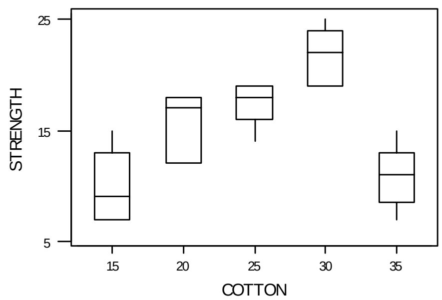
CHAPTER 13

Section 13-2

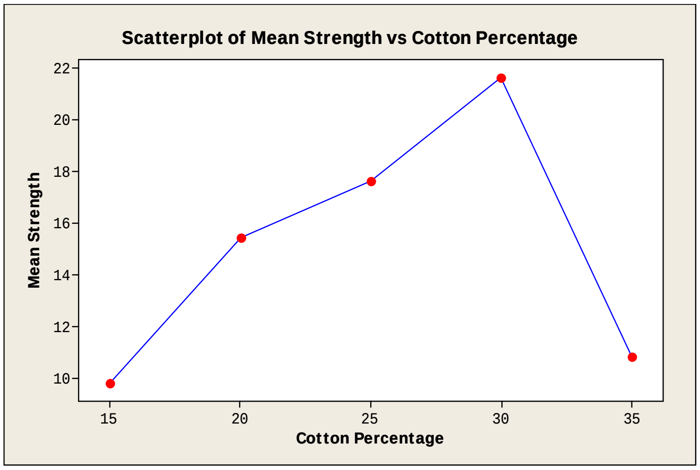
13-1. a) Analysis of Variance for STRENGTH

Source	DF	SS	MS	F	P
COTTON	4	475.76	118.94	14.76	0.000
Error	20	161.20	8.06		
Total	24	636.96			

Reject H_0 and conclude that cotton percentage affects mean breaking strength.

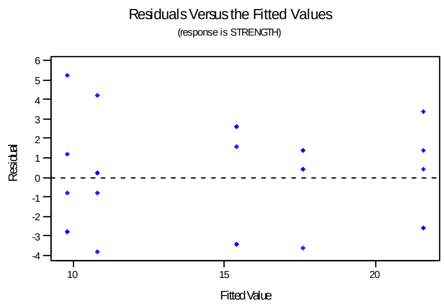


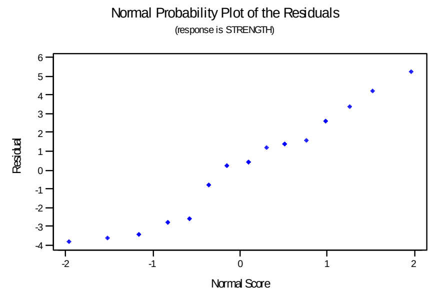
b) Tensile strength seems to increase up to 30% cotton and declines at 35% cotton.

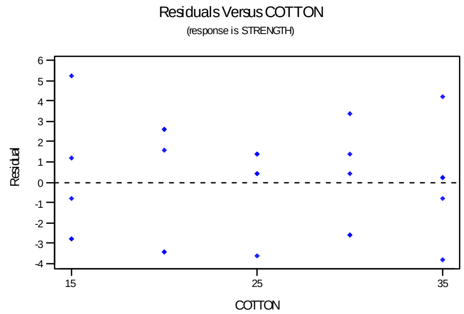



```
                                          Individual 95% CIs For Mean
                                          Based on Pooled StDev
Level       N      Mean     StDev   ------+---------+---------+---------+
15          5     9.800     3.347   (-----*----)
20          5    15.400     3.130               (----*----)
25          5    17.600     2.074                   (----*----)
30          5    21.600     2.608                           (----*----)
35          5    10.800     2.864     (-----*----)
                                    ------+---------+---------+---------+
Pooled StDev =    2.839                 10.0      15.0      20.0      25.0
```

c) The normal probability plot and the residual plots show that the model assumptions are reasonable.



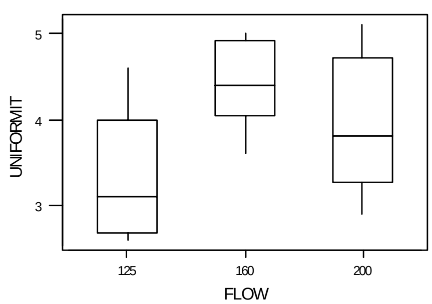




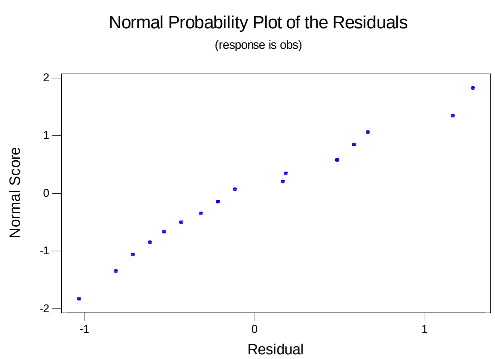
3-2 a) Analysis of Variance for FLOW

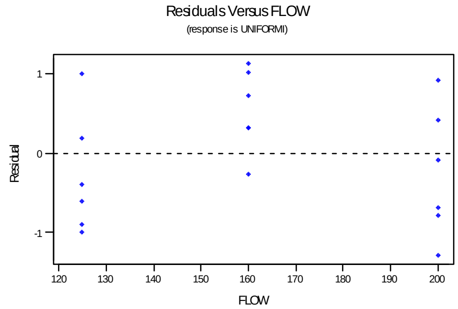
```
Source       DF        SS         MS        F       P
FLOW          2     3.6478     1.8239    3.59   0.053
Error        15     7.6300     0.5087
Total        17    11.2778
```

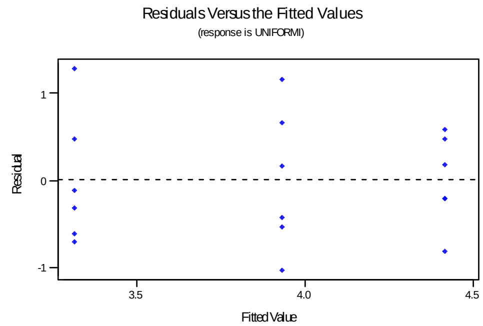
Do not reject H_0. There is no evidence that flow rate affects etch uniformity.



b) Residuals are acceptable.







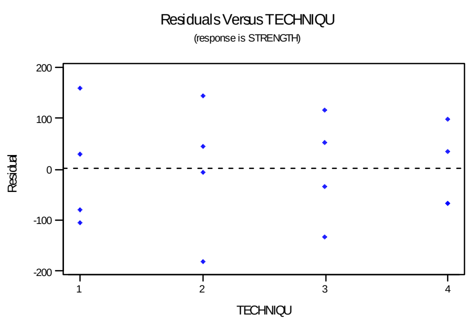
13-3. a) Analysis of Variance for STRENGTH

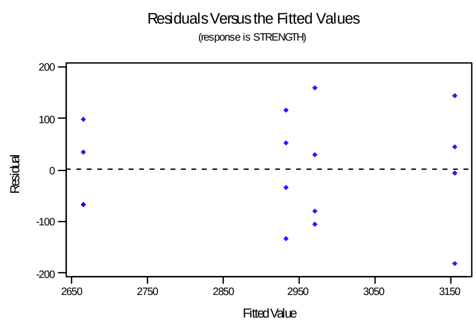
Source	DF	SS	MS	F	P
TECHNIQU	3	489740	163247	12.73	0.000
Error	12	153908	12826		
Total	15	643648			

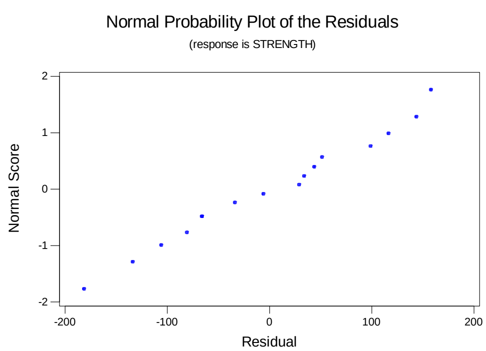
Reject H_0. Techniques affect the mean strength of the concrete.

b) *P-value* $\cong 0$

c) Residuals are acceptable







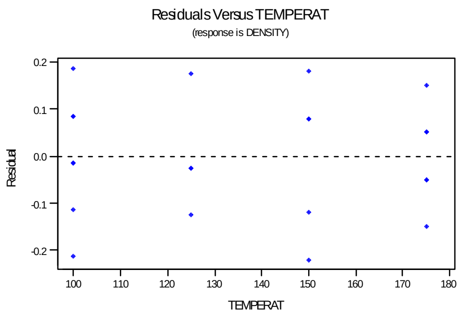
13-4 a) Analysis of Variance for TEMPERATURE

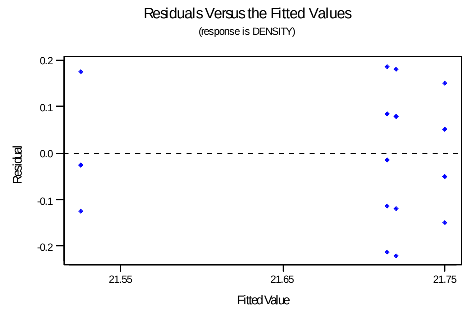
Source	DF	SS	MS	F	P
TEMPERAT	3	0.1391	0.0464	2.62	0.083
Error	18	0.3191	0.0177		
Total	21	0.4582			

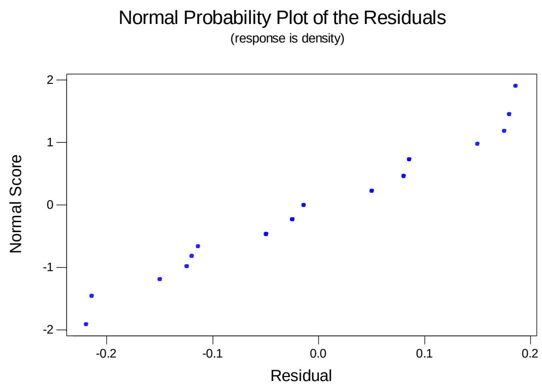
Do not reject H_0

b) *P*-value = 0.083

c) Residuals are acceptable





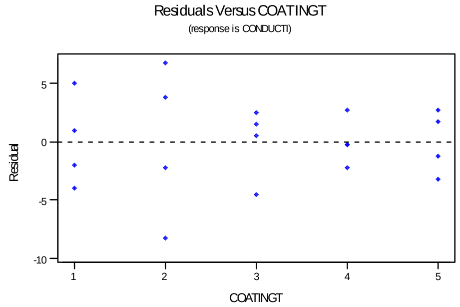


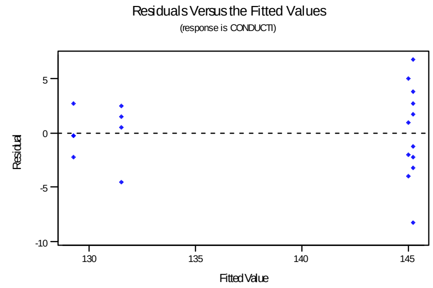
13-5. a) Analysis of Variance for CONDUCTIVITY

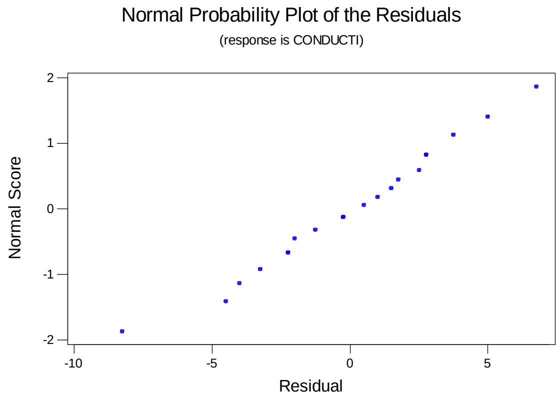
Source	DF	SS	MS	F	P
COATINGTYPE	4	1060.5	265.1	16.35	0.000
Error	15	243.3	16.2		
Total	19	1303.8			

Reject H_0, *P*-value $\cong 0$

b) There is some indication of that the variability of the response may be increasing as the mean response increases. There appears to be an outlier on the normal probability plot.







c) 95% Confidence interval on the mean of coating type 1

$$\bar{y}_1 - t_{0.025,15}\sqrt{\frac{MS_E}{n}} \le \mu_i \le \bar{y}_1 + t_{0.015,15}\sqrt{\frac{MS_E}{n}}$$

$$145.00 - 2.131\sqrt{\frac{16.2}{4}} \le \mu_1 \le 145.00 + 2.131\sqrt{\frac{16.2}{4}}$$

$$140.71 \le \mu_1 \le 149.29$$

99% confidence interval on the difference between the means of coating types 1 and 4.

$$\bar{y}_1 - \bar{y}_4 - t_{0.005,15}\sqrt{\frac{2MS_E}{n}} \le \mu_1 - \mu_4 \le \bar{y}_1 - \bar{y}_4 + t_{0.005,15}\sqrt{\frac{2MS_E}{n}}$$

$$(145.00 - 129.25) - 2.947\sqrt{\frac{2(16.2)}{4}} \le \mu_1 - \mu_4 \le (145.00 - 129.25) - 2.947\sqrt{\frac{2(16.2)}{4}}$$

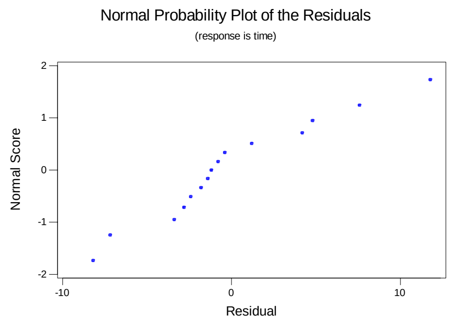
$$7.36 \le \mu_1 - \mu_4 \le 24.14$$

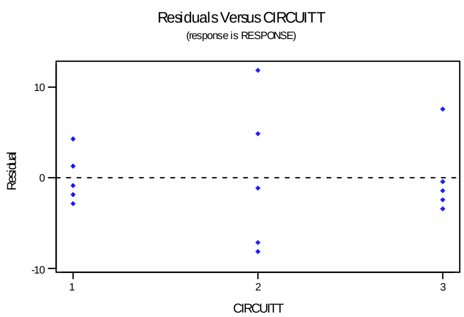
13-6 a) Analysis of Variance for CIRCUIT TYPE

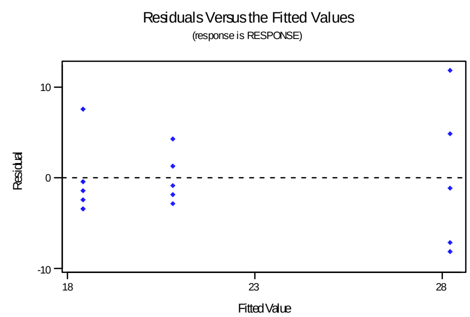
Source	DF	SS	MS	F	P
CIRCUITT	2	260.9	130.5	4.01	0.046
Error	12	390.8	32.6		
Total	14	651.7			

Reject H_0

b)There is some indication of greater variability in circuit two. There is some curvature in the normal probability plot.







c) 95% Confidence interval on the mean of circuit type 3.

$$\bar{y}_3 - t_{0.025,12}\sqrt{\frac{MS_E}{n}} \le \mu_i \le \bar{y}_3 + t_{0.015,12}\sqrt{\frac{MS_E}{n}}$$

$$18.4 - 2.179\sqrt{\frac{32.6}{5}} \le \mu_1 \le 18.4 + 2.179\sqrt{\frac{32.6}{5}}$$

$$12.84 \le \mu_1 \le 23.96$$

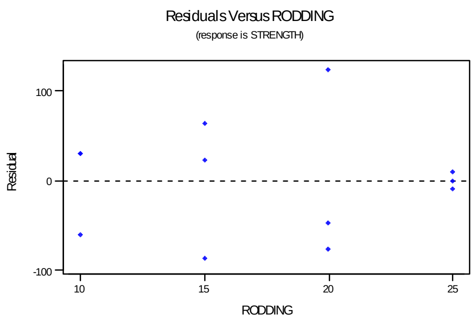
13-7. a) Analysis of Variance for STRENGTH

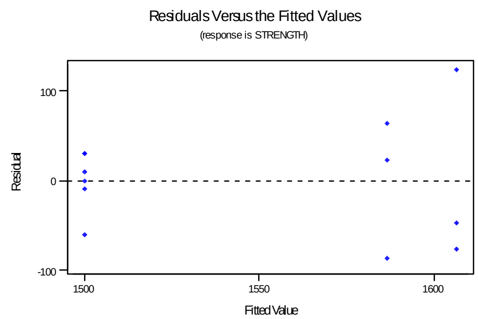
Source	DF	SS	MS	F	P
RODDING	3	28633	9544	1.87	0.214
Error	8	40933	5117		
Total	11	69567			

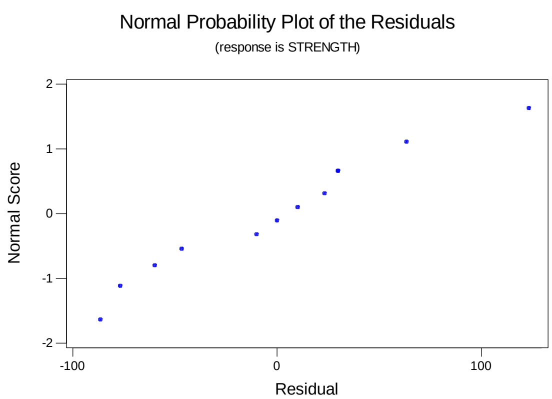
Do not reject H_0

b) *P-value* = 0.214

c) The residual plot suggests nonconstant variance. The normal probability plot looks acceptable.







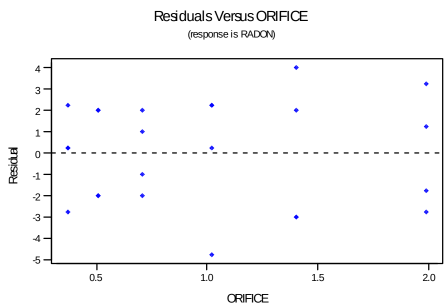
13-8 a) Analysis of Variance for ORIFICE

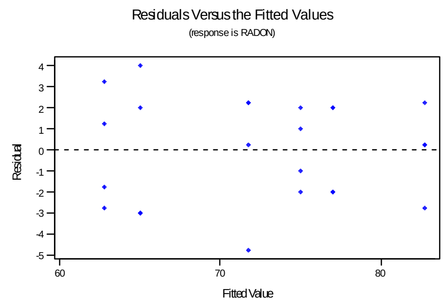
Source	DF	SS	MS	F	P
ORIFICE	5	1133.37	226.67	30.85	0.000
Error	18	132.25	7.35		
Total	23	1265.63			

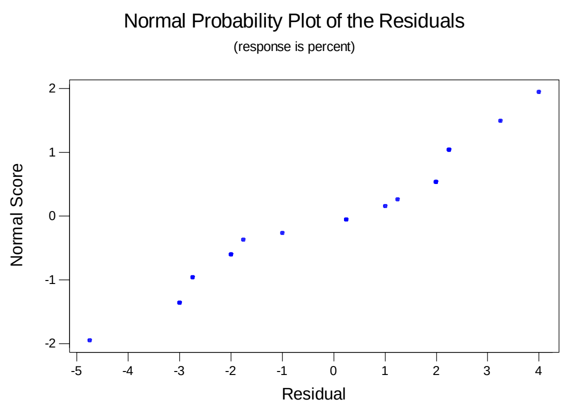
Reject H_0

b) *P*-value $\cong 0$

c)







d) 95% CI on the mean radon released when diameter is 1.40

$$\bar{y}_5 - t_{0.025,18}\sqrt{\frac{MS_E}{n}} \le \mu_i \le \bar{y}_5 + t_{0.015,18}\sqrt{\frac{MS_E}{n}}$$

$$65 - 2.101\sqrt{\frac{7.35}{4}} \le \mu_1 \le 65 + 2.101\sqrt{\frac{7.35}{4}}$$

$$62.15 \le \mu_1 \le 67.84$$

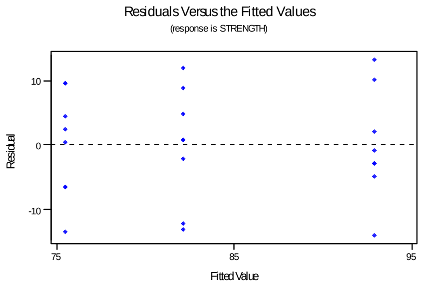
13-9. a) Analysis of Variance for STRENGTH

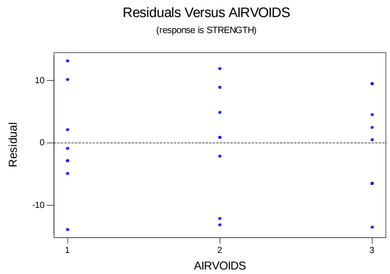
Source	DF	SS	MS	F	P
AIRVOIDS	2	1230.3	615.1	8.30	0.002
Error	21	1555.8	74.1		
Total	23	2786.0			

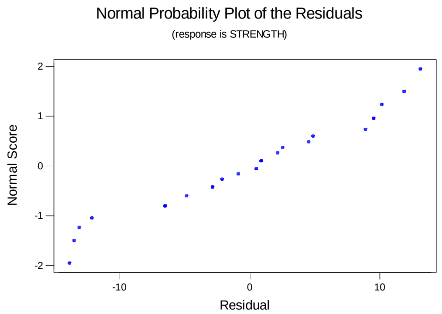
Reject H_0

b) P-value = 0.002

c) The residual plots indicate that the constant variance assumption is reasonable. The normal probability plot has some curvature in the tails but appears reasonable.







d) 95% Confidence interval on the mean of retained strength where there is a high level of air voids

$$\bar{y}_3 - t_{0.025,21}\sqrt{\frac{MS_E}{n}} \le \mu_i \le \bar{y}_3 + t_{0.015,21}\sqrt{\frac{MS_E}{n}}$$

$$75.5 - 2.080\sqrt{\frac{74.1}{8}} \le \mu_3 \le 75.5 + 2.080\sqrt{\frac{74.1}{8}}$$

$$69.17 \le \mu_1 \le 81.83$$

e) 95% confidence interval on the difference between the means of retained strength at the high level and the low levels of air voids.

$$\bar{y}_1 - \bar{y}_3 - t_{0.025,21}\sqrt{\frac{2MS_E}{n}} \le \mu_1 - \mu_3 \le \bar{y}_1 - \bar{y}_3 + t_{0.025,21}\sqrt{\frac{2MS_E}{n}}$$

$$(92.875 - 75.5) - 2.080\sqrt{\frac{2(74.1)}{8}} \le \mu_1 - \mu_4 \le (92.875 - 75.5) - 2.080\sqrt{\frac{2(74.1)}{8}}$$

$$8.42 \le \mu_1 - \mu_4 \le 26.33$$

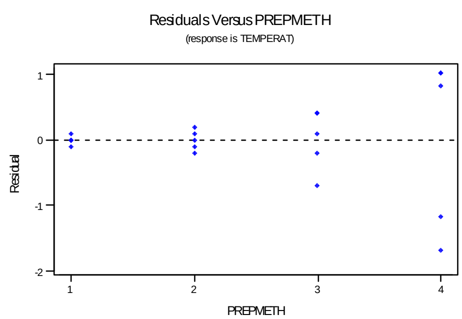
13-10 a) Analysis of Variance of PREPARATION METHOD

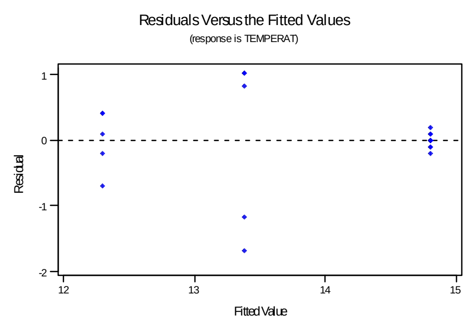
Source	DF	SS	MS	F	P
PREPMETH	3	22.124	7.375	14.85	0.000
Error	16	7.948	0.497		
Total	19	30.072			

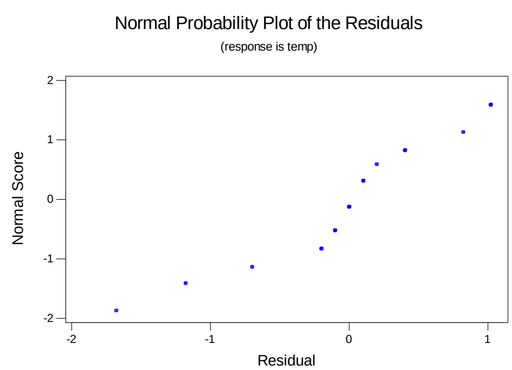
Reject H_0

b) *P*-value $\cong 0$

c) There are some differences in the amount variability at the different preparation methods and there is some curvature in the normal probability plot. There are also some potential problems with the constant variance assumption apparent in the fitted value plot.







d) 95% Confidence interval on the mean of temperature for preparation method 1

$$\bar{y}_1 - t_{0.025,16}\sqrt{\frac{MS_E}{n}} \le \mu_i \le \bar{y}_1 + t_{0.015,16}\sqrt{\frac{MS_E}{n}}$$

$$14.8 - 2.120\sqrt{\frac{0.497}{5}} \le \mu_3 \le 14.8 + 2.120\sqrt{\frac{0.497}{5}}$$

$$14.13 \le \mu_1 \le 15.47$$

13-11. Fisher's pairwise comparisons

Family error rate = 0.264
Individual error rate = 0.0500
Critical value = 2.086
Intervals for (column level mean) - (row level mean)

	15	20	25	30
20	-9.346			
	-1.854			
25	-11.546	-5.946		
	-4.054	1.546		
30	-15.546	-9.946	-7.746	
	-8.054	-2.454	-0.254	
35	-4.746	0.854	3.054	7.054
	2.746	8.346	10.546	14.546

Significant differences between levels 15 and 20, 15 and 25, 15 and 30, 20 and 30, 20 and 35, 25 and 30, 25 and 35, and 30 and 35.

13-12 Fisher's pairwise comparisons

Family error rate = 0.117
Individual error rate = 0.0500
Critical value = 2.131

Intervals for (column level mean) - (row level mean)

	125	160
160	-1.9775	
	-0.2225	
250	-1.4942	-0.3942
	0.2608	1.3608

There are significant differences between levels 125 and 160.

13-13. Fisher's pairwise comparisons

Family error rate = 0.184
Individual error rate = 0.0500
Critical value = 2.179
Intervals for (column level mean) - (row level mean)

	1	2	3
2	-360		
	-11		
3	-137	48	
	212	397	
4	130	316	93
	479	664	442

Significance differences between levels 1 and 2, 1 and 4, 2 and 3, 2 and 4, and 3 and 4.

13-14 Fisher's pairwise comparisons

Family error rate = 0.0649
Individual error rate = 0.0100
Critical value = 2.947
Intervals for (column level mean) - (row level mean)

	1	2	3	4
2	-8.642			
	8.142			
3	5.108	5.358		
	21.892	22.142		
4	7.358	7.608	-6.142	
	24.142	24.392	10.642	
5	-8.642	-8.392	-22.142	-24.392

8.142	8.392	-5.358	-7.608

Significant differences b/t 1 and 3, 1 and 4, 2 and 3, 2 and 4, 3 and 5, 4 and 5.

13-15. Fisher's pairwise comparisons

Family error rate = 0.0251
Individual error rate = 0.0100

Critical value = 3.055

Intervals for (column level mean) - (row level mean)

	1	2
2	-18.426 3.626	
3	-8.626 13.426	-1.226 20.826

No significant differences at $\alpha = 0.01$.

13-16 Fisher's pairwise comparisons
Family error rate = 0.330
Individual error rate = 0.0500
Critical value = 2.101
Intervals for (column level mean) - (row level mean)

	0.37	0.51	0.71	1.02	1.40
0.51	1.723 9.777				
0.71	3.723 11.777	-2.027 6.027			
1.02	6.973 15.027	1.223 9.277	-0.777 7.277		
1.40	13.723 21.777	7.973 16.027	5.973 14.027	2.723 10.777	
1.99	15.973 24.027	10.223 18.277	8.223 16.277	4.973 13.027	-1.777 6.277

Significant differences between levels 0.37 and all other levels; 0.51 and 1.02, 1.40, and 1.99; 0.71 and 1.40 and 1.99; 1.02 and 1.40 and 1.99

13-17. Fisher's pairwise comparisons
Family error rate = 0.118
Individual error rate = 0.0500
Critical value = 2.080
Intervals for (column level mean) - (row level mean)

	1	2
2	1.799 19.701	
3	8.424 26.326	-2.326 15.576

Significant differences between levels 1 and 2; and 1 and 3.

13-18 Fisher's pairwise comparisons
Family error rate = 0.189
Individual error rate = 0.0500
Critical value = 2.120
Intervals for (column level mean) - (row level mean)

	1	2	3
2	-0.9450		

```
          0.9450
   3      1.5550      1.5550
          3.4450      3.4450
   4      0.4750      0.4750     -2.0250
          2.3650      2.3650     -0.1350
```

```
There are significant differences between levels 1 and 3, 4; 2 and 3, 4;
and 3 and 4.
```

13-19. $\overline{\mu} = 55$, $\tau_1 = -5$, $\tau_2 = 5$, $\tau_3 = -5$, $\tau_4 = 5$.

$$\Phi^2 = \frac{n(100)}{4(25)} = n, \quad a-1=3 \quad a(n-1)=4(n-1)$$

Various choices for *n* yield:

n	Φ^2	Φ	a(n-1)	Power=1-β
4	4	2	12	0.80
5	5	2.24	16	0.90

Therefore, $n = 5$ is needed.

13-20 $\overline{\mu} = 188$, $\tau_1 = -13$, $\tau_2 = 2$, $\tau_3 = -28$, $\tau_4 = 12$, $\tau_5 = 27$.

$$\Phi^2 = \frac{n(1830)}{5(100)} = n, \quad a-1=4 \quad a(n-1)=5(n-1)$$

Various choices for *n* yield:

n	Φ^2	Φ	a(n-1)	Power=1-β
2	7.32	2.7	5	0.55
3	10.98	3.13	10	0.95

Therefore, $n = 3$ is needed.

Section 13-3

13-21 a)

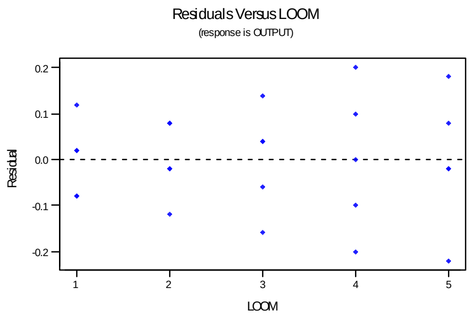
```
Analysis of Variance for OUTPUT
Source     DF        SS        MS        F        P
LOOM        4    0.3416    0.0854     5.77    0.003
Error      20    0.2960    0.0148
Total      24    0.6376
```

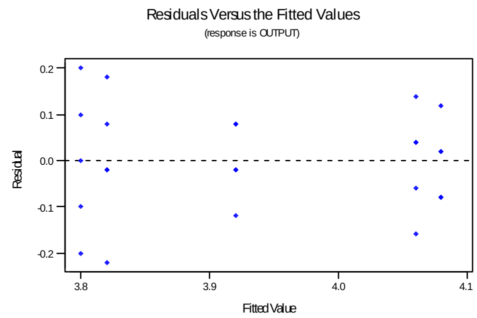
Reject H_0, there are significant differences among the looms.

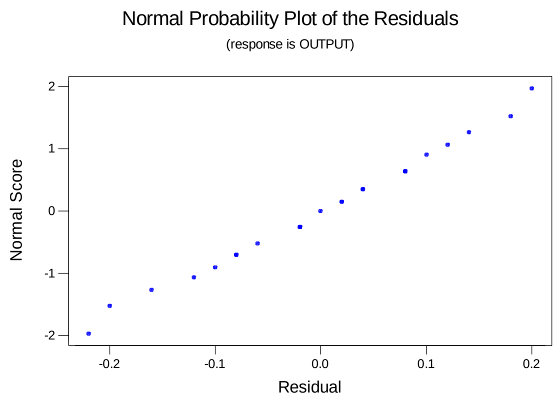
b) $$\hat{\sigma}_\tau^2 = \frac{MS_{Treatments} - MS_E}{n} = \frac{0.0854 - 0.0148}{5} = 0.01412$$

c) $\hat{\sigma}^2 = MS_E = 0.0148$

d) Residuals are acceptable







13-22 a)

Analysis of Variance for UNIFORMITY

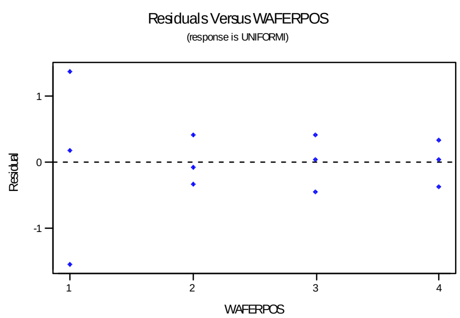
Source	DF	SS	MS	F	P
WAFERPOS	3	16.220	5.407	8.29	0.008
Error	8	5.217	0.652		
Total	11	21.437			

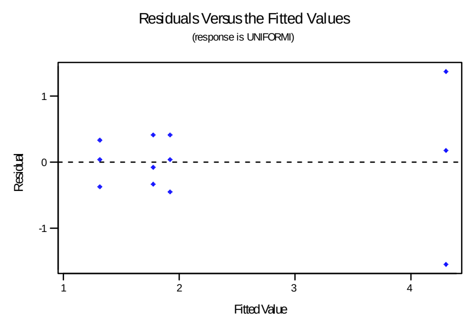
Reject H_0, and conclude that there are significant differences among wafer positions.

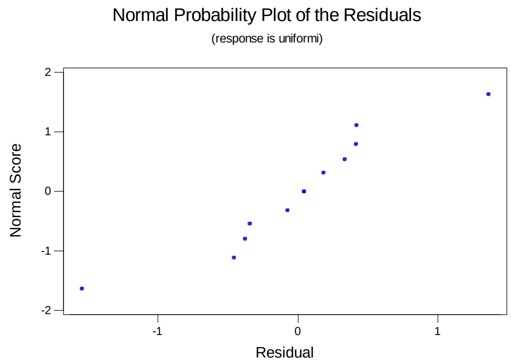
b) $\hat{\sigma}_\tau^2 = \dfrac{MS_{Treatments} - MS_E}{n} = \dfrac{5.407 - 0.652}{4} = 1.189$

c) $\hat{\sigma}^2 = MS_E = 0.652$

d) Greater variability at wafer position 1. There is some slight curvature in the normal probability plot.







13-23. a) Analysis of Variance for BRIGHTNENESS

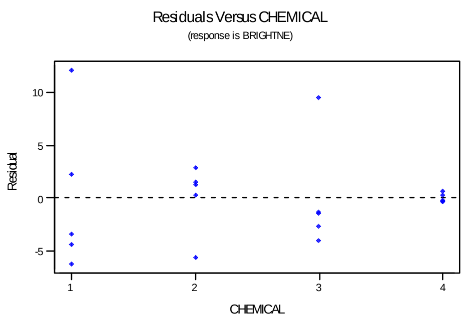
Source	DF	SS	MS	F	P
CHEMICAL	3	54.0	18.0	0.75	0.538
Error	16	384.0	24.0		
Total	19	438.0			

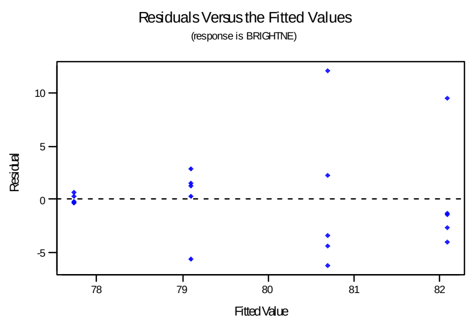
Do not reject H_0, there is no significant difference among the chemical types.

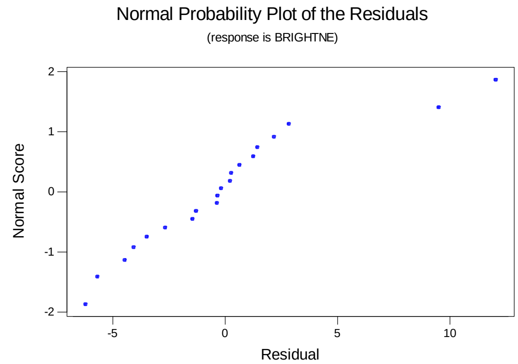
b) $\hat{\sigma}_\tau^2 = \dfrac{18.0 - 24.0}{5} = -1.2$ set equal to 0

c) $\hat{\sigma}^2 = 24.0$

d) Variability is smaller in chemical 4. There is some curvature in the normal probability plot.







13-24 a) $\hat{\sigma}^2_{total} = \hat{\sigma}^2_{position} + \hat{\sigma}^2 = 1.841$

b) $\dfrac{\hat{\sigma}^2_{position}}{\hat{\sigma}^2_{total}} = 0.646$

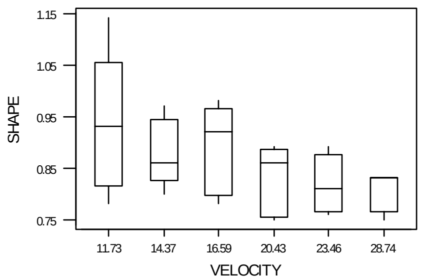
c) It could be reduced to 0.6522. This is a reduction of approximately 65%.

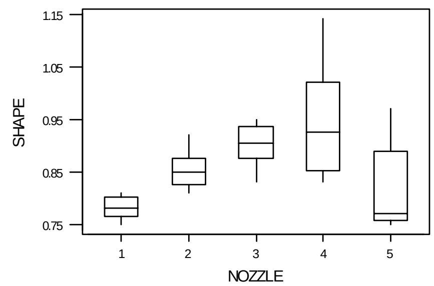
Section 13-4

13-25. a) Analysis of Variance for SHAPE

Source	DF	SS	MS	F	P
NOZZLE	4	0.102180	0.025545	8.92	0.000
VELOCITY	5	0.062867	0.012573	4.39	0.007
Error	20	0.057300	0.002865		
Total	29	0.222347			

Reject H_0, nozzle type affects shape measurement.



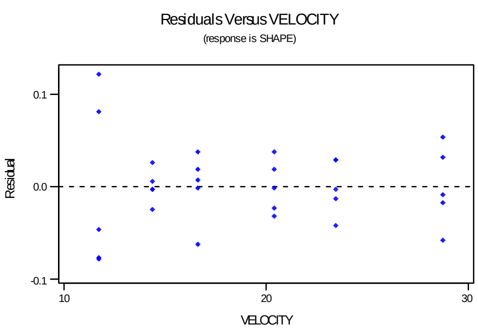


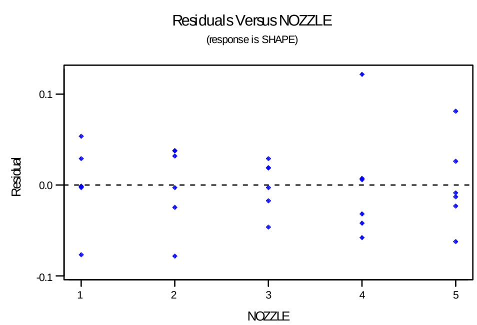
b) Fisher's pairwise comparisons
Family error rate = 0.268
Individual error rate = 0.0500
Critical value = 2.060
Intervals for (column level mean) - (row level mean)

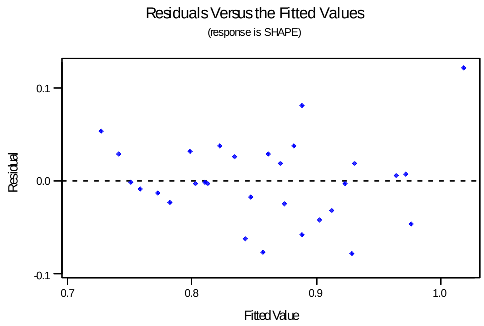
	1	2	3	4
2	-0.15412 0.01079			
3	-0.20246 -0.03754	-0.13079 0.03412		
4	-0.24412 -0.07921	-0.17246 -0.00754	-0.12412 0.04079	
5	-0.11412 0.05079	-0.04246 0.12246	0.00588 0.17079	0.04754 0.21246

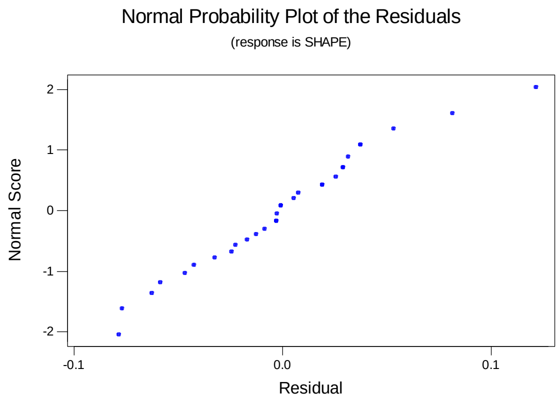
There are significant differences between levels 1 and 3, 4; 2 and 4; 3 and 5; and 4 and 5.

c) The residual analysis shows that there is some inequality of variance. The normal probability plot is acceptable.









13-26 a) Analysis of Variance of HARDNESS

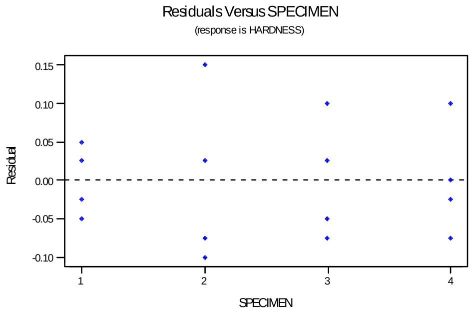
Source	DF	SS	MS	F	P
TIPTYPE	3	0.38500	0.12833	14.44	0.001
SPECIMEN	3	0.82500	0.27500	30.94	0.000
Error	9	0.08000	0.00889		
Total	15	1.29000			

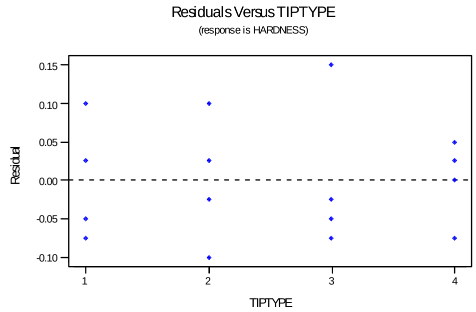
Reject H_0, and conclude that there are significant differences in hardness measurements between the tips.

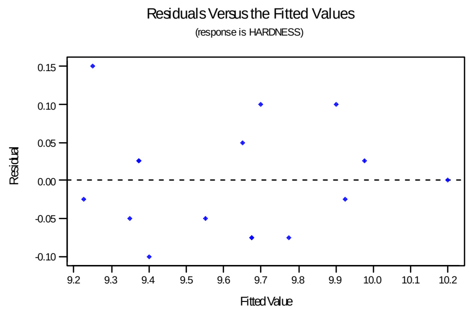
b)

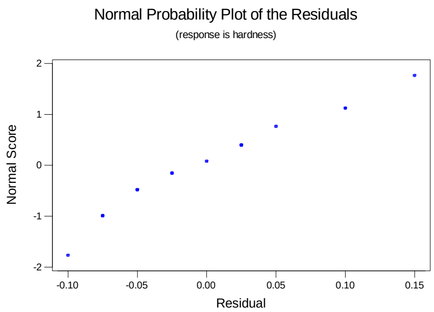
```
Fisher's pairwise comparisons
Family error rate = 0.184
Individual error rate = 0.0500
Critical value = 2.179
Intervals for (column level mean) - (row level mean)
             1           2           3
2      -0.4481
        0.3981
3      -0.2981     -0.2731
        0.5481      0.5731
4      -0.7231     -0.6981     -0.8481
        0.1231      0.1481     -0.0019

Significant difference between tip types 3 and 4
```

c) Residuals are acceptable.







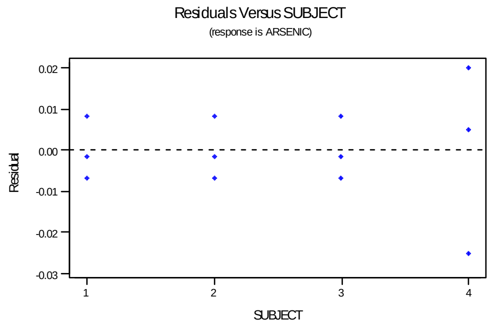


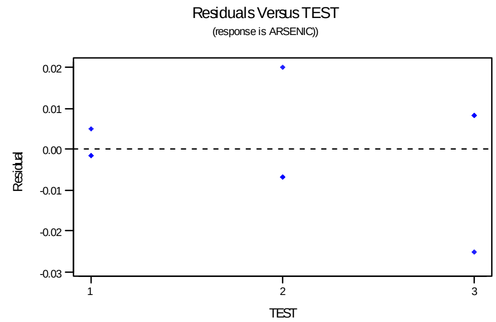
13-27. a) Analysis of Variance for ARSENIC

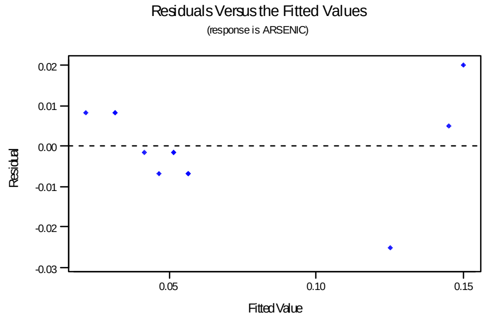
Source	DF	SS	MS	F	P
TEST	2	0.0014000	0.0007000	3.00	0.125
SUBJECT	3	0.0212250	0.0070750	30.32	0.001
Error	6	0.0014000	0.0002333		
Total	11	0.0240250			

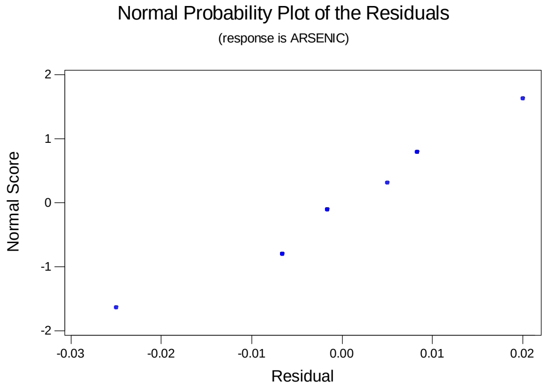
Do not reject H_0, there is no evidence of differences between the tests.

b) Some indication of variability increasing with the magnitude of the response.









13-28 a)Analysis of Variance of PROPECTIN

Source	DF	SS	MS	F	P
STORAGE	3	1972652	657551	4.33	0.014
LOT	8	1980499	247562	1.63	0.169
Error	24	3647150	151965		
Total	35	7600300			

Reject H_0, and conclude that the storage times affect the mean level of propectin.

b) P-value = 0.014

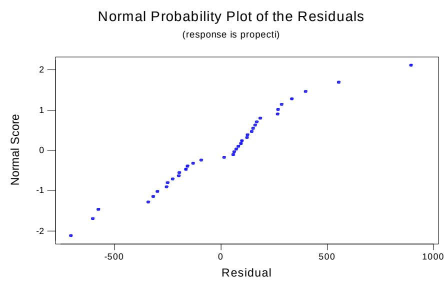
c)

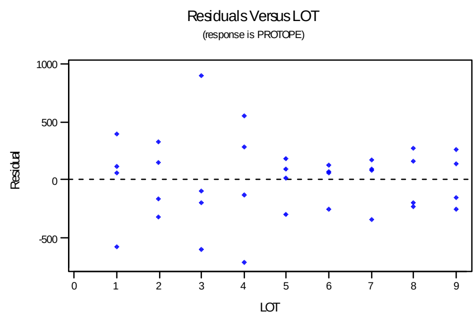
Fisher's pairwise comparisons
Family error rate = 0.196
Individual error rate = 0.0500
Critical value = 2.037
Intervals for (column level mean) - (row level mean)

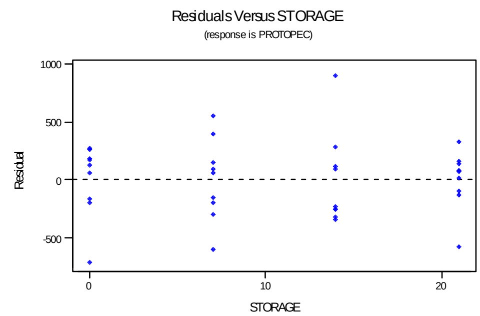
	0	7	14
7	-171 634		
14	-214 592	-445 360	
21	239 1045	8 813	50 856

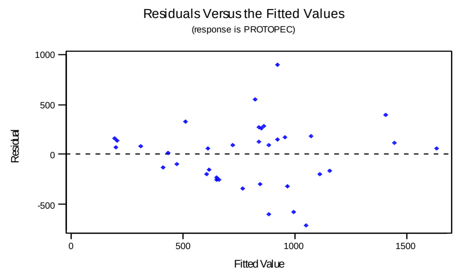
There are differences between 0 and 21 days; 7 and 21 days; and 14 and 21 days. The propectin levels are significantly different at 21 days from the other storage times so there is evidence that the mean level of propectin decreases with storage time. However, differences such as between 0 and 7 days and 7 and 14 days were not significant so that the level is not simply a linear function of storage days.

d) Observations from lot 3 at 14 days appear unusual. Otherwise, the residuals are acceptable.









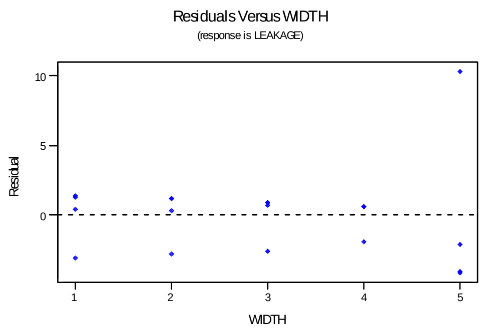
13-29. A version of the electronic data file has the reading for length 4 and width 5 as 2. It should be 20.

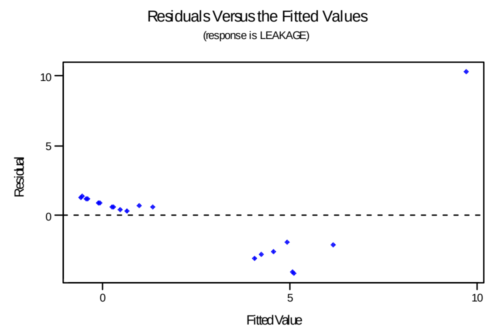
a) Analysis of Variance for LEAKAGE

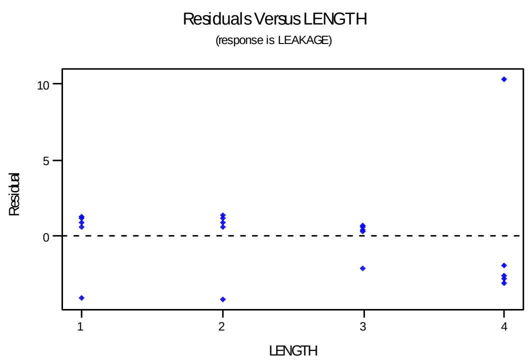
Source	DF	SS	MS	F	P
LENGTH	3	72.66	24.22	1.61	0.240
WIDTH	4	90.52	22.63	1.50	0.263
Error	12	180.83	15.07		
Total	19	344.01			

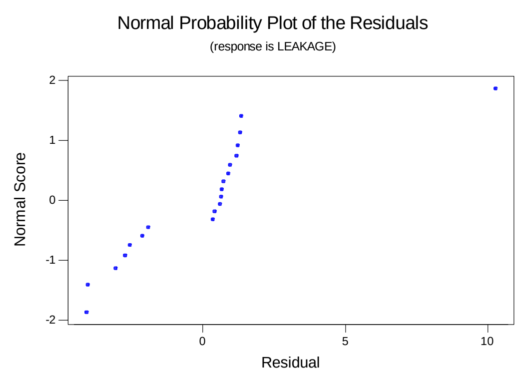
Do not reject H_0, mean leakage voltage does not depend on the channel length.

b) One unusual observation in width 5, length 4. There are some problems with the normal probability plot, including the unusual observation.









13-30 Analysis of Variance for LEAKAGE VOLTAGE

Source	DF	SS	MS	F	P
LENGTH	3	8.1775	2.7258	6.16	0.009
WIDTH	4	6.8380	1.7095	3.86	0.031
Error	12	5.3100	0.4425		
Total	19	20.3255			

Reject H_0. And conclude that the mean leakage voltage does depend on channel length. By removing the data point that was erroneous, the analysis results in a conclusion. The erroneous data point that was an obvious outlier had a strong effect the results of the experiment.

Supplemental Exercises

13-31. a) Analysis of Variance for RESISTANCE

Source	DF	SS	MS	F	P
ALLOY	2	10941.8	5470.9	76.09	0.000
Error	27	1941.4	71.9		
Total	29	12883.2			

Reject H_0, the type of alloy has a significant effect on mean contact resistance.

b)
```
Fisher's pairwise comparisons
  Family error rate = 0.119
  Individual error rate = 0.0500
  Critical value = 2.052
  Intervals for (column level mean) - (row level mean)
                  1           2
       2      -13.58
                1.98
       3      -50.88      -45.08
               -35.32     -29.52
```

There are differences in the mean resistance for alloy types 1 and 3; and types 2 and 3.

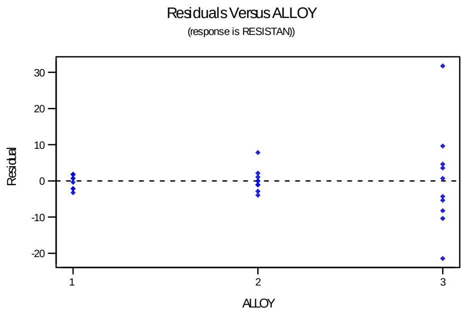
c) 99% confidence interval on the mean contact resistance for alloy 3

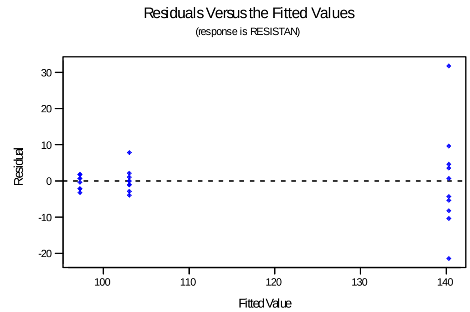
$$\bar{y}_3 - t_{0.005,27}\sqrt{\frac{MS_E}{n}} \le \mu_i \le \bar{y}_3 + t_{0.005,27}\sqrt{\frac{MS_E}{n}}$$

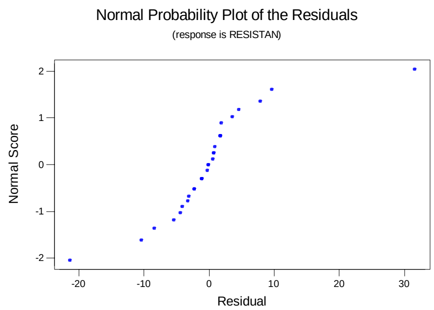
$$140.4 - 2.771\sqrt{\frac{71.9}{10}} \le \mu_3 \le 140.4 + 2.771\sqrt{\frac{71.9}{10}}$$

$$132.97 \le \mu_1 \le 147.83$$

d) Variability of the residuals increases with the response. The normal probability plot has some curvature in the tails, indicating a problem with the normality assumption. A transformation of the response should be conducted.







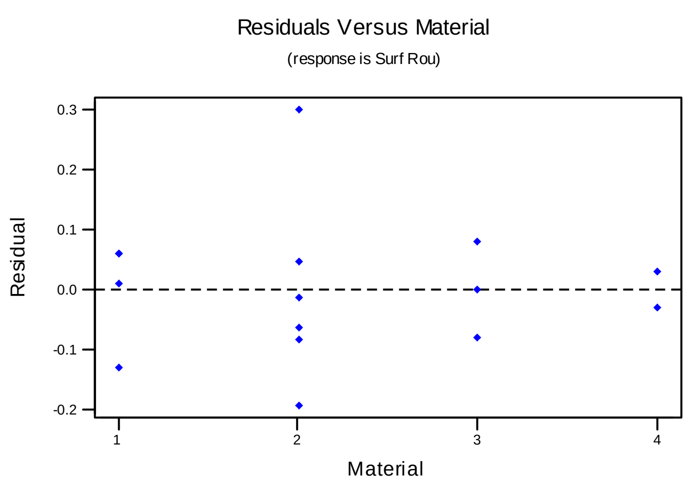
13-32 a)Analysis of Variance for SURFACE ROUGNESS

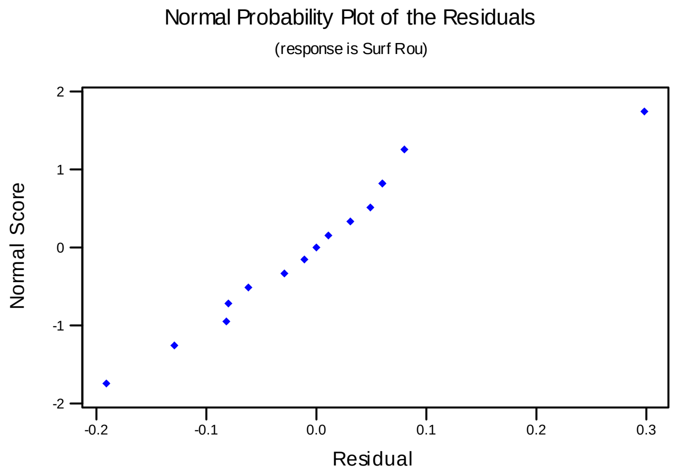
Analysis of Variance for y

Source	DF	SS	MS	F	P
Material	3	0.2402	0.0801	4.96	0.020
Error	11	0.1775	0.0161		
Total	14	0.4177			

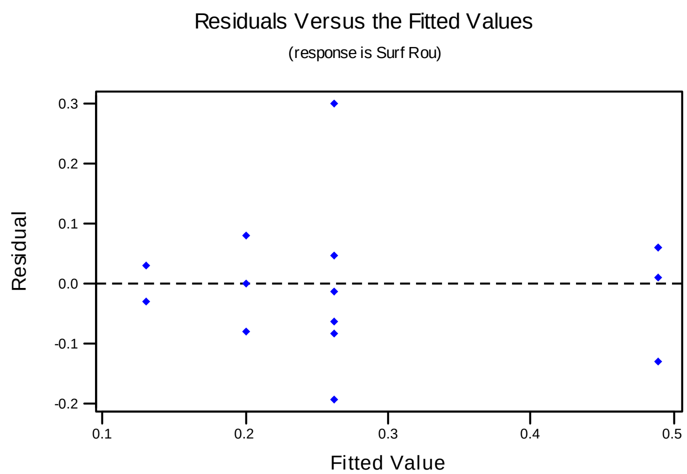
Reject H_0

b) One observation is an outlier.





c.) There appears to be a problem with constant variance. This may be due to the outlier in the data.



d) 95% confidence interval on the difference in the means of EC10 and EC1

$$\bar{y}_1 - \bar{y}_4 - t_{0.025,11}\sqrt{\frac{MS_E}{n_1} + \frac{MS_E}{n_2}} \le \mu_1 - \mu_3 \le \bar{y}_1 - \bar{y}_4 + t_{0.025,11}\sqrt{\frac{MS_E}{n_1} + \frac{MS_E}{n_2}}$$

$$(0.490 - 0.130) - 2.201\sqrt{\frac{(0.0161)}{4} + \frac{(0.0161)}{2}} \le \mu_1 - \mu_4 \le (0.490 - 0.130) + 2.201\sqrt{\frac{(0.0161)}{4} + \frac{(0.0161)}{2}}$$

$$0.118 \le \mu_1 - \mu_4 \le 0.602$$

13-33 Fisher's pairwise comparisons

```
Family error rate = 0.183
Individual error rate = 0.0500
Critical value = 2.201
Intervals for (column level mean) - (row level mean)
      1           2           3

2     0.0479
      0.4088

3     0.0765     -0.1360
      0.5035      0.2594

4     0.1179     -0.0966     -0.1852
      0.6021      0.3599      0.3252
```

There are significant in differences between the mean surface roughness for carbon material types 1 and 2; 1 and 3; and 1 and 4. Therefore, material type 1 is different from all the others and produces a higher mean surface roughness.

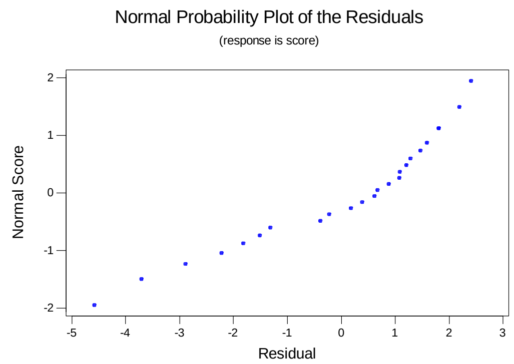
13-34 .a)Analysis of Variance for SCORE

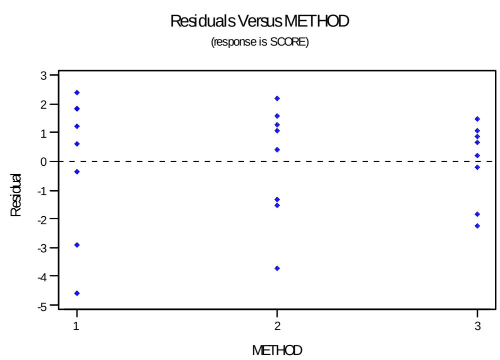
Source	DF	SS	MS	F	P
METHOD	2	13.55	6.78	1.68	0.211
Error	21	84.77	4.04		
Total	23	98.32			

Do not reject H_0

b) *P*-value = 0.211

c) There is some curvature in the normal probability plot. There appears to be some differences in the variability for the different methods. The variability for method one is larger than the variability for method 3.





d.) $\hat{\sigma}_{\tau}^{2} = \dfrac{MS_{Treatments} - MS_E}{n} = \dfrac{6.78 - 4.04}{8} = 0.342$

$\hat{\sigma}^{2} = MS_E = 4.04$

13-35. a)Analysis of Variance for VOLUME

Source	DF	SS	MS	F	P
TEMPERATURE	2	16480	8240	7.84	0.007
Error	12	12610	1051		
Total	14	29090			

Reject H_0.

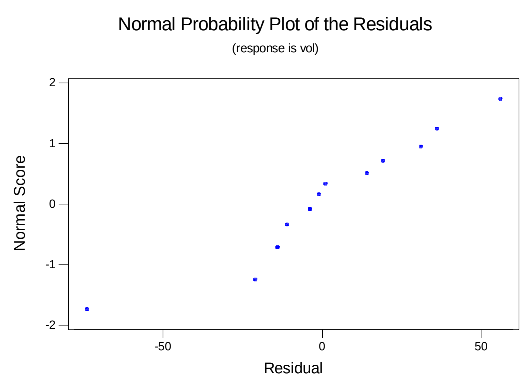
b) *P-value* = 0.007

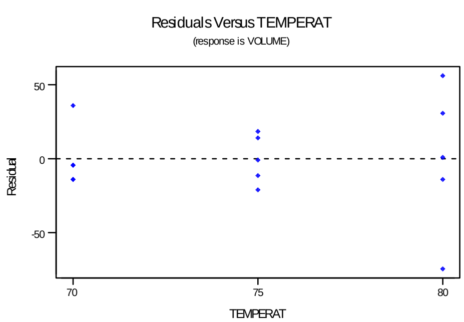
c) Fisher's pairwise comparisons

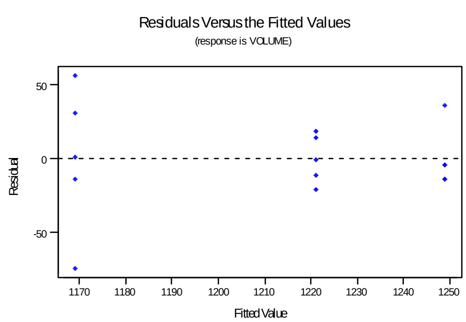
```
Family error rate = 0.116
Individual error rate = 0.0500
Critical value = 2.179
Intervals for (column level mean) - (row level mean)
              70          75
    75      -16.7
             72.7
    80       35.3         7.3
            124.7        96.7
```

There are significant differences in the mean volume for temperature levels 70 and 80; and 75 and 80. The highest temperature results in the smallest mean volume.

d) There are some relatively small differences in the variability at the different levels of temperature. The variability decreases with the fitted values. There is an unusual observation on the normal probability plot.







13-36 a) Analysis of Variance of Weight Gain

Source	DF	SS	MS	F	P
MEANWEIG	2	0.2227	0.1113	1.48	0.273
AIRTEMP	5	10.1852	2.0370	27.13	0.000
Error	10	0.7509	0.0751		
Total	17	11.1588			

Reject H_0 and conclude that the air temperature has an effect on the mean weight gain.

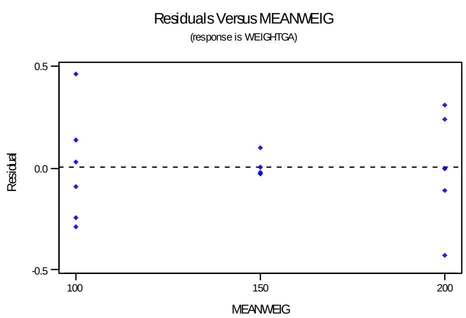
b) Fisher's pairwise comparisons

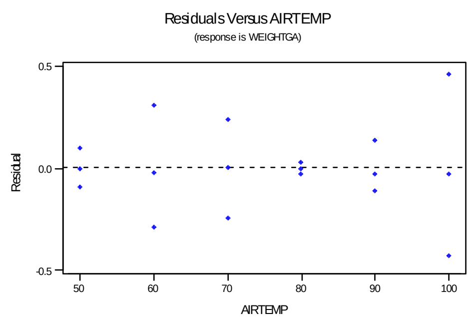
```
   Family error rate = 0.314
Individual error rate = 0.0500
Critical value = 2.179
Intervals for (column level mean) - (row level mean)
```

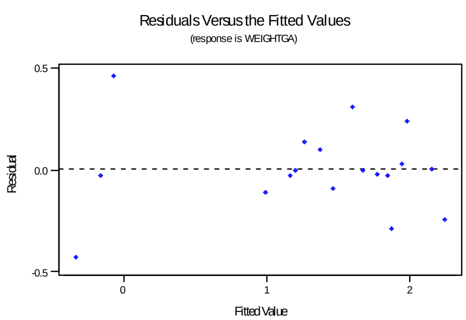
	50	60	70	80	90
60	-0.9101				
	0.1034				
70	-1.2901	-0.8868			
	-0.2766	0.1268			
80	-0.9834	-0.5801	-0.2001		
	0.0301	0.4334	0.8134		
90	-0.3034	0.0999	0.4799	0.1732	
	0.7101	1.1134	1.4934	1.1868	
100	1.0266	1.4299	1.8099	1.5032	0.8232
	2.0401	2.4434	2.8234	2.5168	1.8368

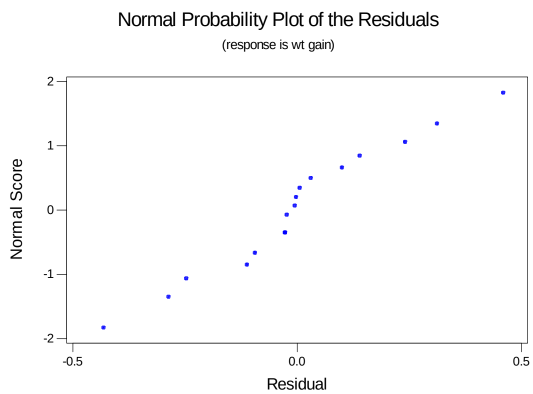
There are significant differences in the mean air temperature levels 50 and 70, 100; 60 and 90, 100; 70 and 90, 100; 80 and 90, 100; and 90 and 100. The mean of temperature level 100 is different from all the other temperatures.

c) There appears to be some problems with the assumption of constant variance.







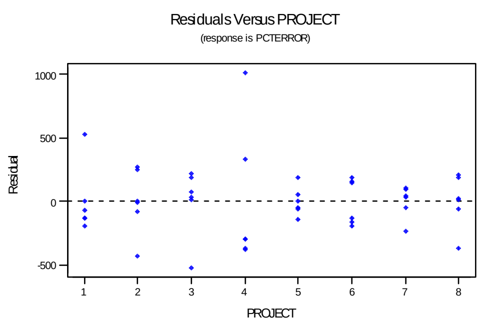


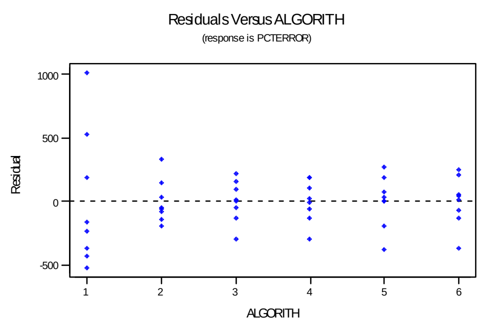
13-37. a) Analysis of Variance for PCTERROR

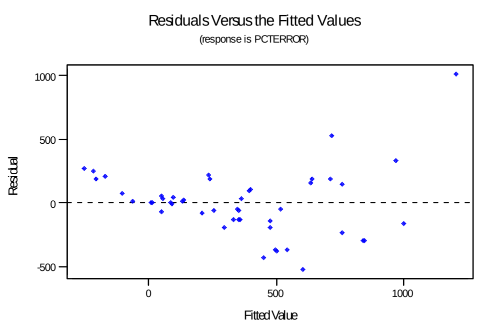
Source	DF	SS	MS	F	P
ALGORITH	5	2825746	565149	6.23	0.000
PROJECT	7	2710323	387189	4.27	0.002
Error	35	3175290	90723		
Total	47	8711358			

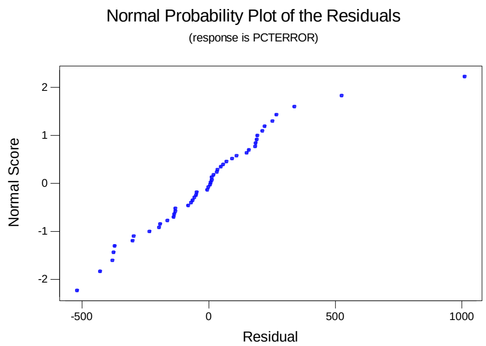
Reject H_0 , the algorithms are significantly different.

b) The residuals look acceptable, except there is one unusual point.









c) The best choice is algorithm 5 because it has the smallest mean and a low variability.

13-38 a) μ = (1+5+8+4)/4 = 4.5 and

$$\Phi^2 = \frac{4[(1-4.5)^2 + (5-4.5)^2 + (8-4.5)^2 + (4-4.5)^2]}{4(4)} = 6.25$$

$$\Phi = 2.5$$

Numerator degrees of freedom = $a - 1 = 3 = v_1$

Denominator degrees of freedom = $a(n-1) = 12 = v_2$

From Figure 13-6, β = 0.05 and the power = 1 - β = 0.95

b)

n	Φ^2	Φ	a(n-1)	β	Power = 1-β
4	6.25	2.5	12	0.05	0.95
3	4.6875	2.165	8	0.25	0.75

The sample size should be approximately *n* = 4.

13-39 a) μ=1.6, Φ^2 =0.284, Φ =0.5333

Numerator degrees of freedom = $a - 1 = 4 = v_1$

Denominator degrees of freedom = $a(n-1) = 15 = v_2$

From Chart Figure 13-6, $\beta \approx 0.8$ and the power = 1 - β = 0.2

b)

n	Φ^2	Φ	a(n-1)	β	Power = 1-β
50	3.56	1.89	245	0.05	0.95

The sample size should be approximately *n* = 50.

Mind Expanding Exercises

13-40 $MS_E = \dfrac{\sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij} - \bar{y}_i)^2}{a(n-1)}$ and $y_{ij} = \mu + a_i + \varepsilon_{ij}$. Then $y_{ij} - \bar{y}_i = \varepsilon_{ij} - \bar{\varepsilon}_{i.}$ and

$\dfrac{\sum_{j=1}^{n}(\varepsilon_{ij} - \bar{\varepsilon}_{i.})}{n-1}$ is recognized to be the sample variance of the independent random variables

$\varepsilon_{i1}, \varepsilon_{i2}, \ldots, \varepsilon_{in}$. Therefore, $E = \left[\dfrac{\sum_{j=1}^{n}(\varepsilon_{ij} - \bar{\varepsilon}_{i.})^2}{n-1}\right] = \sigma^2$ and $E(MS_E) = \sum_{i=1}^{a}\dfrac{\sigma^2}{a} = \sigma^2$.

The development would not change if the random effects model had been specified because $y_{ij} - \bar{y}_i = \varepsilon_{ij} - \bar{\varepsilon}_{i.}$ for this model also.

13-41 The two sample t-test rejects equality of means if the statistic

$t = \dfrac{|\bar{y}_1 - \bar{y}_2|}{s_p\sqrt{\frac{1}{n} + \frac{1}{n}}} = \dfrac{|\bar{y}_1 - \bar{y}_2|}{s_p\sqrt{\frac{2}{n}}}$ is too large. The ANOVA F-test rejects equality of means if

$F = \dfrac{n\sum_{i=1}^{2}(\bar{y}_{i.} - \bar{y}_{..})^2}{MS_E}$ is too large.

Now, $F = \dfrac{\frac{n}{2}(\bar{y}_{1.} - \bar{y}_{2.})^2}{MS_E} = \dfrac{(\bar{y}_{1.} - \bar{y}_{2.})^2}{MS_E\,\frac{2}{n}}$ and $MS_E = s_p^2$.

Consequently, $F = t^2$. Also, the distribution the square of a t random variable with a(n - 1) degrees of freedom is an F distribution with 1 and a(n - 1) degrees of freedom. Therefore, if the tabulated t value for a two-sided t-test of size α **is t_0**, then the tabulated F value for the F test above is t_0^2. Therefore, $t > t_0$ whenever $F = t^2 > t_0^2$ and the two tests are identical.

13-42 $MS_E = \dfrac{\sum_{i=1}^{2}\sum_{j=1}^{n}(y_{ij} - \bar{y}_{i.})^2}{2(n-1)}$ and $\dfrac{\sum_{j=1}^{n}(y_{ij} - \bar{y}_{i.})^2}{n-1}$ is recognized as the sample standard deviation

calculated from the data from population i. Then, $MS_E = \dfrac{s_1^2 + s_2^2}{2}$ which is the pooled variance estimate

used in the t-test.

13-43 $V(\sum_{i=1}^{a} c_i Y_{i.}) = \sum_{i=1}^{a} c_i^2 V(Y_{i.})$ from the independence of $Y_{1.}, Y_{2.}, \ldots, Y_{a.}$.

Also, $V(Y_{i.}) = n_i\sigma_i^2$. Then, $V(\sum_{i=1}^{a} c_i Y_{i.}) = \sigma^2\sum_{i=1}^{a} c_i^2 n_i$

13-44 If b, c, and d are the coefficients of three orthogonal contrasts, it can be shown that

$$\frac{(\sum_{i=1}^{4} b_i y_{i.})^2}{\sum_{i=1}^{a} b_i^2} + \frac{(\sum_{i=1}^{a} c_i y_{i.})^2}{\sum_{i=1}^{a} c_i^2} + \frac{(\sum_{i=1}^{a} d_i y_{i.})^2}{\sum_{i=1}^{a} d_i^2} = \sum_{i=1}^{a} y_{i.}^2 - \frac{(\sum_{i=1}^{a} y_{i.})^2}{a}$$

always holds. Upon dividing both sides by n, we have $Q_1^2 + Q_2^2 + Q_3^2 = \sum_{i=1}^{a} \frac{y_i^2}{n} - \frac{y_{..}^2}{N}$ which equals $SS_{treatments}$. The equation above can be obtained from a geometrical argument. The square of the distance of any point in four-dimensional space from the zero point can be expressed as the sum of the squared distance along four orthogonal axes. Let one of the axes be the 45 degree line and let the point be ($y_{1.}, y_{2.}, y_{3.}, y_{4.}$). The three orthogonal contrasts are the other three axes. The square of the distance of the point from the origin is $\sum_{i=1}^{a} y_{i.}^2$ and this equals the sum of the squared distances along each of the four axes.

13-45 Because $\Phi^2 = \frac{n\sum_{i=1}^{a} (\mu_i - \overline{\mu})^2}{a\sigma^2}$, we only need to shows that $\frac{D^2}{2} \le \sum_{i=1}^{a} (\mu_i - \overline{\mu})^2$.

Let μ_1 and μ_2 denote the means that differ by D. Now, $(\mu_1 - x)^2 + (\mu_2 - x)^2$ is minimized for x equal to the mean of μ_1 and μ_2. Therefore,

$$(\mu_1 - \frac{\mu_1 + \mu_2}{2})^2 + (\mu_2 - \frac{\mu_1 + \mu_2}{2})^2 \le (\mu_1 - \overline{\mu})^2 + (\mu_2 - \overline{\mu})^2 \le \sum_{i=1}^{a} (\mu_i - \overline{\mu})^2$$

Then, $\left(\frac{\mu_1 - \mu_2}{2}\right)^2 + \left(\frac{\mu_2 - \mu_1}{2}\right)^2 = \frac{D^2}{4} + \frac{D^2}{4} = \frac{D^2}{2} \le \sum_{i=1}^{a} (\mu_i - \overline{\mu})^2$.

13-46 $MS_E = \frac{\sum_{i=1}^{a}\sum_{j=1}^{n} (y_{ij} - \bar{y}_{i.})^2}{a(n-1)} = \frac{\sum_{i=1}^{a} s_i^2}{a}$ where $s_i^2 = \frac{\sum_{j=1}^{n} (y_{ij} - \bar{y}_{i.})^2}{n-1}$. Because s_i^2 is the sample variance of $y_{i1}, y_{i2}, ..., y_{in}$, $\frac{(n-1)S_i^2}{\sigma^2}$ has a chi-square distribution with n-1 degrees of freedom. Then, $\frac{a(n-1)MS_E}{\sigma^2}$ is a sum of independent chi-square random variables. Consequently, $\frac{a(n-1)MS_E}{\sigma^2}$ has a chi-square distribution with a(n - 1) degrees of freedom. Consequently,

$$P(\chi^2_{1-\frac{\alpha}{2},a(n-1)} \le \frac{a(n-1)MS_E}{\sigma^2} \le \chi^2_{\frac{\alpha}{2},a(n-1)}) = 1-\alpha$$

$$= P\left(\frac{a(n-1)MS_E}{\chi^2_{\frac{\alpha}{2},a(n-1)}} \le \sigma^2 \le \frac{a(n-1)MS_E}{\chi^2_{1-\frac{\alpha}{2},a(n-1)}}\right)$$

Using the fact that a(n - 1) = N - a completes the derivation.

13-47 From Exercise 13-46, $\frac{(N-a)MS_E}{\sigma^2}$ has a chi-square distribution with *N* - *a* degrees of freedom. Now,

$V(\bar{Y}_{i.}) = \sigma_\tau^2 + \frac{\sigma^2}{n}$ and mean square treatment = MS_T is *n* times the sample variance of

$\bar{y}_{1.}, \bar{y}_{2.}, \ldots, \bar{y}_{a.}$. Therefore, $\frac{(a-1)MS_T}{n(\sigma_\tau^2 + \frac{\sigma^2}{n})} = \frac{(a-1)MS_T}{n\sigma_\tau^2 + \sigma^2}$ has a chi-squared distribution with *a* - 1 degrees of freedom. Using the independence of MS_T and MS_E, we conclude that

$\left(\frac{MS_T}{n\sigma_\tau^2 + \sigma^2}\right) / \left(\frac{MS_E}{\sigma^2}\right)$ has an $F_{(a-1),(N-a)}$ distribution.

Therefore,

$$P\left(f_{1-\frac{\alpha}{2},a-1,N-a} \le \frac{MS_T}{MS_E}\frac{\sigma^2}{n\sigma_\tau^2 + \sigma^2} \le f_{\frac{\alpha}{2},a-1,N-a}\right) = 1-\alpha$$

$$= P\left(\frac{1}{n}\left[\frac{1}{f_{\frac{\alpha}{2},a-1,N-a}}\frac{MS_T}{MS_E} - 1\right] \le \frac{\sigma_\tau^2}{\sigma^2} \le \frac{1}{n}\left[\frac{1}{f_{1-\frac{\alpha}{2},a-1,N-a}}\frac{MS_T}{MS_E} - 1\right]\right)$$

by an algebraic solution for $\frac{\sigma_\tau^2}{\sigma^2}$ and $P(L \le \frac{\sigma_\tau^2}{\sigma^2} \le U)$.

13-48 As in Exercise 13-47, $\frac{MS_T}{MS_E}\frac{\sigma^2}{n\sigma_\tau^2 + \sigma^2}$ has an $F_{(a-1),(N-a)}$ distribution.

and

$$1-\alpha = P(L \le \frac{\sigma_\tau^2}{\sigma^2} \le U)$$

$$= P(\frac{1}{U} \le \frac{\sigma^2}{\sigma_\tau^2} \le \frac{1}{L})$$

$$= P(\frac{1}{U} + 1 \le \frac{\sigma^2}{\sigma_\tau^2} + 1 \le \frac{1}{L} + 1)$$

$$= P(\frac{L}{L+1} \le \frac{\sigma_\tau^2}{\sigma^2 + \sigma_\tau^2} \le \frac{U}{U+1})$$

13-49 From Exercise 13-48,

$$1-\alpha = P(L \le \frac{\sigma_\tau^2}{\sigma^2} \le U)$$

$$= P(L+1 \le \frac{\sigma_\tau^2 + 1}{\sigma^2} \le U+1)$$

$$= P(L+1 \le \frac{\sigma_\tau^2 + \sigma^2}{\sigma^2} \le U+1)$$

$$= P(\frac{1}{U+1} \le \frac{\sigma^2}{\sigma^2 + \sigma_\tau^2} \le \frac{1}{L+1})$$

Therefore, $(\frac{1}{U+1}, \frac{1}{L+1})$ is a confidence interval for $\frac{\sigma^2}{\sigma_\tau^2 + \sigma^2}$

13-50 $MS_T = \frac{\sum_{i=1}^{a} n_i(\bar{y}_{i.} - \bar{y}_{..})^2}{a-1}$ and for any random variable *X*, $E(X^2) = V(X) + [E(X)]^2$.

Then,

$$E(MS_T) = \frac{\sum_{i=1}^{a} n_i \{V(\bar{Y}_{i.} - \bar{Y}_{..}) + [E(\bar{Y}_{i.} - \bar{Y}_{..})]^2\}}{a-1}$$

Now, $\bar{Y}_{1.} - \bar{Y}_{..} = (\frac{1}{n_1} - \frac{1}{N})Y_{11} + \ldots + (\frac{1}{n_1} - \frac{1}{N})Y_{1n_1} - \frac{1}{N}Y_{21} - \ldots - \frac{1}{N}Y_{2n_2} - \ldots - \frac{1}{N}Y_{a1} - \ldots - \frac{1}{N}Y_{an_a}$

and

$$V(\bar{Y}_{1.} - \bar{Y}_{..}) = \left((\tfrac{1}{n_1} - \tfrac{1}{N})^2 n_1 + \frac{N - n_1}{N^2} \right)\sigma^2 = (\tfrac{1}{n_1} - \tfrac{1}{N})\sigma^2$$

$$E(\bar{Y}_{1.} - \bar{Y}_{..}) = (\tfrac{1}{n_1} - \tfrac{1}{N})n_1\lambda_1 - \frac{n_2}{N}\lambda_2 - \ldots - \frac{n_a}{N}\lambda_a = \lambda_1 \text{ from the constraint}$$

Then,

$$E(MS_T) = \frac{\sum_{i=1}^{a} n_i \{(\frac{1}{n_i} - \frac{1}{N})\sigma^2 + \lambda_i^2\}}{a-1} = \frac{\sum_{i=1}^{a} [(1 - \frac{n_i}{N})\sigma^2 + n_i\lambda_i^2]}{a-1}$$

$$= \sigma^2 + \frac{\sum_{i=1}^{a} n_i\lambda_i^2}{a-1}$$

Because $E(MS_E) = \sigma^2$, this does suggest that the null hypothesis is as given in the exercise.

13-51 a) If A is the accuracy of the interval, then $t_{\frac{\alpha}{2},a(n-1)}\sqrt{\frac{2MS_E}{n}} = A$

Squaring both sides yields $t^2_{\frac{\alpha}{2},a(n-1)}\frac{2MS_E}{n} = A^2$

As in Exercise 13-41, $t^2_{\frac{\alpha}{2},a(n-1)} = F_{\alpha,1,a(n-1)}$. Then,

$$n = \frac{2MS_E F_{\alpha,1,a(n-1)}}{A^2}$$

b) Because n determines one of the degrees of freedom of the tabulated F value on the right-side of the equation in part (a), some approximation is needed. Because the value for a 95% confidence interval based on a normal distribution is 1.96, we approximate $t_{\frac{\alpha}{2},a(n-1)}$ by 2 and we approximate

$t^2_{\frac{\alpha}{2},a(n-1)} = F_{\alpha,1,a(n-1)}$ by 4.

Then, $n = \frac{2(4)(4)}{4} = 8$. With n = 8, $a(n - 1) = 35$ and $F_{0.05,1,35} = 4.12$.

The value 4.12 can be used for F in the equation for n and a new value can be computed for n as

$$n = \frac{2(4)(4.12)}{4} = 8.24 \cong 8$$

Because the solution for n did not change, we can use n = 8. If needed, another iteration could be used to refine the value of n.