existence of the integral $\int_1^\infty \left[t^{-1}F(t) - a\right]t^{-1}dt$.

To use this conclusion, suppose that $u^{-1}F(u) - a \geq \delta > 0$ for arbitrarily large u. Then with $\lambda > 1$ chosen so that $a\lambda \leq a + \delta$, we have for such u

$$\begin{aligned}\int_u^{\lambda u}\left[\frac{1}{t}F(t) - a\right]\frac{1}{t}\,dt &= \int_u^{\lambda u}\frac{1}{t^2}[F(t) - at]\,dt \\ &\geq \int_u^{\lambda u}\frac{1}{t^2}[F(u) - at]\,dt \\ &\geq \int_u^{\lambda u}\frac{1}{t^2}[au + \delta u - at]\,dt \\ &= \int_1^{\lambda}\frac{1}{(ur)^2}[au + \delta u - aur]u\,dr \\ &= \int_1^{\lambda}\frac{1}{r^2}[a + \delta - ar]\,dr \\ &> 0.\end{aligned}$$

The fact that $\int_u^{\lambda u}\left[t^{-1}F(t) - a\right]t^{-1}dt$ does not converge to 0 as u tends to ∞ contradicts the existence of $\int_1^\infty \left[t^{-1}F(t) - a\right]t^{-1}dt$. Hence, the inequality $u^{-1}F(u) - a < \delta$ must hold for all sufficiently large u.

One can likewise demonstrate that $u^{-1}F(u) - a > -\delta$ for all large u by assuming the contrary. If u is a point where $u^{-1}F(u) - a \leq -\delta$, then choose $0 < \theta < 1$ so that $a - \delta < a\theta$. The inequality

$$\int_{\theta u}^{u}\left[\frac{1}{t}F(t) - a\right]\frac{1}{t}\,dt \leq \int_\theta^1 \frac{1}{r^2}\Big[a - \delta - ar\Big]dr$$

for arbitrarily large u again contradicts the existence of the integral

$$\int_1^\infty \left[\frac{1}{t}F(t) - a\right] t^{-1} dt$$

and completes the proof of the proposition. ∎

Example 15.5.1 *Application to the Prime Number Theorem*

Let us verify that the assumptions of Proposition 15.5.3 hold for the function

$$f(n) \quad = \quad \begin{cases} \ln n & n \text{ is prime} \\ 0 & \text{otherwise.} \end{cases}$$

This function is clearly nonnegative, and according to Proposition 15.5.1 its running averages $n^{-1}\vartheta(n)$ are bounded. The remaining hypothesis is more delicate. The key to its verification is to write

$$\begin{aligned} \sum_p \frac{\ln p}{p^s} &= \sum_p \frac{\ln p}{p^s - 1} - \sum_p \frac{\ln p}{p^s(p^s - 1)} \\ &= -\frac{\zeta'(s)}{\zeta(s)} - \sum_p \frac{\ln p}{p^s(p^s - 1)} \end{aligned}$$

using equation (15.8). The comparisons

$$\sum_p \left| \frac{\ln p}{p^s(p^s - 1)} \right| \le \sum_p \frac{\ln p}{p^{\mathsf{Re}(s)}(p^{\mathsf{Re}(s)} - 1)} \le \sum_{n=2}^\infty \frac{\ln n}{n^{\mathsf{Re}(s)}(n^{\mathsf{Re}(s)} - 1)}$$

show that the series $\sum_p \ln p/[p^s(p^s - 1)]$ is analytic for $\mathsf{Re}(s) > \frac{1}{2}$.

Consider the difference

$$\frac{1}{s}\sum_p \frac{\ln p}{p^s} - \frac{1}{s-1} = -\frac{\zeta'(s)}{s\zeta(s)} - \frac{1}{s-1} - \frac{1}{s}\sum_p \frac{\ln p}{p^s(p^s - 1)},$$

and set $\eta(s) = \zeta(s) - (s-1)^{-1}$. It suffices to demonstrate that

$$\begin{aligned} -\frac{\zeta'(s)}{s\zeta(s)} - \frac{1}{s-1} &= -\frac{(s-1)\zeta'(s) + s\zeta(s)}{s(s-1)\zeta(s)} \\ &= -\frac{(s-1)\eta'(s) + s\eta(s) + 1}{s(s-1)\zeta(s)} \end{aligned}$$

is analytic in a neighborhood of any point s with $\mathsf{Re}(s) = 1$. According to Proposition (15.4.2), both the numerator and the denominator are analytic for $\mathsf{Re}(s) > 0$. The denominator is nonzero on the point $s = 1$ because of Proposition 15.4.2. At all other points of the vertical line $\mathsf{Re}(s) = 1$, Proposition 15.4.4 guarantees that $\zeta(s) \neq 0$. This completes the proof of the prime number theorem. ∎

15.6 Problems

1. Demonstrate that

$$\begin{aligned}\prod_{n=2}^{\infty}\left(1-\frac{1}{n^2}\right) &= \frac{1}{2}\\ \prod_{n=0}^{\infty}\left(1+z^{2^n}\right) &= \frac{1}{1-z}, \qquad |z|<1.\end{aligned}$$

2. Show that the gap g between successive prime numbers can be arbitrarily large. (Hint: Consider the sequence of numbers beginning with $(g+1)!+2$.)

3. Let N follow Zipf's distribution. Demonstrate that

$$\begin{aligned}\mathrm{E}_s(N^k) &= \frac{\zeta(s-k)}{\zeta(s)}\\ \mathrm{E}_s(N^k \mid q \text{ divides } N) &= \frac{q^k\zeta(s-k)}{\zeta(s)}\\ \mathrm{E}_s(N^k \mid N \text{ prime}) &= \frac{\sum_p p^{k-s}}{\sum_p p^{-s}}\end{aligned}$$

   for k a natural number with $s > k+1$.

4. Choose two independent random numbers M and N according to Zipf's distribution. Prove that M and N are relatively prime with probability $\zeta(2s)^{-1}$. (Hints: Let X_p and Y_p be the powers of the prime p occurring in M and N, respectively. Calculate $\mathrm{E}_s(\prod_p 1_{B_p})$, where B_p is the event $\{X_pY_p = 0\}$.)

5. Suppose the arithmetic function $Y = f(N)$ satisfies $\mathrm{E}_s(Y) = 0$ for all $s \geq r > 1$. Show that $Y \equiv 0$. (Hint: Prove that $f(n) = 0$ for all n by induction and sending s to ∞ in the equation $n^s\,\mathrm{E}_s(Y) = 0$.)

6. Let $N_1, \ldots, N_m$ be an i.i.d. sample from the Zipf distribution with values $n_1, \ldots, n_m$. If at least one $n_i > 1$, then prove that the maximum likelihood estimate of s is uniquely determined by the equation

$$-\frac{d}{ds}\ln\zeta(s) = \mathrm{E}_s(\ln N) = \frac{1}{m}\sum_{i=1}^{m}\ln n_i.$$

   What happens in the exceptional case when all $n_i = 1$?

7. Suppose the independent realizations M and N of Zipf's distribution generate arithmetic functions $Y = f(M)$ and $Z = g(N)$ with finite

expectations. Show that the random variables $L = MN$ and $W = YZ$ satisfy

$$\begin{aligned} \Pr(L = l) &= \frac{\tau(l)}{l^s \zeta(s)^2} \\ \mathrm{E}_s(W \mid L = l) &= \tau(l)^{-1} f * g(l). \end{aligned}$$

Recall that $\tau(l)$ is the number of divisors of l. Use these results to demonstrate that $\mathrm{E}_s(W) = \mathrm{E}_s(Y)\,\mathrm{E}_s(Z)$ entails

$$\sum_{l=1}^{\infty} \frac{f * g(l)}{l^s} = \left[\sum_{m=1}^{\infty} \frac{f(m)}{m^s}\right]\left[\sum_{n=1}^{\infty} \frac{g(n)}{n^s}\right].$$

8. Prove the two inequalities (15.6). (Hints: Show that the function $f(t) = t^{-s} \ln t$ is decreasing on the interval $[e^{1/s}, \infty)$. Use this to prove that

$$\begin{aligned} 2\frac{\ln(2n)}{(2n)^s} &\leq \frac{\ln(2n-1)}{(2n-1)^s} + \frac{\ln(2n)}{(2n)^s} \\ 2\frac{\ln(2n)}{(2n)^s} &\geq \frac{\ln(2n+1)}{(2n+1)^s} + \frac{\ln(2n)}{(2n)^s} \end{aligned} \tag{15.11}$$

for appropriate values of n and s. Use these to prove that

$$\begin{aligned} \sum_{n=1}^{\infty} \frac{\ln n}{n^s} &\geq 2\sum_{n=1}^{\infty} \frac{\ln(2n)}{(2n)^s} - \frac{\ln 2}{2^s} \\ \sum_{n=1}^{\infty} \frac{\ln n}{n^s} &\leq 2\sum_{n=1}^{\infty} \frac{\ln(2n)}{(2n)^s}. \end{aligned}$$

Note that the second inequality in (15.11) can fail when $n = 1$ and $s < 1/\ln 2$. In this special case, apply the more delicate inequality

$$\frac{\ln 1}{1^s} + \frac{\ln 2}{2^s} + \frac{\ln 3}{3^s} + \frac{\ln 4}{4^s} + \frac{\ln 5}{5^s} + \quad \leq \quad 2\left(\frac{\ln 2}{2^s} + \frac{\ln 4}{4^s}\right),$$

which can be demonstrated by showing that

$$g(s) = \frac{\ln 2}{2^s} - \frac{\ln 3}{3^s} + \frac{\ln 4}{4^s} - \frac{\ln 5}{5^s}$$

is positive. For this purpose, write $g(s)$ in the equivalent form

$$g(s) = \frac{1}{(s-1)^2}\left(\int_2^3 \frac{1 + \ln t^{s-1}}{t^{s-1}} dt - \int_4^5 \frac{1 + \ln t^{s-1}}{t^{s-1}} dt\right)$$

and prove that the indicated integrand is decreasing in t.)

9. Check that the Dirichlet inverse $f^{[-1]}$ of a multiplicative arithmetic function f is multiplicative. (Hints: Assume otherwise, and consider the least product mn of two relatively prime natural numbers m and n with $f^{[-1]}(mn) \neq f^{[-1]}(m)f^{[-1]}(n)$. Now apply the inductive definition of $f^{[-1]}(mn)$.)

10. Verify that the identities

$$\begin{aligned} 1 &= \sum_{d|n} \mu(d)\tau(n/d) \\ n &= \sum_{d|n} \mu(d)\sigma(n/d) \end{aligned}$$

hold for all natural numbers n.

11. Demonstrate that neither $\varphi(n)$ nor $\mu(n)$ is completely multiplicative. Show that the Dirichlet convolution of two completely multiplicative functions need not be completely multiplicative. (Hint: $\sigma = id * \mathbf{1}$.)

12. Liouville's arithmetic function is defined by

$$\lambda(n) = \begin{cases} 1, & n = 1 \\ (-1)^{e_1+\cdots+e_k}, & n = p_1^{e_1} \cdots p_k^{e_k}. \end{cases}$$

Prove that

$$\sum_{d|n} \lambda(d) = \begin{cases} 1 & \text{if } n \text{ is a square} \\ 0 & \text{otherwise.} \end{cases}$$

13. Suppose q is a positive integer and c_m is a sequence of numbers with $c_{m+q} = c_m$ for all m. Prove that the series $\sum_{m=1}^{n} \frac{c_m}{m}$ converges if and only if $\sum_{m=1}^{q} c_m = 0$. (Hint: Apply Proposition 13.4.1.)

14. If $\lim_{n\to\infty} f(n) = c$, then demonstrate that

$$A(f) = \lim_{s\to 1} \mathrm{E}_s[f(N)] = c.$$

15. Derive formula (15.9).

16. In Proposition 15.5.3, show that the assumption that $f(n)$ is nonnegative can be relaxed to $\inf_n f(n) > -\infty$ or $\sup_n f(n) < \infty$.

17. Apply Proposition 15.5.3 and prove that $A[\varphi(N)/N] = \zeta(2)^{-1}$.

Appendix: Mathematical Review

A.1 Elementary Number Theory

Our first result is the standard division algorithm.

Proposition A.1.1 *Given two integers a and b with b positive, there is a unique pair of integers q and r such that $a = qb + r$ and $0 \leq r < b$.*

Proof: Consider the set $S = \{a - nb : n \text{ an integer}\}$. This set contains nonnegative elements such as $a+|a|b$, so there is a least nonnegative element $r = a - qb$. If $r \geq b$, then $r - b = a - (q + 1)b$ is nonnegative and belongs to S, contradicting the choice of r. Hence, $0 \leq r < b$. If $a = q'b + r'$ is a different representation of a, then $r - r' = (q' - q)b$. But this is absurd because $|r - r'| < b$ while $|q' - q|b > q$. ■

The next proposition refers to an additive group. This is just a nonempty set of integers closed under both addition and subtraction. An additive group contains 0 and possibly other elements as well.

Proposition A.1.2 *An additive group G of integers can be represented as $G = \{nd : n \text{ is an integer}\}$ for some nonnegative integer d.*

Proof: If G consists of 0 alone, then take $d = 0$. Otherwise, consider $g \neq 0$ in G. Because either g or $-g$ is positive, G contains at least one positive element. Let d be the smallest positive element. Given that G is closed under addition and subtraction, all multiples of d belong to G. It remains to prove that every element g of G is a multiple of d. This follows from the representation $g = qd + r$, where $0 \leq r < d$. If $r = 0$, then we are done. If

K. Lange, *Applied Probability,*
DOI 10.1007/978-1-4419-7165-4, © Springer Science+Business Media, LLC 2010

$r > 0$, then the equation $r = g - qd$ identifies r as a positive element of G less than d, contradicting the choice of d. ■

The integer d is said to be a divisor of the integer a if $a = qd$ for some integer q.

Proposition A.1.3 *The greatest common divisor d of a set of integers $\{n_1, \ldots, n_m\}$ can be expressed as a linear combination $d = \sum_{k=1}^{m} c_k n_k$ with integer coefficients $c_1, \ldots, c_m$.*

Proof: The set G of such linear combinations clearly forms an additive group. Let e be the generator of G identified in Proposition A.1.2. On the one hand, because e equals a linear combination of the n_k, the greatest common divisor d divides e. On the other hand, because each n_k belongs to G, e divides d. It follows that $e = d$. ■

Proposition A.1.4 *Suppose S is an infinite set of positive integers closed under addition. If d is the greatest common divisor of S, then S contains all but a finite number of positive multiples of d.*

Proof: Note that S is not a group because it is not closed under subtraction. We can assume that $d = 1$ by dividing all elements of S by d if necessary. Given the assumption that $d = 1$, take integers $n_1, \ldots, n_m$ from S lacking any common factor. According to Proposition A.1.3, there exists integer coefficients $c_1, \ldots, c_m$ such that $1 = \sum_{k=1}^{m} c_k n_k$. The two integers

$$p = \sum_{k:c_k>0} c_k n_k, \qquad q = -\sum_{k:c_k<0} c_k n_k$$

clearly belong to S, and $1 = p - q$. Let us show that any number $r \geq q(q-1)$ belongs to S. If we write $r = aq + b$, where $0 \leq b < q$, then it follows that $a \geq q - 1 \geq b$. The representation

$$r = aq + b(p - q) = (a - b)q + bp$$

now expresses r as an element of S. ■

An integer $p > 1$ is said to be prime if the only positive divisors of p are 1 and p itself.

Proposition A.1.5 *If a prime p divides the product ab, then p divides a or b.*

Proof: Suppose that p does not divide a. Then the greatest common divisor of p and a is 1. According to Proposition A.1.3, there exist integers c and d such that $1 = ca + dp$. Multiplying this equality by b produces $b = cab + dpb$. Since p divides both cab and dpb, it must divide b. ■

Our final result is the well-known fundamental theorem of arithmetic.

Proposition A.1.6 *Every integer $n > 1$ has a prime power factorization*

$$n = p_1^{m_1} \cdots p_k^{m_k},$$

where p_1 through p_k are prime numbers and m_1 through m_k are positive integers. This factorization is unique up to permutation of the factors.

Proof: The proof is by induction on n. The assertion is obvious for $n = 2$. Therefore consider a general $n > 2$, and assume the assertion is true for all positive integers between 2 and $n - 1$. If n is prime, then we are done. Otherwise, $n = ab$ for $a > 1$ and $b > 1$. By the induction hypothesis, a and b can both be represented as a product of primes. It follows that n can be so represented. To prove uniqueness, suppose

$$n = p_1 p_2 \cdots p_s = q_1 q_2 \cdots q_t$$

are two prime factorizations of n. According to Proposition A.1.5, p_1 either divides q_1 or the product $q_2 \cdots q_t$. In the former case, $p_1 = q_1$ because both are primes. Otherwise, p_1 divides $q_2 \cdots q_t$. Again either p_1 equals q_2 or divides $q_3 \cdots q_t$. Eventually we exhaust the partial products in the second factorization. Thus, $p_1 = q_k$ for some k, and reordering the primes in the second factorization if necessary, we can take $k = 1$. If we apply the uniqueness part of the induction hypothesis to

$$\frac{n}{p_1} = p_2 \cdots p_s = q_2 \cdots q_t,$$

then we can conclude that $s = t$ and the remaining prime factors agree up to order. ■

A.2 Nonnegative Matrices

In this section we survey the theory of nonnegative matrices and prove the celebrated Perron-Frobenius theorem [99, 106, 179, 180]. This theorem deals with matrices $M = (m_{ij})$ that are square, nonnegative, and irreducible. Irreducibility means that given any two indices i and j, there is a path $k_1, \ldots, k_m$ from $k_1 = i$ to $k_m = j$ such that $m_{k_1 k_2} \cdots m_{k_{m-1} k_m} > 0$. Irreducibility has some immediate consequences. For example, the transpose M^t of $M \geq \mathbf{0}$ is irreducible whenever M is irreducible. Somewhat more useful is the fact that the sum of powers $S = \sum_{k=1}^{q} M^k$ has all entries positive when q is large enough.

Most of the theory erected by Perron and Frobenius carries over to matrices whose off-diagonal entries are nonnegative but whose diagonal entries may have either sign. A square matrix M is said to be a Metzler-Leontief matrix if $M + \mu I$ is nonnegative and irreducible for some constant $\mu \geq 0$. The modern extensions of the Perron-Frobenius theorem assert

that a Metzler-Leontief matrix has a dominant eigenvalue with an essentially unique eigenvector having positive entries. By dominant eigenvalue we mean an eigenvalue with largest real part.

It turns out that the dominant eigenvalue ρ of a Metzler-Leontief $n \times n$ matrix M is given by the formula

$$\begin{aligned} \rho &= \sup\{\lambda : Mv \geq \lambda v, \text{ for some } v \geq \mathbf{0},\ v \neq \mathbf{0}\} \\ &= \sup\{\lambda : Mv \geq \lambda v, \text{ for some } v \in T_{n-1}\}, \end{aligned} \tag{A.1}$$

where all vector inequalities are interpreted entry by entry and the set $T_{n-1} = \{v \in \mathsf{R}^n_+ : \|v\|_1 = 1\}$ is the unit simplex equipped with the standard ℓ_1 norm $\|v\|_1 = \sum_{i=1}^n |v_i| = \sum_{i=1}^n v_i$. The formula (A.1) is stated for column vectors. It can be restated for row vectors, and all subsequent theoretical results are valid under this substitution. Observe that the row version of ρ and the column version of ρ coincide because M and M^t share all eigenvalues.

We will prove the next few propositions under the assumption that the matrix M is nonnegative and irreducible. The general case is recovered by adding a large enough positive multiple μI of the identity to M so that $M+\mu I$ is nonnegative and irreducible. All but one claim of the propositions deduced for $M + \mu I$ remain true for M since $\rho(M) = \rho(M + \mu I) - \mu$ and inequalities such as $Mv \geq \lambda v$ and $(M+\mu I)v \geq (\lambda+\mu)v$ are equivalent. The exception is the claim in Proposition A.2.3 that the dominant eigenvalue also dominates all other eigenvalues in absolute value.

Our first proposition highlights the importance of irreducibility.

Proposition A.2.1 *Let M be a Metzler-Leontief matrix. Suppose that $Mv \geq \lambda v$ and $Mv \neq \lambda v$, where $v \geq \mathbf{0}$ and $v \neq \mathbf{0}$. Then there exists a number $\gamma > \lambda$ and a vector w with $w \geq \mathbf{0}$ and $w \neq \mathbf{0}$ such that $Mw \geq \gamma w$. Similarly, if $Mv \leq \lambda v$ and $Mv \neq \lambda v$, then there exists a number $\gamma < \lambda$ and an appropriate vector w such that $Mw \leq \gamma w$.*

Proof: Assume that M is nonnegative. If $Mv \geq \lambda v$ and $Mv \neq \lambda v$, then we multiply the vector $Mv - \lambda v \geq \mathbf{0}$ by the sum $S = \sum_{k=1}^q M^k$ chosen so that S has all entries positive. The vector $w = Sv$ satisfies the strict inequality $Mw > \lambda w$, and we can choose $\epsilon > 0$ sufficiently small so that $Mw > (\lambda + \epsilon)w$. The second assertion is demonstrated similarly. ■

Proposition A.2.2 *Suppose M is a Metzler-Leontief matrix. Then the real number ρ defined by formula (A.1) is an eigenvalue of M with a corresponding eigenvector v having positive entries. When M is nonnegative, ρ is positive. Furthermore, any other eigenvector u associated with ρ is a multiple of v.*

Proof: Under the assumption that M is nonnegative, the set

$$\mathcal{A} = \{\lambda : Mv \geq \lambda v, \text{ for some } v \in T_{n-1}\}$$

clearly contains the number 0. Let

$$\|M\|_1 \quad = \quad \max_{1\le j\le n} \sum_{i=1}^{n} m_{ij}$$

denote the matrix norm induced by the ℓ_1 vector norm. If $Mv \ge \lambda v$ and $\lambda \ge 0$, then the inequality

$$\lambda\|v\|_1 \quad \le \quad \|M\|_1\|v\|_1$$

shows that $\lambda \le \|M\|_1$. Hence, $\mathcal{A}$ is bounded above by $\|M\|_1$, and its supremum ρ exists. Let λ_n be a sequence of scalars and v_n a sequence of vectors in T_{n-1} such that λ_n tends to ρ and $Mv_n \ge \lambda_n v_n$. Since T_{n-1} is compact, by passing to a subsequence if necessary, we can assume that v_n converges to a vector v in T_{n-1}. An appeal to continuity shows that $Mv \ge \rho v$. If $Mv \ne \rho v$, then application of Proposition A.2.1 produces a substitute vector w with $Mw \ge \gamma w$ for $\gamma > \rho$. But this contradicts the definition of ρ, and we are forced to conclude that $Mv = \rho v$. Finally, the equality $M^k v = \rho^k v$, valid for all positive integers k, demonstrates that $(\rho + \cdots + \rho^q)v = Sv > \mathbf{0}$ and therefore that ρ and all entries of v are positive.

Consider another eigenvector $u \ne v$ associated with ρ. If u is complex, then the real and imaginary parts of u are also eigenvectors associated with ρ. To prove that u is a multiple of v, it suffices to prove that both of these are real multiples of v. Thus, we take u to have real entries and define the eigenvector $w = v - tu$ for a scalar t. Given that the entries of v are positive, we can take $t \ne 0$ so that $w \ge \mathbf{0}$ and at least one entry w_i is 0. If $w = \mathbf{0}$ we are done; otherwise applying the matrix S to w leads to the strict inequality $(\rho + \cdots + \rho^q)w = Sw > \mathbf{0}$, contradicting the assumption $w_i = 0$. Hence, $w = \mathbf{0}$ and $u = t^{-1}v$. ■

Proposition A.2.3 *The eigenvalue ρ discussed in Proposition A.2.2 dominates all other eigenvalues in the sense of having largest real part. If M is nonnegative and some power M^k of M has all entries positive, then ρ dominates all other eigenvalues in absolute value as well.*

Proof: Assuming M is nonnegative and taking absolute values entry by entry in the equality $Mu = \lambda u$ produce the inequality $|\lambda|w \le Mw$ for the vector w with entries $w_i = |u_i|$. In view of the definition (A.1) of ρ, this inequality implies that $|\lambda| \le \rho$. Unless $\lambda = \rho$, the real part of λ must be strictly less than ρ. This demonstrates the first claim.

For the second claim, assume that $|\lambda| = \rho$. If $|\lambda|w \ne Mw$, then Proposition A.2.1 proves that $|\lambda| < \rho$. Thus, suppose $Mw = |\lambda|w$. Successive multiplications of this equality by M yield $M^k w = |\lambda|^k w$. Comparison to the similar equality $M^k u = \lambda^k u$ shows that

$$\sum_{j=1}^{n} m_{ij}^k |u_j| \quad = \quad \left| \sum_{j=1}^{n} m_{ij}^k u_j \right| \tag{A.2}$$

for all i, where M^k has entries $m_{ij}^k > 0$. Equation (A.2) involving the complex entries of u can only be true if all u_j are positive multiples of the same complex number. Dividing $Mu = \lambda u$ by this complex number shows that w is an eigenvector of both λ and $|\lambda|$. This can occur for $w \neq \mathbf{0}$ only if $\lambda = |\lambda| = \rho$. In other words, an eigenvalue λ either coincides with ρ or has absolute value $|\lambda| < \rho$. ■

There is also a dual characterization of the dominant eigenvalue.

Proposition A.2.4 *The dominant eigenvalue ρ of a Metzler-Leontief matrix M satisfies*

$$\begin{aligned} \rho &= \inf\{\lambda : Mv \leq \lambda v, \text{ for some } v \geq \mathbf{0},\ v \neq \mathbf{0}\} \\ &= \inf\{\lambda : Mv \leq \lambda v, \text{ for some } v \in T_{n-1}\}. \end{aligned}$$

Proof: According to Propositions A.2.2 and A.2.3, the dominant eigenvalue ρ is a real number possessing row and column eigenvectors u^t and w with positive entries. Clearly, ρ and w satisfy $Mw \leq \rho w$. If λ and $v \in T_{n-1}$ satisfy $Mv \leq \lambda v$, then multiplying both sides of this inequality by u^t yields $\rho u^t v \leq \lambda u^t v$. Because $u^t v$ is positive, division by it produces $\rho \leq \lambda$. ■

As an example of the theory just developed, consider a continuous-time branching process with n types. The generator Ω of the mean matrix for the process can be written as the product $\Omega = \Lambda(F - I)$, where F is the reproduction matrix, Λ is the diagonal matrix with inverse life expectancies λ_i along its diagonal, and I is the n-dimensional identity matrix.

Proposition A.2.5 *The dominant eigenvalues $\rho(\Omega)$ and $\rho(F - I)$ are simultaneously strictly less than 0, equal to 0, or strictly greater than 0.*

Proof: The inequality $v^t\Omega \geq \mathbf{0}^t$ is valid if and only if the inequality $w^t(F - I) \geq \mathbf{0}^t$ is valid, where $w^t = v^t\Lambda$. Similarly, the conditions $v \geq \mathbf{0}$ and $w \geq \mathbf{0}$ are equivalent, and the conditions $v \neq \mathbf{0}$ and $w \neq \mathbf{0}$ are equivalent. Proposition A.2.2 therefore implies that if either $\rho(\Omega)$ or $\rho(F - I)$ is nonnegative, then the other is also nonnegative. Similarly, if either is nonpositive, then Proposition A.2.4 implies that the other is also nonpositive. We complete the proof by excluding the possibility that one of $\rho(\Omega)$ and $\rho(F - I)$ equals zero but the other is nonzero. Suppose for the sake of argument that $\rho(\Omega) > 0$ while $\rho(F - I) = 0$. Then there exist $\lambda > 0$ and $v \geq \mathbf{0}$ with $v \neq \mathbf{0}$ and $v^t\Omega \geq \lambda v^t$. But this means $w^t(F - I) \geq \mathbf{0}^t$ and $w^t(F-I) \neq \mathbf{0}^t$. Proposition A.2.1 now forces $\rho(F-I) > 0$, a contradiction. Other cases where one dominant eigenvalue is zero and the other is nonzero are handled similarly. ■

A.3 The Finite Fourier Transform

Periodic sequences $\{c_j\}_{j=-\infty}^{\infty}$ of period n constitute the natural domain of the finite Fourier transform. The transform of such a sequence is defined by

$$\hat{c}_k = \frac{1}{n}\sum_{j=0}^{n-1} c_j e^{-2\pi i \frac{jk}{n}}. \tag{A.3}$$

From this definition it follows immediately that the finite Fourier transform is linear and maps periodic sequences into periodic sequences with the same period. The inverse transform turns out to be

$$\check{d}_j = \sum_{k=0}^{n-1} d_k e^{2\pi i \frac{jk}{n}}. \tag{A.4}$$

It is fruitful to view each of these operations as a matrix times vector multiplication. Thus, if we let $u_n = e^{2\pi i/n}$ denote the principal nth root of unity, then the finite Fourier transform represents multiplication by the matrix (u_n^{-kj}/n) and the inverse transform multiplication by the matrix (u_n^{jk}). To warrant the name "inverse transform," the second matrix should be the inverse of the first. Indeed, we have

$$\begin{aligned}
\sum_{l=0}^{n-1} u_n^{jl}\frac{1}{n}u_n^{-kl} &= \frac{1}{n}\sum_{l=0}^{n-1} u_n^{(j-k)l} \\
&= \begin{cases} \frac{1}{n}\frac{1-u_n^{(j-k)n}}{1-u_n^{j-k}} & j \neq k \bmod n \\ \frac{1}{n}n & j = k \bmod n \end{cases} \\
&= \begin{cases} 0 & j \neq k \bmod n \\ 1 & j = k \bmod n. \end{cases}
\end{aligned}$$

More symmetry in the finite Fourier transform (A.3) and its inverse (A.4) can be achieved by replacing the factor $1/n$ in the transform by the factor $1/\sqrt{n}$. The inverse transform then includes the $1/\sqrt{n}$ factor as well, and the matrix $(u_n^{-kj}/\sqrt{n})$ is unitary.

We modify periodic sequences of period n by convolution, translation, reversion, and stretching. The convolution of two periodic sequences c_j and d_j is the sequence

$$c * d_k = \sum_{j=0}^{n-1} c_{k-j}d_j = \sum_{j=0}^{n-1} c_j d_{k-j}$$

with the same period. The translate of the periodic sequence c_j by index r is the periodic sequence $T_r c_j$ defined by $T_r c_j = c_{j-r}$. Thus, the operator T_r

translates a sequence r places to the right. The reversion operator R takes a sequence c_j into $Rc_j = c_{-j}$. Finally, the stretch operator S_r interpolates $r-1$ zeros between every pair of adjacent entries of a sequence c_j. In symbols,

$$S_r c_j = \begin{cases} c_{\frac{j}{r}} & r \mid j \\ 0 & r \not| j, \end{cases}$$

where $r \mid j$ indicates r divides j without remainder. The sequence $S_r c_j$ has period rn, not n. For instance, if $n = 2$ and $r = 2$, the periodic sequence $\ldots, 1, 2, 1, 2 \ldots$ becomes $\ldots, 1, 0, 2, 0, 1, 0, 2, 0, \ldots$.

Proposition A.3.1 *The finite Fourier transform satisfies the rules:*

(a) $\widehat{c * d}_k = n\hat{c}_k\hat{d}_k$

(b) $\widehat{T_r c}_k = u_n^{-rk}\hat{c}_k$

(c) $\widehat{Rc}_k = R\hat{c}_k = \hat{c}^*{}_k^*$

(d) $\widehat{S_r c}_k = \frac{\hat{c}_k}{r}$.

In (d) the transform on the left has period rn.

Proof: To prove rule (d), note that

$$\begin{aligned} \widehat{S_r c}_k &= \frac{1}{rn}\sum_{j=0}^{rn-1} S_r c_j u_{rn}^{-jk} \\ &= \frac{1}{rn}\sum_{l=0}^{n-1} c_{\frac{lr}{r}} u_{rn}^{-lrk} \\ &= \frac{1}{rn}\sum_{l=0}^{n-1} c_l u_n^{-lk} \\ &= \frac{\hat{c}_k}{r}. \end{aligned}$$

Verification of rules (a) through (c) is left to the reader. ■

The naive approach to computing the finite Fourier transform (A.3) takes $3n^2$ arithmetic operations (additions, multiplications, and complex exponentiations). The fast Fourier transform accomplishes the same task in $O(n \log n)$ operations when n is a power of 2. Proposition A.3.1 lays the foundation for deriving this useful and clever result.

Consider a sequence c_j of period n, and suppose n factors as $n = rs$. For $k = 0, 1, \ldots, r-1$, define related sequences c_j^k according to the recipe $c_j^k = c_{jr+k}$. Each of these secondary sequences has period s. We now argue

that we can recover the primary sequence through

$$c_j = \sum_{k=0}^{r-1} T_k S_r c_j^k. \tag{A.5}$$

In fact, $T_k S_r c_j^k = 0$ unless $r \mid j-k$. The condition $r \mid j-k$ occurs for exactly one value of k between 0 and $r-1$. For the chosen k,

$$\begin{aligned} T_k S_r c_j^k &= c^k_{\frac{j-k}{r}} \\ &= c_{\frac{j-k}{r}r+k} \\ &= c_j. \end{aligned}$$

In view of properties (b) and (d) of Proposition A.3.1, taking the finite Fourier transform of equation (A.5) gives

$$\hat{c}_j = \sum_{k=0}^{r-1} u_n^{-kj} \widehat{S_r c^k}_j = \sum_{k=0}^{r-1} u_n^{-kj} \frac{1}{r} \widehat{c^k}_j. \tag{A.6}$$

Now let $Op(n)$ denote the number of operations necessary to compute a finite Fourier transform of period n. From equation (A.6) it evidently takes $3r$ operations to compute each $\hat{c}_j$ once the $\widehat{c^k}_j$ are computed. Since there are n numbers $\hat{c}_j$ to compute and r sequences c_j^k, it follows that

$$Op(n) = 3nr + rOp(s). \tag{A.7}$$

If r is prime but s is not, then the same procedure can be repeated on each c_j^k to further reduce the amount of arithmetic. A simple inductive argument based on (A.7) that splits off one prime factor at a time yields

$$Op(n) = 3n(p_1 + \cdots + p_d),$$

where $n = p_1 \cdots p_d$ is the prime factorization of n. In particular, if $n = 2^d$, then $Op(n) = 6n \log_2 n$. In this case, it is noteworthy that all computations can be done in place without requiring computer storage beyond that allotted to the original vector $(c_0, \ldots, c_{n-1})^t$ [31, 87, 163, 206].

A.4 The Fourier Transform

The Fourier transform can be defined on a variety of function spaces [54, 65, 117, 135, 174]. For our purposes, it suffices to consider complex-valued, integrable functions whose domain is the real line. The Fourier transform of such a function $f(x)$ is defined according to the recipe

$$\hat{f}(y) = \int_{-\infty}^{\infty} e^{iyx} f(x) dx$$

for all real numbers y. By integrable we mean

$$\int_{-\infty}^{\infty} |f(x)|dx \quad < \quad \infty.$$

In the sequel we usually omit the limits of integration. If $f(x)$ is a probability density, then the Fourier transform $\hat{f}(y)$ coincides with the characteristic function of $f(x)$.

Table A.1 summarizes the operational properties of the Fourier transform. In the table, a, b, x_0, and y_0 are constants, and the functions $f(x)$, $xf(x)$, $\frac{d}{dx}f(x)$, and $g(x)$ are assumed integrable as needed. In entry (g), $f(x)$ is taken to be absolutely continuous. This is a technical condition permitting integration by parts and application of the fundamental theorem of calculus.

TABLE A.1. Fourier Transform Pairs

Function	**Transform**	**Function**	**Transform**
$(a) \quad af(x) + bg(x)$	$a\hat{f}(y) + b\hat{g}(y)$	$(e) \quad f(x)^*$	$\hat{f}(-y)^*$
$(b) \quad f(x - x_0)$	$e^{iyx_0}\hat{f}(y)$	$(f) \quad ixf(x)$	$\frac{d}{dy}\hat{f}(y)$
$(c) \quad e^{iy_0x}f(x)$	$\hat{f}(y + y_0)$	$(g) \quad \frac{d}{dx}f(x)$	$-iy\hat{f}(y)$
$(d) \quad f(\frac{x}{a})$	$\lvert a\rvert\hat{f}(ay)$	$(h) \quad f * g(x)$	$\hat{f}(y)\hat{g}(y)$

The next two propositions present deeper properties of the Fourier transform.

Proposition A.4.1 (Riemann-Lebesgue) *If the function $f(x)$ is integrable, then its Fourier transform $\hat{f}(y)$ is bounded, continuous, and tends to 0 as $|y|$ tends to ∞.*

Proof: The transform $\hat{f}(y)$ is bounded because

$$\begin{aligned} |\hat{f}(y)| &= \Big|\int e^{iyx}f(x)dx\Big| \\ &\leq \int |e^{iyx}||f(x)|dx \\ &= \int |f(x)|dx. \end{aligned} \tag{A.8}$$

To prove continuity, let $\lim_{n\to\infty} y_n = y$. Then the sequence of functions $g_n(x) = e^{iy_n x} f(x)$ is bounded in absolute value by $|f(x)|$ and satisfies

$$\lim_{n\to\infty} g_n(x) \quad = \quad e^{iyx} f(x).$$

Hence the dominated convergence theorem implies that

$$\lim_{n\to\infty} \int g_n(x)dx \quad = \quad \int e^{iyx} f(x)dx.$$

To prove the last assertion, we use the fact that the space of step functions with bounded support is dense in the space of integrable functions. Thus, given any $\epsilon > 0$, there exists a step function

$$g(x) \quad = \quad \sum_{j=1}^{m} c_j 1_{[x_{j-1},x_j)}(x)$$

vanishing off some finite interval and satisfying $\int |f(x) - g(x)|dx < \epsilon$. The Fourier transform $\hat{g}(y)$ has the requisite behavior at ∞ because the indicator function $1_{[x_{j-1},x_j)}(x)$ has Fourier transform

$$\int_{x_{j-1}}^{x_j} e^{iyx)} dx \quad = \quad e^{i\frac{1}{2}(x_{j-1}+x_j)y} \frac{\sin[\frac{1}{2}(x_j - x_{j-1})y]}{\frac{1}{2}y}.$$

This allows us to calculate

$$\hat{g}(y) \quad = \quad \sum_{j=1}^{m} c_j e^{i\frac{1}{2}(x_{j-1}+x_j)y} \frac{\sin[\frac{1}{2}(x_j - x_{j-1})y]}{\frac{1}{2}y},$$

and this finite sum clearly tends to 0 as $|y|$ tends to ∞. The original transform $\hat{f}(y)$ exhibits the same behavior because the bound (A.8) entails the inequality

$$\begin{aligned} |\hat{f}(y)| &\leq |\hat{f}(y) - \hat{g}(y)| + |\hat{g}(y)| \\ &\leq \epsilon + |\hat{g}(y)|. \end{aligned}$$

This completes the proof. ■

Proposition A.4.2 *Let $f(x)$ be a bounded, continuous function. If $f(x)$ and $\hat{f}(y)$ are both integrable, then*

$$f(x) \quad = \quad \frac{1}{2\pi} \int e^{-iyx} \hat{f}(y)dy. \tag{A.9}$$

Proof: Consider the identities

$$\begin{aligned}
\frac{1}{2\pi}\int e^{-iyx}e^{-\frac{y^2}{2\sigma^2}}\hat{f}(y)dy &= \frac{1}{2\pi}\int e^{-iyx}e^{-\frac{y^2}{2\sigma^2}}\int e^{iyu}f(u)du\,dy \\
&= \int f(u)\frac{1}{2\pi}\int e^{iy(u-x)}e^{-\frac{y^2}{2\sigma^2}}dy\,du \\
&= \int f(u)\frac{\sigma}{\sqrt{2\pi}}e^{-\frac{\sigma^2(u-x)^2}{2}}du \\
&= \frac{1}{\sqrt{2\pi}}\int f\Big(x+\frac{v}{\sigma}\Big)e^{-\frac{v^2}{2}}dv,
\end{aligned}$$

which involve Example 2.4.1 and the change of variables $u = x + v/\sigma$. As σ tends to ∞, the last integral tends to

$$\frac{1}{\sqrt{2\pi}}\int f(x)e^{-\frac{v^2}{2}}dv = f(x),$$

while the original integral tends to

$$\frac{1}{2\pi}\int e^{-iyx}\lim_{\sigma\to\infty}e^{-\frac{y^2}{2\sigma^2}}\hat{f}(y)dy = \frac{1}{2\pi}\int e^{-iyx}\hat{f}(y)dy.$$

Equating these two limits yields the inversion formula (A.9). ■

A.5 Fourier Series

The space $L^2[0,1]$ of square-integrable functions on [0,1] is the prototype for all Hilbert spaces [54, 174, 189]. The structure of $L^2[0,1]$ is determined by the inner product

$$\langle f, g\rangle = \int_0^1 f(x)g(x)^*dx$$

and its associated norm $\|f\| = \langle f, f\rangle^{1/2}$. It is well known that the complex exponentials $\{e^{2\pi inx}\}_{n=-\infty}^{\infty}$ provide an orthonormal basis for $L^2[0,1]$. Indeed, the calculation

$$\begin{aligned}
\int_0^1 e^{2\pi imx}e^{-2\pi inx}dx &= \begin{cases} 1 & m = n \\ \frac{e^{2\pi i(m-n)x}}{2\pi i(m-n)}\Big|_0^1 & m \neq n \end{cases} \\
&= \begin{cases} 1 & m = n \\ 0 & m \neq n \end{cases}
\end{aligned}$$

shows that the sequence is orthonormal. Completeness is essentially a consequence of Fejér's theorem [54], which says that any periodic, continuous function can be uniformly approximated by a linear combination of sines

and cosines. (Fejér's theorem is a special case of the more general Stone-Weierstrass theorem [90].) The Fourier coefficients of $f(x)$ are computed according to the standard recipe

$$c_n = \int_0^1 f(x)e^{-2\pi inx}dx.$$

The Fourier series $\sum_{n=-\infty}^{\infty} c_n e^{2\pi inx}$ is guaranteed to converge to $f(x)$ in mean square. The more delicate issue of pointwise convergence is partially covered by the next proposition. In the proposition and subsequent discussion, we implicitly view every function $f(x)$ defined on [0,1] as extended periodically to the whole real line via the equation $f(x+1) = f(x)$.

Proposition A.5.1 *Assume that the square-integrable function $f(x)$ on [0,1] is continuous at x_0 and possesses both one-sided derivatives there. Then*

$$\lim_{m\to\infty} \sum_{n=-m}^{m} c_n e^{2\pi inx_0} = f(x_0).$$

Proof: Extend $f(x)$ to be periodic, and consider the associated periodic function

$$g(x) = \frac{f(x+x_0) - f(x_0)}{e^{-2\pi ix} - 1}.$$

Applying l'Hôpital's rule yields

$$\lim_{x\to 0^+} g(x) = \frac{\frac{d}{dx}f(x_0^+)}{-2\pi i},$$

where $\frac{d}{dx}f(x_0^+)$ denotes the one-sided derivative from the right. A similar expression holds for the limit from the left. Since these two limits are finite and $\int_0^1 |f(x)|^2 dx < \infty$, we have $\int_0^1 |g(x)|^2 dx < \infty$ as well.

Now let d_n be the nth Fourier coefficient of $g(x)$. Because

$$f(x+x_0) = f(x_0) + (e^{-2\pi ix} - 1)g(x),$$

it follows that

$$\begin{aligned} c_n e^{2\pi inx_0} &= \int_0^1 f(x)e^{-2\pi in(x-x_0)}dx \\ &= \int_0^1 f(x+x_0)e^{-2\pi inx}dx \\ &= f(x_0)1_{\{n=0\}} + d_{n+1} - d_n. \end{aligned}$$

Therefore,

$$\begin{aligned}\sum_{n=-m}^{m} c_n e^{2\pi i n x_0} &= f(x_0) + \sum_{n=-m}^{m} (d_{n+1} - d_n) \\ &= f(x_0) + d_{m+1} - d_{-m}.\end{aligned}$$

To complete the proof, observe that

$$\begin{aligned}\lim_{|m|\to\infty} d_m &= \lim_{|m|\to\infty} \int_0^1 g(x) e^{-2\pi i m x} dx \\ &= 0\end{aligned}$$

by Proposition A.4.1. ■

Example A.5.1 *Bernoulli Functions*

To derive the Euler-Maclaurin summation formula, one must introduce Bernoulli polynomials $B_n(x)$ and periodic Bernoulli functions $b_n(x)$. Let us start with the Bernoulli polynomials. These are defined by the three conditions

$$\begin{aligned}B_0(x) &= 1 \\ \frac{d}{dx} B_n(x) &= n B_{n-1}(x), \quad n > 0 \\ \int_0^1 B_n(x) dx &= 0, \quad n > 0.\end{aligned} \tag{A.10}$$

For example, we calculate recursively

$$\begin{aligned}B_1(x) &= x - \frac{1}{2} \\ B_2(x) &= 2\Big(\frac{x^2}{2} - \frac{x}{2} + \frac{1}{12}\Big).\end{aligned}$$

The Bernoulli function $b_n(x)$ coincides with $B_n(x)$ on $[0, 1)$. Outside $[0, 1)$, $b_n(x)$ is extended periodically. In particular, $b_0(x) = B_0(x) = 1$ for all x. Note that $b_1(x)$ is discontinuous at $x = 1$ while $b_2(x)$ is continuous. All subsequent $b_n(x)$ are continuous at $x = 1$ because

$$B_n(1) - B_n(0) = \int_0^1 \frac{d}{dx} B_n(x) dx = n \int_0^1 B_{n-1}(x) dx = 0$$

by assumption.

To compute the Fourier series expansion $\sum_k c_{nk} e^{2\pi i k x}$ of $b_n(x)$ for $n > 0$, note that $c_{n0} = \int_0^1 B_n(x) dx = 0$. For $k \neq 0$, we have

$$c_{nk} = \int_0^1 b_n(x) e^{-2\pi i k x} dx$$

$$\begin{aligned} &= b_n(x)\frac{e^{-2\pi ikx}}{-2\pi ik}\Big|_0^1 + \frac{1}{2\pi ik}\int_0^1 \frac{d}{dx}b_n(x)e^{-2\pi ikx}dx \qquad \text{(A.11)} \\ &= b_n(x)\frac{e^{-2\pi ikx}}{-2\pi ik}\Big|_0^1 + \frac{n}{2\pi ik}\int_0^1 b_{n-1}(x)e^{-2\pi ikx}dx. \end{aligned}$$

From the integration-by-parts formula (A.11), we deduce that $b_1(x)$ has Fourier series expansion

$$-\frac{1}{2\pi i}\sum_{k\neq 0}\frac{e^{2\pi ikx}}{k}.$$

This series converges pointwise to $b_1(x)$ except at $x = 0$ and $x = 1$. For $n > 1$, the boundary terms in (A.11) cancel, and

$$c_{nk} = \frac{nc_{n-1,k}}{2\pi ik}. \qquad \text{(A.12)}$$

Formula (A.12) and Proposition A.5.1 together imply that

$$b_n(x) = -\frac{n!}{(2\pi i)^n}\sum_{k\neq 0}\frac{e^{2\pi ikx}}{k^n} \qquad \text{(A.13)}$$

for all $n > 1$ and all x.

The constant term $B_n = B_n(0)$ is known as a Bernoulli number. One can compute B_{n-1} recursively by expanding $B_n(x)$ in a Taylor series around $x = 0$. In view of the defining properties (A.10),

$$\begin{aligned} B_n(x) &= \sum_{k=0}^{n}\frac{1}{k!}\frac{d^k}{dx^k}B_n(0)x^k \\ &= \sum_{k=0}^{n}\frac{1}{k!}n^{\underline{k}}B_{n-k}x^k \\ &= \sum_{k=0}^{n}\binom{n}{k}B_{n-k}x^k, \end{aligned}$$

where

$$n^{\underline{k}} = n(n-1)\cdots(n-k+1)$$

denotes a falling power. The continuity and periodicity of $b_n(x)$ for $n \geq 2$ therefore imply that

$$B_n = B_n(1) = \sum_{k=0}^{n}\binom{n}{k}B_{n-k}.$$

Subtracting B_n from both sides of this equality gives the recurrence relation

$$0 = \sum_{k=1}^{n} \binom{n}{k} B_{n-k}$$

for computing B_{n-1} from $B_0, \ldots, B_{n-2}$. For instance, starting from $B_0 = 1$, we calculate $B_1 = -1/2$, $B_2 = 1/6$, $B_3 = 0$, and $B_4 = -1/30$. From the expansion (A.13), evidently $B_n = 0$ for all odd integers $n > 1$. ■

A.6 Laplace's Method and Watson's Lemma

Here we undertake a formal proof of the second Laplace asymptotic formula (12.6). Proof of the first formula (12.4) is similar.

Proposition A.6.1 *If the conditions*

(a) for every $\delta > 0$ there exists a $\rho > 0$ with $g(y) - g(0) \geq \rho$ for $|y| \geq \delta$,

(b) $g(y)$ is twice continuously differentiable in a neighborhood of 0 and $g''(0) > 0$,

(c) $f(y)$ is continuous in a neighborhood of 0 and $f(0) > 0$,

(d) the integral $\int_{-\infty}^{\infty} f(y)e^{-xg(y)}dy$ is absolutely convergent for $x \geq x_1$,

are satisfied, then the asymptotic relation (12.6) obtains.

Proof: By multiplying both sides of the asymptotic relation (12.6) by $e^{xg(0)}$, we can assume without loss of generality that $g(0) = 0$. Because $g(y)$ has its minimum at $y = 0$, l'Hôpital's rule implies $g(y) - \frac{1}{2}g''(0)y^2 = o(y^2)$ as $y \to 0$. Now let a small $\epsilon > 0$ be given, and choose $\delta > 0$ sufficiently small so that the inequalities

$$\begin{aligned} (1-\epsilon)f(0) &\leq f(y) \\ &\leq (1+\epsilon)f(0) \\ \left| g(y) - \frac{1}{2}g''(0)y^2 \right| &\leq \epsilon y^2 \end{aligned}$$

hold for $|y| \leq \delta$. Assumption (a) guarantees the existence of a $\rho > 0$ with $g(y) \geq \rho$ for $|y| \geq \delta$.

We next show that the contributions to the Laplace integral from the region $|y| \geq \delta$ are negligible as $x \to \infty$. Indeed, for $x \geq x_1$,

$$\begin{aligned} \left| \int_{\delta}^{\infty} f(y)e^{-xg(y)}dy \right| &\leq \int_{\delta}^{\infty} |f(y)|e^{-(x-x_1)g(y)}e^{-x_1 g(y)}dy \\ &\leq e^{-(x-x_1)\rho} \int_{\delta}^{\infty} |f(y)|e^{-x_1 g(y)}dy \\ &= O(e^{-\rho x}). \end{aligned}$$

Likewise, $\int_{-\infty}^{-\delta} f(y)e^{-xg(y)}dy = O(e^{-\rho x})$.

Owing to our choice of δ, the central portion of the integral satisfies

$$\int_{-\delta}^{\delta} f(y)e^{-xg(y)}dy \quad \leq \quad (1+\epsilon)f(0)\int_{-\delta}^{\delta} e^{-\frac{x}{2}[g''(0)-2\epsilon]y^2}dy.$$

Duplicating the above reasoning,

$$\int_{-\infty}^{-\delta} e^{-\frac{x}{2}[g''(0)-2\epsilon]y^2}dy + \int_{\delta}^{\infty} e^{-\frac{x}{2}[g''(0)-2\epsilon]y^2}dy \quad = \quad O(e^{-\omega x}),$$

where $\omega = \frac{1}{2}[g''(0) - 2\epsilon]\delta^2$. Thus,

$$\begin{aligned} & (1+\epsilon)f(0)\int_{-\delta}^{\delta} e^{-\frac{x}{2}[g''(0)-2\epsilon]y^2}dy \\ = \quad & (1+\epsilon)f(0)\int_{-\infty}^{\infty} e^{-\frac{x}{2}[g''(0)-2\epsilon]y^2}dy + O(e^{-\omega x}) \\ = \quad & (1+\epsilon)f(0)\sqrt{\frac{2\pi}{x[g''(0)-2\epsilon]}} + O(e^{-\omega x}). \end{aligned}$$

Assembling all of the relevant pieces, we now conclude that

$$\begin{aligned} \int_{-\infty}^{\infty} f(y)e^{-xg(y)}dy \quad \leq \quad & (1+\epsilon)f(0)\sqrt{\frac{2\pi}{x[g''(0)-2\epsilon]}} \\ & + O(e^{-\rho x}) + O(e^{-\omega x}). \end{aligned}$$

Hence,

$$\limsup_{x\to\infty} \sqrt{x}\int_{-\infty}^{\infty} f(y)e^{-xg(y)}dy \quad \leq \quad (1+\epsilon)f(0)\sqrt{\frac{2\pi}{[g''(0)-2\epsilon]}},$$

and sending $\epsilon \to 0$ produces

$$\limsup_{x\to\infty} \sqrt{x}\int_{-\infty}^{\infty} f(y)e^{-xg(y)}dy \quad \leq \quad f(0)\sqrt{\frac{2\pi}{g''(0)}}.$$

A similar argument gives

$$\liminf_{x\to\infty} \sqrt{x}\int_{-\infty}^{\infty} f(y)e^{-xg(y)}dy \quad \geq \quad f(0)\sqrt{\frac{2\pi}{g''(0)}}$$

and proves the proposition. ■

Proof of Watson's lemma in Section 12.3.2 is easier to sketch. The asymptotic expansion for $f(x)$ entails the bound

$$\left| f(x) - \sum_{k=0}^{n-1} a_k x^{\lambda_k - 1} \right| \quad \leq \quad b x^{\lambda_n - 1}$$

for some $b > 0$ and x close to 0. Together with the assumption that $f(x)$ is $O(e^{cx})$ as $x \to \infty$, this implies the existence of another constant d such that

$$\left| f(x) - \sum_{k=0}^{n-1} a_k x^{\lambda_k - 1} \right| \quad \leq \quad d e^{cx} x^{\lambda_n - 1}$$

for all x. Taking Laplace transforms now gives

$$\left| \tilde{f}(t) - \sum_{k=0}^{n-1} a_k \frac{\Gamma(\lambda_k)}{t^{\lambda_k}} \right| \quad \leq \quad d \frac{\Gamma(\lambda_n)}{(t-c)^{\lambda_n}} \quad \asymp \quad d \frac{\Gamma(\lambda_n)}{t^{\lambda_n}}$$

as $t \to \infty$.

A.7 A Tauberian Theorem

In this section we prove Proposition 15.5.2 using Cauchy's integral formula. For $T > 0$ set

$$\hat{g}_T(s) \quad = \quad \int_0^T g(t) e^{-st} dt.$$

This function is entire, that is analytic throughout the complex plane, and the proposition asserts that $\lim_{T \to \infty} \hat{g}_T(0) = \hat{g}(0)$. The assumption that $\hat{g}(s)$ can be continued to the imaginary axis guarantees that $\hat{g}(0)$ exists.

To apply Cauchy's formula, we let R be a large positive radius and define the contour C as the boundary of the region $\{s : |s| \leq R, \mathsf{Re}(s) \geq -\delta\}$ in the complex plane. Here $\delta > 0$ is chosen small enough so that $\hat{g}(s)$ is analytic inside and on C. In general, δ will depend on R. Now Cauchy's theorem implies that

$$\hat{g}(0) - \hat{g}_T(0) \quad = \quad \frac{1}{2\pi i} \int_C [\hat{g}(s) - \hat{g}_T(s)] e^{sT} \left(1 + \frac{s^2}{R^2}\right) \frac{1}{s} \, ds. \quad \text{(A.14)}$$

On the semicircle $C_+ = C \cap \{\mathsf{Re}(s) > 0\}$, the integrand is bounded by $2B/R^2$, where $B = \sup_{t \geq 0} |g(t)|$. This bound follows from the inequality

$$\left| \hat{g}(s) - \hat{g}_T(s) \right| \quad = \quad \left| \int_T^\infty g(t) e^{-st} dt \right|$$

$$\begin{aligned} &\leq B\int_T^\infty e^{-\operatorname{Re}(s)t}dt \\ &= \frac{Be^{-\operatorname{Re}(s)T}}{\operatorname{Re}(s)} \end{aligned}$$

and the equality

$$\begin{aligned} \left|e^{sT}\left(1+\frac{s^2}{R^2}\right)\frac{1}{s}\right| &= e^{\operatorname{Re}(s)T}\left|\frac{1}{s}+\frac{s}{|s|^2}\right| \\ &= e^{\operatorname{Re}(s)T}\left|\frac{s^*+s}{|s|^2}\right| \\ &= e^{\operatorname{Re}(s)T}\frac{2|\operatorname{Re}(s)|}{|R|^2}, \end{aligned} \tag{A.15}$$

with the understanding that $|s| = R$ on C_+ and s^* is the complex conjugate of s. Because the length of C_+ equals πR, the contribution of C_+ to the integral (A.14) is dominated by B/R.

For the integral over $C_- = C \cap \{\operatorname{Re}(s) < 0\}$, we deal with the terms involving $\hat{g}(s)$ and $\hat{g}_T(s)$ separately. Since $\hat{g}_T(s)$ is an entire function, we can deform the contour C_- to the semicircle $\{s : |s| = R, \operatorname{Re}(s) \leq 0\}$ without changing the value of the integral. With this change in mind,

$$\begin{aligned} \left|\hat{g}_T(s)\right| &= \left|\int_0^T g(t)e^{-st}dt\right| \\ &\leq B\int_0^T e^{-\operatorname{Re}(s)t}dt \\ &\leq B\int_{-\infty}^T e^{-\operatorname{Re}(s)t}dt \\ &= \frac{Be^{-\operatorname{Re}(s)T}}{|\operatorname{Re}(s)|}. \end{aligned}$$

Because equality (A.15) is still valid, the integrand is again bounded by $2B/R^2$ and the integral along the semicircle by B/R.

If we can show that the integral along C_- involving $\hat{g}(s)$ tends to 0 as T tends to ∞, then it is clear that $\limsup_{T\to\infty}|\hat{g}(0)-\hat{g}_T(0)| \leq 2B/R$. Hence, sending R to ∞ gives the desired conclusion. Therefore, consider the final integral. Its integrand is the product of the function $\hat{g}(s)(1+s^2/R^2)/s$, which does not depend on T, and the function e^{sT}, which converges exponentially fast to 0 along C_- as T tends to ∞. Thus for any fixed R, the integral along C_- involving $\hat{g}(s)$ tends to 0 as T tends to ∞.

References

[1] Aigner M, Ziegler GM (1999) *Proofs from the Book*. Springer, New York

[2] Aldous D (1989) *Probability Approximations via the Poisson Clumping Heuristic.* Springer, New York

[3] Aldous D, Diaconis P (1986) Shuffling cards and stopping times. *Amer Math Monthly* 93:333–348

[4] Alon N, Spencer JH, Erdös P (1992) *The Probabilistic Method*. Wiley, New York

[5] Anděl J (2001) *Mathematics of Chance*. Wiley, New York

[6] Apostol T (1974) *Mathematical Analysis*, 2nd ed. Addison-Wesley, Reading, MA

[7] Angus J (1994) The probability integral transform and related results. *SIAM Review* 36:652–654

[8] Apostol T (1976) *Introduction to Analytic Number Theory.* Springer, New York

[9] Arnold BC, Balakrishnan N, Nagaraja HN (1992) *A First Course in Order Statistics.* Wiley, New York

[10] Arratia R, Goldstein L, Gordon L (1989) Two moments suffice for Poisson approximations: the Chen-Stein method. *Ann Prob* 17:9–25

K. Lange, *Applied Probability,*
DOI 10.1007/978-1-4419-7165-4, © Springer Science+Business Media, LLC 2010

[11] Arratia R, Goldstein L, Gordon L (1990) Poisson approximation and the Chen-Stein method. *Stat Sci* 5:403–434

[12] Asmussen S, Glynn PW (2007) *Stochastic Simulation: Algorithms and Analysis.* Springer, New York

[13] Asmussen S, Hering H (1983) *Branching Processes.* Birkhäuser, Boston

[14] Athreya KB, Ney PE (1972) *Branching Processes.* Springer, New York

[15] Baclawski K, Rota G-C, Billey S (1989) *An Introduction to the Theory of Probability.* Massachusetts Institute of Technology, Cambridge, MA

[16] Baker JA (1997) Integration over spheres and the divergence theorem for balls. *Amer Math Monthly* 64:36–47

[17] Balasubramanian K, Balakrishnan N (1993) Duality principle in order statistics. *J Roy Stat Soc B* 55:687–691

[18] Barbour AD, Holst L, Janson S (1992) *Poisson Approximation.* Oxford University Press, Oxford

[19] Barndorff-Nielsen OE, Cox DR (1989) *Asymptotic Techniques for Use in Statistics.* Chapman & Hall, London

[20] Bender CM, Orszag SA (1978) *Advanced Mathematical Methods for Scientists and Engineers.* McGraw-Hill, New York

[21] Benjamin AT, Quinn JJ (2003) *Proofs That Really Count: The Art of Combinatorial Proof.* Mathematical Association of America, Washington, DC

[22] Berge C (1971) *Principles of Combinatorics.* Academic Press, New York

[23] Bhattacharya RN, Waymire EC (1990) *Stochastic Processes with Applications.* Wiley, New York

[24] Billingsley P (1986) *Probability and Measure,* 2nd ed. Wiley, New York

[25] Blom G, Holst L (1991) Embedding procedures for discrete problems in probability. *Math Scientist* 16:29–40

[26] Blom G, Holst L, Sandell D (1994) *Problems and Snapshots from the World of Probability.* Springer, New York

[27] Boas RP Jr (1977) Partial sums of infinite series, and how they grow. *Amer Math Monthly* 84:237–258

[28] Bradley RA, Terry ME (1952), Rank analysis of incomplete block designs, *Biometrika* 39:324–345

[29] Bragg LR (1991) Parametric differentiation revisited. *Amer Math Monthly* 98:259–262

[30] Brezeźniak Z, Zastawniak T (1999) *Basic Stochastic Processes.* Springer, New York

[31] Brigham EO (1974) *The Fast Fourier Transform.* Prentice-Hall, Englewood Cliffs, NJ

[32] Brualdi RA (1977) *Introductory Combinatorics.* North-Holland, New York

[33] Cao Y, Gillespie DT, Petzold LR (2006). Efficient leap-size selection for accelerated stochastic simulation. *J Phys Chem* 124:1–11

[34] Casella G, Berger RL (1990) *Statistical Inference.* Wadsworth and Brooks/Cole, Pacific Grove, CA

[35] Casella G, George EI (1992) Explaining the Gibbs sampler. *Amer Statistician* 46:167–174

[36] Chen LHY (1975) Poisson approximation for dependent trials. *Ann Prob* 3:534–545

[37] Chib S, Greenberg E (1995) Understanding the Metropolis-Hastings algorithm. *Amer Statistician* 49:327–335

[38] Chung KL, Williams RJ (1990) *Introduction to Stochastic Integration,* 2nd ed. Birkhäuser, Boston

[39] Ciarlet PG (1989) *Introduction to Numerical Linear Algebra and Optimization.* Cambridge University Press, Cambridge

[40] Cochran WG (1977) *Sampling Techniques*, 3rd ed. Wiley, New York

[41] Coldman AJ, Goldie JH (1986) A stochastic model for the origin and treatment of tumors containing drug resistant cells. *Bull Math Biol* 48:279–292

[42] Comet L (1974) *Advanced Combinatorics.* Reidel, Dordrecht

[43] Cressie N, Davis AS, Folks JL, Polocelli GE Jr (1981) The moment-generating function and negative integer moments. *Amer Statistician* 35:148–150

[44] Crow, JF, Kimura M (1970) *An Introduction to Population Genetics Theory.* Harper & Row, New York

[45] David HA (1993) A note on order statistics for dependent variates. *Amer Statistician* 47:198–199

[46] de Bruijn NG (1981) *Asymptotic Methods in Analysis.* Dover, New York

[47] De Pierro AR (1993) On the relation between the ISRA and EM algorithm for positron emission tomography. *IEEE Trans Med Imaging* 12:328–333

[48] D'Eustachio P, Ruddle FH (1983) Somatic cell genetics and gene families. *Science* 220:919–924

[49] Diaconis P (1988) *Group Representations in Probability and Statistics.* Institute of Mathematical Statistics, Hayward, CA

[50] Dieudonné J (1971) *Infinitesimal Calculus.* Houghton-Mifflin, Boston

[51] Dorman K, Sinsheimer JS, Lange K (2004) In the garden of branching processes. *SIAM Review* 46:202–229

[52] Doyle PG, Snell JL (1984) *Random Walks and Electrical Networks.* The Mathematical Association of America, Washington, DC

[53] Durrett R (1991) *Probability: Theory and Examples.* Wadsworth & Brooks/Cole, Pacific Grove, CA

[54] Dym H, McKean HP (1972) *Fourier Series and Integrals.* Academic Press, New York

[55] Erdös P, Fűredi Z (1981) The greatest angle among n points in the d-dimensional Euclidean space. *Ann Discrete Math* 17:275–283

[56] Ewens, W.J. (1979). *Mathematical Population Genetics.* Springer, New York

[57] Fan R, Lange K (1999) Diffusion process calculations for mutant genes in nonstationary populations. In *Statistics in Molecular Biology and Genetics.* edited by Seillier-Moiseiwitsch F, The Institute of Mathematical Statistics and the American Mathematical Society, Providence, RI, pp 38-55

[58] Fan R, Lange K, Peña EA (1999) Applications of a formula for the variance function of a stochastic process. *Stat Prob Letters* 43:123–130

[59] Feller W (1968) *An Introduction to Probability Theory and Its Applications, Vol 1*, 3rd ed. Wiley, New York

[60] Feller W (1971) *An Introduction to Probability Theory and Its Applications, Vol 2*, 2nd ed. Wiley, New York

[61] Ferguson TS (1996) *A Course in Large Sample Theory.* Chapman & Hall, London

[62] Fix E, Neyman J (1951) A simple stochastic model of recovery, relapse, death and loss of patients. *Hum Biol* 23:205–241

[63] Flanders H (2001) From Ford to Faà. *Amer Math Monthly* 108:559–561

[64] Flatto L, Konheim AG (1962) The random division of an interval and the random covering of a circle. *SIAM Review* 4:211–222

[65] Folland GB (1992) *Fourier Analysis and its Applications.* Wadsworth and Brooks/Cole, Pacific Grove, CA

[66] Friedlen DM (1990) More socks in the laundry. Problem E 3265. *Amer Math Monthly* 97:242–244

[67] Galambos J, Simonelli I (1996) *Bonferroni-type Inequalities with Applications.* Springer, New York

[68] Gani J, Saunders IW (1977) Fitting a model to the growth of yeast colonies. *Biometrics* 33:113–120

[69] Gelfand AE, Smith AFM (1990) Sampling-based approaches to calculating marginal densities. *J Amer Stat Assoc* 85:398–409

[70] Gelman A, Carlin JB, Stern HS, Rubin DB (1995) *Bayesian Data Analysis.* Chapman & Hall, London

[71] Geman S, Geman D (1984) Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Trans Pattern Anal Machine Intell* 6:721–741

[72] Geman S, McClure D (1985) Bayesian image analysis: An application to single photon emission tomography. *Proceedings of the Statistical Computing Section*, American Statistical Association, Washington, DC, pp 12–18

[73] Gilks WR, Richardson S, Spiegelhalter DJ (editors) (1996) *Markov Chain Monte Carlo in Practice.* Chapman & Hall, London

[74] Gillespie DT (1977) Exact stochastic simulation of coupled chemical reactions. *J Phys Chem* 81:2340–2361

[75] Gillespie DT (2001) Approximate accelerated stochastic simulation of chemically reacting systems. *J Chem Phys* 115:1716–1733

[76] Goradia TM (1991) *Stochastic Models for Human Gene Mapping.* Ph.D. Thesis, Division of Applied Sciences, Harvard University

[77] Goradia TM, Lange K (1988) Applications of coding theory to the design of somatic cell hybrid panels. *Math Biosci* 91:201–219

[78] Graham RL, Knuth DE, Patashnik O (1988) *Concrete Mathematics: A Foundation for Computer Science*. Addison-Wesley, Reading, MA

[79] Green P (1990) Bayesian reconstruction for emission tomography data using a modified EM algorithm. *IEEE Trans Med Imaging* 9:84–94

[80] Grimmett GR, Stirzaker DR (2001) *Probability and Random Processes*, 3rd ed. Oxford University Press, Oxford

[81] Guttorp P (1991) *Statistical Inference for Branching Processes.* Wiley, New York

[82] Gzyl H, Palacios JL (1997) The Weierstrass approximation theorem and large deviations. *Amer Math Monthly* 104:650–653

[83] Hardy GH, Wright EM (1960) *An Introduction to the Theory of Numbers,* 4th ed. Clarendon Press, Oxford

[84] Harris TE (1989) *The Theory of Branching Processes.* Dover, New York

[85] Hastings WK (1970) Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* 57:97–109

[86] Henrici P (1979) Fast Fourier transform methods in computational complex analysis. *SIAM Review* 21:481–527

[87] Henrici P (1982) *Essentials of Numerical Analysis with Pocket Calculator Demonstrations.* Wiley, New York

[88] Herman GT (1980) *Image Reconstruction from Projections: The Fundamentals of Computerized Tomography.* Springer, New York

[89] Hestenes MR (1981) *Optimization Theory: The Finite Dimensional Case.* Robert E Krieger Publishing, Huntington, NY

[90] Hewitt E, Stromberg K (1965) *Real and Abstract Analysis.* Springer, New York

[91] Higham DJ, (2008) Modeling and simulating chemical reactions. *SIAM Review* 50:347–368

[92] Higham NJ (2009) The scaling and squaring method for matrix exponentiation. *SIAM Review* 51:747–764

[93] Hille E (1959) *Analytic Function Theory, Vol 1.* Blaisdell, New York

[94] Hirsch MW, Smale S (1974) *Differential Equations, Dynamical Systems, and Linear Algebra.* Academic Press, New York

[95] Hochstadt H (1986) *The Functions of Mathematical Physics.* Dover, New York

[96] Hoel PG, Port SC, Stone CJ (1971) *Introduction to Probability Theory.* Houghton Mifflin, Boston

[97] Hoffman K (1975) *Analysis in Euclidean Space.* Prentice-Hall, Englewood Cliffs, NJ

[98] Hoppe FM (2008) Faà di Bruno's formula and the distribution of random partitions in population genetics and physics. *Theor Pop Biol* 73:543–551

[99] Horn RA, Johnson CR (1991) *Topics in Matrix Analysis.* Cambridge University Press, Cambridge

[100] Hwang JT (1982) Improving on standard estimators in discrete exponential families with applications to the Poisson and negative binomial cases. *Ann Statist* 10:857–867

[101] Isaacson E, Keller HB (1966) *Analysis of Numerical Methods.* Wiley, New York

[102] Jacquez JA, Simon CP, Koopman JS (1991) The reproduction number in deterministic models of contagious diseases. *Comments Theor Biol* 2:159–209

[103] Jagers P (1975) *Branching Processes with Biological Applications.* Wiley, New York

[104] Jones GA, Jones JM (1998) *Elementary Number Theory.* Springer, New York

[105] Kac M (1959) *Statistical Independence in Probability, Analysis and Number Theory.* Mathematical Association of America, Washington, DC

[106] Karlin S, Taylor HM (1975) *A First Course in Stochastic Processes,* 2nd ed. Academic Press, New York

[107] Karlin S, Taylor HM (1981) *A Second Course in Stochastic Processes.* Academic Press, New York

[108] Katz B (1966) *Nerve, Muscle, and Synapse.* McGraw-Hill, New York

[109] Keener JP (1993), The Perron-Frobenius theorem and the ranking of football teams, *SIAM Review*, 35:80–93

[110] Keiding N (1991) Age-specific incidence and prevalence: a statistical perspective. *J Royal Stat Soc Series A* 154:371–412

[111] Kelly FP (1979) *Reversibility and Stochastic Networks.* Wiley, New York

[112] Kennedy WJ Jr, Gentle JE (1980) *Statistical Computing.* Marcel Dekker, New York

[113] Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol* 16:111–120

[114] Kingman JFC (1993) *Poisson Processes.* Oxford University Press, Oxford

[115] Kirkpatrick S, Gelatt CD, Vecchi MP (1983) Optimization by simulated annealing. *Science* 220:671–680

[116] Klain DA, Rota G-C (1997) *Introduction to Geometric Probability.* Cambridge University Press, Cambridge

[117] Körner TW (1988) *Fourier Analysis.* Cambridge University Press, Cambridge

[118] Lamperti J (1977) *Stochastic Processes. A Survey of the Mathematical Theory.* Springer, New York

[119] Lando SK (2003) *Lectures on Generating Functions.* American Mathematical Society, Providence, RI

[120] Lange K (1982) Calculation of the equilibrium distribution for a deleterious gene by the finite Fourier transform. *Biometrics* 38:79–86

[121] Lange K (1995) A gradient algorithm locally equivalent to the EM algorithm. *J Roy Stat Soc B* 57:425–437

[122] Lange K (1999) *Numerical Analysis for Statisticians.* Springer, New York

[123] Lange K (2002) *Mathematical and Statistical Methods for Genetic Analysis,* 2nd ed. Springer, New York

[124] Lange K, Boehnke M (1982) How many polymorphic genes will it take to span the human genome? *Amer J Hum Genet* 34:842–845

[125] Lange K, Carson R (1984) EM reconstruction algorithms for emission and transmission tomography. *J Computer Assist Tomography* 8:306–316

[126] Lange K, Fessler JA (1995) Globally convergent algorithms for maximum a posteriori transmission tomography. *IEEE Trans Image Processing* 4:1430–1438

[127] Lange K, Gladstien K (1980) Further characterization of the long-run population distribution of a deleterious gene. *Theor Pop Bio* 18:31-43

[128] Lange K, Hunter DR, Yang I (2000) Optimization transfer using surrogate objective functions (with discussion). *J Comput Graph Statist* 9:1–59

[129] Lawler GF (1995) *Introduction to Stochastic Processes.* Chapman & Hall, London

[130] Lawler GF, Coyle LN (1999) *Lectures on Contemporary Probability.* American Mathematical Society, Providence, RI

[131] Lazzeroni LC, Lange K (1997) Markov chains for Monte Carlo tests of genetic equilibrium in multidimensional contingency tables. *Ann Statist* 25:138–168

[132] Lehmann EL (2004) *Elements of Large-Sample Theory*, Springer, New York

[133] Li W-H, Graur D (1991) *Fundamentals of Molecular Evolution.* Sinauer Associates, Sunderland, MA

[134] Liggett TM (1985) *Interacting Particle Systems.* Springer, New York

[135] Lighthill MJ (1958) *An Introduction to Fourier Analysis and Generalized Functions.* Cambridge University Press, Cambridge

[136] Lindvall T (1992) *Lectures on the Coupling Method.* Wiley, New York

[137] Liu JS (1996) Metropolized independent sampling with comparisons to rejection sampling and importance sampling. *Stat Comput* 6:113–119

[138] Lotka AJ (1931) Population analysis—the extinction of families I. *J Wash Acad Sci* 21:377–380

[139] Lozansky E, Rousseau C (1996) *Winning Solutions.* Springer, New York

[140] Luenberger DG (1984) *Linear and Nonlinear Programming,* 2nd ed. Addison-Wesley, Reading, MA

[141] McLeod RM (1980) *The Generalized Riemann Integral.* Mathematical Association of America, Washington, DC

[142] Metropolis N, Rosenbluth A, Rosenbluth M, Teller A, Teller E (1953) Equations of state calculations by fast computing machines. *J Chem Physics* 21:1087–1092

[143] Minin VN, Suchard MA (2008) Counting labeled transitions in continuous-time Markov models of evolution. *Math Biol* 56:391–412

[144] Mitzenmacher M, Upfal E (2005) *Probability and Computing: Randomized Algorithms and Probabilistic Analysis.* Cambridge University Press, Cambridge

[145] Moler C, Van Loan C (1978) Nineteen dubious ways to compute the exponential of a matrix. *SIAM Review* 20:801–836

[146] Muldoon ME, Ungar AA (1996) Beyond sin and cos. *Mathematics Magazine* 69:3–14

[147] Murray JD (1984) *Asymptotic Analysis.* Springer, New York

[148] Nedelman J, Wallenuis T (1986) Bernoulli trials, Poisson trials, surprising variances, and Jensen's inequality. *Amer Statistician* 40:286–289

[149] Neuts MF (1995) *Matrix-Geometric Solutions in Stochastic Models: An Algorithmic Approach.* Dover, New York

[150] Newman DJ (1980) Simple analytic proof of the prime number theorem. *Amer Math Monthly* 87:693–696

[151] Niven I, Zuckerman HS, Montgomery HL (1991) *An Introduction to the Theory of Numbers.* Wiley, New York

[152] Norris JR (1997) *Markov Chains.* Cambridge University Press, Cambridge

[153] Paige CC, Styan GPH, Wachter PG (1975) Computation of the stationary distribution of a Markov chain. *J Stat Comput Simul* 4:173–186

[154] Perelson AS, Macken CA (1985) *Branching Processes Applied to Cell Surface Aggregation Phenomena.* Springer, Berlin

[155] Peressini AL, Sullivan FE, Uhl JJ Jr (1988) *The Mathematics of Nonlinear Programming.* Springer, New York

[156] Phillips AN (1996) Reduction of HIV concentration during acute infection: Independence from a specific immune response. *Science* 271:497–499

[157] Pinsker IS, Kipnis V, Grechanovsky E (1986) A recursive formula for the probability of occurrence of at least m out of N events. *Amer Statistician* 40:275–276

[158] Pittenger AO (1985) The logarithmic mean in n dimensions. *Amer Math Monthly* 92:99–104

[159] Pollard JH (1973) *Mathematical Models for the Growth of Human Populations.* Cambridge University Press, Cambridge

[160] Port SC, Stone CJ (1978) *Brownian Motion and Classical Potential Theory.* Academic Press, New York

[161] Post E (1930) Generalized differentiation. *Trans Amer Math Soc* 32:723–781

[162] Powell MJD (1981) *Approximation Theory and Methods.* Cambridge University Press, Cambridge

[163] Press WH, Teukolsky SA, Vetterling WT, Flannery BP (1992) *Numerical Recipes in Fortran: The Art of Scientific Computing*, 2nd ed. Cambridge University Press, Cambridge

[164] Rao CR (1973) *Linear Statistical Inference and Its Applications,* 2nd ed. Wiley, New York

[165] Reed TE, Neil JV (1959) Huntington's chorea in Michigan. *Amer J Hum Genet* 11:107–136

[166] Rényi A (1970) *Probability Theory.* North-Holland, Amsterdam

[167] Roman S (1984) *The Umbral Calculus.* Academic Press, New York

[168] Roman S (1997) *Introduction to Coding and Information Theory.* Academic Press, New York

[169] Rosenthal JS (1995) Convergence rates for Markov chains. *SIAM Review* 37:387–405

[170] Ross SM (1996) *Stochastic Processes*, 2nd ed. Wiley, New York

[171] Royden HL (1968) *Real Analysis*, 2nd ed. Macmillan, London

[172] Rubinow SI (1975) *Introduction to Mathematical Biology.* Wiley, New York

[173] Rudin W (1964) *Principles of Mathematical Analysis,* 2nd ed. McGraw-Hill, New York

[174] Rudin W (1973) *Functional Analysis.* McGraw-Hill, New York

[175] Rushton AR (1976) Quantitative analysis of human chromosome segregation in man-mouse somatic cell hybrids. *Cytogenetics Cell Genet* 17:243–253

[176] Schadt EE, Lange K (2002) Codon and rate variation models in molecular phylogeny. *Mol Biol Evol* 19:1534–1549

[177] Sedgewick R (1988) *Algorithms*, 2nd ed. Addison-Wesley, Reading, MA

[178] Sehl ME, Alexseyenko AV, Lange KL (2009) Accurate stochastic simulation via the step anticipation (SAL) algorithm. *J Comp Biol* 16:1195–1208

[179] Seneta E (1973) *Non-negative Matrices: An Introduction to Theory and Applications.* Wiley, New York

[180] Serre D (2002) *Matrices: Theory and Applications.* Springer, New York

[181] Severini TA (2005) *Elements of Distribution Theory.* Cambridge University Press, Cambridge

[182] Shargel BH, D'Orsogna MR, Chou T (2010) Arrival times in a zero-range process with injection and decay. *J Phys A: Math Theor* (in press)

[183] Silver EA, Costa D (1997) A property of symmetric distributions and a related order statistic result. *Amer Statistician* 51:32–33

[184] Solomon H (1978) *Geometric Probability.* SIAM, Philadelphia

[185] Spivak M (1965) *Calculus on Manifolds.* Benjamin/Cummings Publishing Co., Menlo Park, CA

[186] Steele JM (1997) *Probability Theory and Combinatorial Optimization.* SIAM, Philadelphia

[187] Stein C (1986) *Approximate Computation of Expectations.* Institute of Mathematical Statistics, Hayward, CA

[188] Stein EM, Shakarchi R (2003) *Complex Analysis.* Princeton University Press, Princeton, NJ

[189] Stein EM, Shakarchi R (2005) *Real Analysis: Measure Theory, Integration, and Hilbert Spaces.* Princeton University Press, Princeton, NJ

[190] Steutel FW (1985) Poisson processes and a modified Bessel integral. *SIAM Review* 27:73–77

[191] Stewart WJ (1994) *Introduction to the Numerical Solution of Markov Chains.* Princeton University Press, Princeton, NJ

[192] Strichartz R (2000) Evaluating integrals using self-similarity. *Amer Math Monthly* 107::316–326

[193] Sumita U, Igaki N (1997) Necessary and sufficient conditions for global convergence of block Gauss-Seidel iteration algorithm applied to Markov chains. *J Operations Research* 40:283–293

[194] Tanner MA (1993) *Tools for Statistical Inference: Methods for Exploration of Posterior Distributions and Likelihood Functions*, 2nd ed. Springer, New York

[195] Tanny S (1973) A probabilistic interpretation of Eulerian numbers. *Duke Math J* 40:717–722

[196] Taylor HM, Karlin S (1984) *An Introduction to Stochastic Modeling.* Academic Press, Orlando, FL

[197] Tenenbaum G, France MM (2000) *The Prime Numbers and Their Distribution.* translated by Spain PG, American Mathematical Society, Providence, RI

[198] The Huntington's Disease Collaborative Research Group (1993). A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell* 72:971–983

[199] Tuckwell HC (1989) *Stochastic Processes in the Neurosciences.* SIAM, Philadelphia

[200] van den Driessche P, Watmough J (2002) Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Math Biosciences* 180:29–48

[201] Waterman MS (1995) *Introduction to Computational Biology: Maps, Sequences, and Genomes.* Chapman & Hall, London

[202] Weiss M, Green H (1967) Human-mouse hybrid cell lines containing partial complements of human chromosomes and functioning human genes. *Proc Natl Acad Sci USA* 58:1104–1111

[203] Whitmore GA, Seshadri V (1987) A heuristic derivation of the inverse Gaussian distribution. *Amer Statistician* 41:280–281

[204] Wilf HS (1978) *Mathematics for the Physical Sciences.* Dover, New York

[205] Wilf HS (1985) Some examples of combinatorial averaging. *Amer Math Monthly* 92:250–260

[206] Wilf HS (1986) *Algorithms and Complexity.* Prentice-Hall, New York

[207] Wilf HS (1990) *generatingfunctionology.* Academic Press, San Diego

[208] Williams D (1991) *Probability with Martingales*. Cambridge University Press, Cambridge

[209] Williams D (2001) *Weighing the Odds: A Course in Probability and Statistics*. Cambridge University Press, Cambridge

[210] Yee PL, Vyb́orný R (2000) *The Integral: An Easy Approach after Kurzweil and Henstock*. Cambridge University Press, Cambridge

[211] Zagier D (1997) Newman's short proof of the prime number theorem. *Amer Math Monthly* 104:705–708

Index

K. Lange, *Applied Probability,*
DOI 10.1007/978-1-4419-7165-4, © Springer Science+Business Media, LLC 2010