CHAPTER 1

Sections 1-1 and 1-2

1-1. Sample average:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{\sum_{i=1}^{8} x_i}{8}$$
$$= \frac{592.035}{8}$$
$$= 74.0044 \text{ mm}$$

Sample variance:

$$\sum_{i=1}^{8} x_i = 592.035$$
$$\sum_{i=1}^{8} x_i^2 = 43813.18031$$

$$s^2 = \frac{\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}}{n-1}$$
$$= \frac{43813.18031 - \frac{(592.035)^2}{8}}{8-1}$$
$$= \frac{0.0001569}{7}$$
$$= 0.000022414 \text{ (mm)}^2$$

Sample standard deviation:

$$s = \sqrt{0.000022414}$$
$$= 0.00473 \text{ mm}$$

The sample standard deviation could also be found using

$$s = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}}$$

where

$$\sum_{i=1}^{8} (x_i - \bar{x})^2 = 0.0001569$$

Dot Diagram:

```
                   .. ...:           .
-------+---------+---------+---------+---------+---------diameter
 73.9920   74.0000   74.0080   74.0160   74.0240   74.0320
```

There appears to be an outlier in the data set.

1-2. Sample average:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{\sum_{i=1}^{19} x_i}{19}$$
$$= \frac{272.82}{19}$$
$$= 14.359 \text{ min}$$

Sample variance:

$$\sum_{i=1}^{19} x_i = 272.82$$
$$\sum_{i=1}^{19} x_i^2 = 10333.8964$$

$$s^2 = \frac{\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}}{n-1}$$
$$= \frac{10333.8964 - \frac{(272.82)^2}{19}}{19-1}$$
$$= \frac{6416.49}{18}$$
$$= 356.47 \text{ (min)}^2$$

Sample standard deviation:

$$s = \sqrt{356.47}$$
$$= 18.88 \text{ min}$$

The sample standard deviation could also be found using

$$s = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}}$$

where

$$\sum_{i=1}^{19} (x_i - \bar{x})^2 = 6416.46$$

Dot Diagram:

```
  . .
.::::.:. .             ....                      .
+---------+---------+---------+---------+---------+-------Breakdown
0        15        30        45        60        75
```

1-3. Sample average:

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{\sum_{i=1}^{12} x_i}{12}$$
$$= \frac{84727}{12}$$
$$= 7068.1 \quad \text{yards}$$

Sample variance:

$$\sum_{i=1}^{12} x_i = 84817$$
$$\sum_{i=1}^{19} x_i^2 = 600057949$$

$$s^2 = \frac{\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}}{n-1}$$
$$= \frac{600057949 - \frac{(84817)^2}{12}}{12-1}$$
$$= \frac{564324.92}{11}$$
$$= 51302.265 \quad (\text{yards})^2$$

Sample standard deviation:

$$s = \sqrt{51302.265}$$
$$= 226.5 \quad \text{yards}$$

The sample standard deviation could also be found using

$$s = \sqrt{\frac{\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2}{n-1}}$$

where

$$\sum_{i=1}^{12}\left(x_i - \overline{x}\right)^2 = 564324.92$$

Dot Diagram:

```
      .     .  :  .. ..   :                           :
  -+---------+---------+---------+---------+---------+-----C1
6750      6900      7050      7200      7350      7500
```

1-4. Sample mean:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{\sum_{i=1}^{18} x_i}{18}$$

$$= \frac{2272}{18}$$

$$= 126.22 \quad \text{kN}$$

Sample variance:

$$\sum_{i=1}^{18} x_i = 2272$$

$$\sum_{i=1}^{18} x_i^2 = 298392$$

$$s^2 = \frac{\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}}{n-1}$$

$$= \frac{298392 - \frac{(2272)^2}{18}}{18-1}$$

$$= \frac{11615.11}{17}$$

$$= 683.24 \quad (\text{kN})^2$$

Sample standard deviation:

$$s = \sqrt{683.24}$$

$$= 26.14 \quad \text{kN}$$

The sample standard deviation could also be found using

$$s = \sqrt{\frac{\sum_{i=1}^{n} \left(x_i - \bar{x}\right)^2}{n-1}}$$

where

$$\sum_{i=1}^{18} \left(x_i - \bar{x}\right)^2 = 11615.11$$

Dot Diagram:

```
         .
     :   ::  .           ::       .          . : .    .
 +---------+---------+---------+---------+---------+-------yield
90       105       120       135       150       165
```

1-5. Sample average:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{\sum_{i=1}^{8} x_i}{8}$$

$$= \frac{351.8}{8}$$

$$= 43.975$$

Sample variance:

$$\sum_{i=1}^{8} x_i = 351.8$$

$$\sum_{i=1}^{19} x_i^2 = 16528.403$$

$$s^2 = \frac{\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}}{n-1}$$

$$= \frac{16528.043 - \frac{(351.8)^2}{8}}{8-1}$$

$$= \frac{1057.998}{7}$$

$$= 151.143$$

Sample standard deviation:

$$s = \sqrt{151.143}$$

$$= 12.294$$

The sample standard deviation could also be found using

$$s = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}}$$

where

$$\sum_{i=1}^{8} (x_i - \bar{x})^2 = 1057.998$$

Dot Diagram:

```
      .            . ..   .         ..                  .
   +---------+---------+---------+---------+---------+-------
 24.0      32.0      40.0      48.0      56.0      64.0
```

1-6. Sample average:

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{\sum_{i=1}^{35} x_i}{35}$$
$$= \frac{28368}{35}$$
$$= 810.514 \quad \text{watts / m}^2$$

Sample variance:

$$\sum_{i=1}^{19} x_i = 28368$$
$$\sum_{i=1}^{19} x_i^2 = 23552500$$

$$s^2 = \frac{\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}}{n-1}$$
$$= \frac{23552500 - \frac{(28368)^2}{35}}{35-1}$$
$$= \frac{559830.743}{34}$$
$$= 16465.61 \quad (\text{watts / m}^2)^2$$

Sample standard deviation:

$$s = \sqrt{16465.61}$$
$$= 128.32 \quad \text{watts / m}^2$$

The sample standard deviation could also be found using

$$s = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \overline{x})^2}{n-1}}$$

where

$$\sum_{i=1}^{35} (x_i - \overline{x})^2 = 559830.743$$

1-7. The only data set that may have resulted from a designed experiment is in Exercise 1-4.

1-8. $\mu = \frac{6905}{1270} = 5.44$; The value 5.44 is the population mean since the actual physical population of all flight times during the operation is available.

Supplemental Exercises

1-9. a) $\overline{x} = \dfrac{\sum_{i=1}^{n} x_i}{n} = \dfrac{\sum_{i=1}^{9} x_i}{9} = \dfrac{19.56}{9} = 2.17$

b) $s^2 = \dfrac{\sum_{i=1}^{n} x_i^2 - \dfrac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}}{n-1} = \dfrac{45.9532 - \dfrac{(19.56)^2}{9}}{9-1} = \dfrac{559830.743}{8} = 0.43$

$s = \sqrt{0.43} = 0.656$

c) Dot Diagram

```
.     .            .     .  .         .  .          ..
-------+---------+---------+---------+---------+---------C1
    1.40      1.75      2.10      2.45      2.80      3.15
```

1-10. Dot Diagram

```
                                                 . .
     .     :        .:  ... . . .:  :. . .:. .:: :::
-----+---------+---------+---------+---------+---------+-x
   500       600       700       800       900      1000
```

1-11. a) Dot Diagram

```
      .            . ..   .        ..                      .
   +---------+---------+---------+---------+---------+-------x
24.0      32.0      40.0      48.0      56.0      64.0
```

b) Dot Diagram

```
    .
    :.    .        .       .                       .
-------+---------+---------+---------+---------+---------x
      12        24        36        48        60        72
```

The data from the low-resolution screen has a much lower center than the high-resolution data, and it is less symmetric with one large value (64.63).

1-12. a) $\overline{x} = \dfrac{\sum_{i=1}^{n} x_i}{n} = \dfrac{\sum_{i=1}^{8} x_i}{8} = \dfrac{5747}{8} = 7.184$

b) $s^2 = \dfrac{\sum_{i=1}^{n} x_i^2 - \dfrac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}}{n-1} = \dfrac{412.853 - \dfrac{(57.47)^2}{8}}{8-1} = \dfrac{0.003}{7} = 0.000427$

$s = \sqrt{0.000427} = 0.02066$

c) Examples: repeatability of the test equipment, time lag between samples, operator skill.

1-13. $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{\sum_{i=1}^{9} x_i}{9} = \frac{748}{9} = 83.11$

$$s^2 = \frac{\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}}{n-1} = \frac{62572 - \frac{(748)^2}{9}}{9-1} = \frac{405}{8} = 50.6$$

$s = \sqrt{50.6} = 7.114$

Dot Diagram:

```
 .         . . . . . . .                             .
---+---------+---------+---------+---------+---------+---drag
75.0      80.0      85.0      90.0      95.0     100.0
```

1-14. a) $\bar{x} = 65.85$
$s = 12.16$

b) Dot Diagram

```
                                      :  :
  .        .    .   .  ..   .: .: .  .:..: .. :: .... ..
-+---------+---------+---------+---------+---------+-----temp
30        40        50        60        70        80
```

c) Removing the smallest observation (31), the sample mean and standard deviation become
$\bar{x} = 66.86$
$s = 10.74$

<u>Mind-Expanding Exercises</u>

1-15. The new data will have a resulting variance of $(10)^2 s^2$. The subtraction of a constant does not effect the variance of the sample. However, the multiplication by a constant results in a multiplication of the variance by the constant squared.

1-16. $\sum_{i=1}^{n} (x_i - a)^2 = \sum_{i=1}^{n} (x_i - \bar{x}) + n(\bar{x} - a)^2$; The sum written in this form shows that the quantity is minimized when $a = \bar{x}$.

1-17. $\sum_{i=1}^{n} (x_i - \bar{x})^2$ will be the smaller of the two quantities if $\mu \neq \bar{x}$ since $\sum_{i=1}^{n} (x_i - \mu)^2$ is minimized when $\mu = \bar{x}$.

1-18. $\bar{y} = a + b\bar{x}$ $\qquad s_y = bs_x$

1-19. $\bar{x} = 835°F$, $\qquad s_x = 10.5°F$

Let y_i = temperature in °C $\qquad y_i = \frac{5}{9}(x_i - 32) = -\frac{160}{9} + \frac{5}{9} x_i$

Using the results found in Exercise 1-18 with $a = -160/9$ and $b = 5/9$, the sample mean and sample standard deviation in degrees Celsius are

$\bar{y} = -\frac{160}{9} + \frac{5}{9}(835) = 446.11^\circ C$ $\qquad s_y = \frac{5}{9}(10.5) = 5.833^\circ C$

1-20. Using the results found in Exercise 1-18 with $a = -\frac{\bar{x}}{s}$ and $b = 1/s$, the mean and standard deviation of the z_i are

$\bar{z} = 0$ and $s_Z = 1$.

CHAPTER 2

Section 2-2

2-1.

Stem-and-leaf display for Problem 2-1.Octane: unit = 0.1   1|2 represents 1.2

```
   1    83|4
   3    84|33
   4    85|3
   7    86|777
  13    87|456789
  24    88|23334556679
  34    89|0233678899
 (13)   90|0111344456789
  35    91|0001112256688
  22    92|22236777
  14    93|023347
   8    94|2247
   4    95|
   4    96|15

        HI|988,1003
```

2-2

a) Stem-and-leaf display for Problem 2-2 cycles: unit = 100   1|2 represents 1200

```
   1     0T|3
   1     0F|
   5     0S|7777
  10     0o|88899
  22     1*|000000011111
  33     1T|22222223333
 (15)    1F|444445555555555
  22     1S|66667777777
  11     1o|888899
   5     2*|011
   2     2T|22
```

b) No, only 5/70 survived beyond 2000 cycles.

2-3.

Stem-and-leaf display for Problem 2-3.cotton: unit = 0.1   1|2 represents 1.2

```
   1    32*|1
   6    32o|56789
   9    33*|114
  17    33o|56666688
  24    34*|0111223
 (14)   34o|55666667777779
  26    35*|001112344
  17    35o|56789
  12    36*|234
   9    36o|6888
   5    37*|13
   3    37o|689
```

2-4.

Stem-and-leaf display for Problem 2-4.yield: unit = 1   1|2 represents 12

```
   1    7o|8
   1    8*|
   7    8T|223333
  21    8F|44444444555555
  38    8S|66666666667777777
(11)    8o|88888999999
  41    9*|00000000001111
  27    9T|22233333
  19    9F|444444445555
   7    9S|666677
   1    9o|8
```

2-5.

Descriptive Statistics

```
Variable        N  Median    Q1      Q3
octane         82  90.400  88.575  92.209
```

2-6.

Descriptive Statistics

```
Variable        N  Median    Q1        Q3
cycles         70  1436.5  1097.8   1735.0
```

2-7.

Descriptive Statistics

```
Variable        N     Mean   Median   Mode
cotton         64   34.798   34.700   34.7
```

The values of the mean, median, and mode coincide. This indicates that the distribution is symmetric.

2-8.

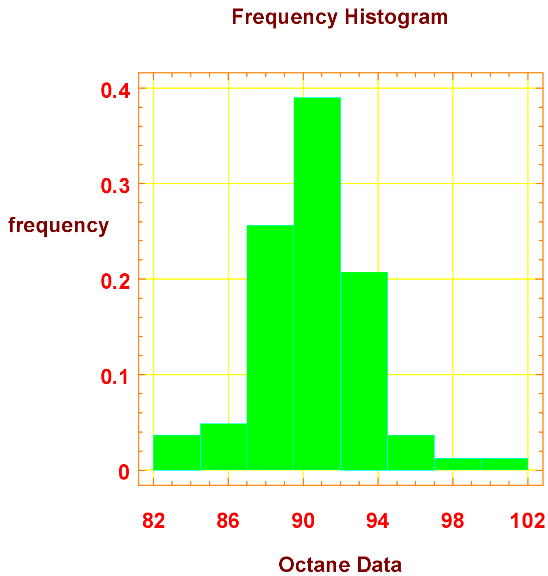
Descriptive Statistics

```
Variable        N Median     Q1        Q3
yield          90   89.250 86.100   93.125
```

Section 2-3

2-9.

```
                    Frequency Tabulation for Problem 2-9.Octane
--------------------------------------------------------------------------------
        Lower    Upper                          Relative  Cumulative  Cum. Rel.
Class   Limit    Limit   Midpoint   Frequency  Frequency  Frequency   Frequency
--------------------------------------------------------------------------------
   at or below   82.000                     0      .0000          0       .0000
   1    82.000   84.500    83.250           3      .0366          3       .0366
   2    84.500   87.000    85.750           4      .0488          7       .0854
   3    87.000   89.500    88.250          21      .2561         28       .3415
   4    89.500   92.000    90.750          32      .3902         60       .7317
   5    92.000   94.500    93.250          17      .2073         77       .9390
   6    94.500   97.000    95.750           3      .0366         80       .9756
   7    97.000   99.500    98.250           1      .0122         81       .9878
   8    99.500  102.000   100.750           1      .0122         82      1.0000
above  102.000                              0      .0000         82      1.0000
--------------------------------------------------------------------------------
Mean = 90.5256    Standard Deviation = 2.9052    Median = 90.4
```

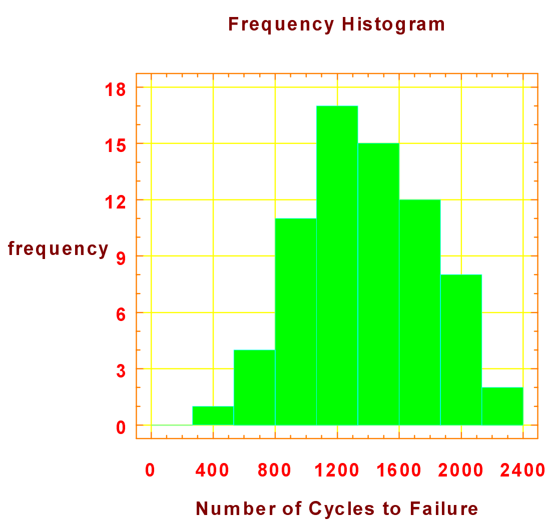


2-10.

Frequency Tabulation for Problem 2-10.Cycles

Class	Lower Limit	Upper Limit	Midpoint	Frequency	Relative Frequency	Cumulative Frequency	Cum. Rel. Frequency
at or below		.000		0	.0000	0	.0000
1	.000	266.667	133.333	0	.0000	0	.0000
2	266.667	533.333	400.000	1	.0143	1	.0143
3	533.333	800.000	666.667	4	.0571	5	.0714
4	800.000	1066.667	933.333	11	.1571	16	.2286
5	1066.667	1333.333	1200.000	17	.2429	33	.4714
6	1333.333	1600.000	1466.667	15	.2143	48	.6857
7	1600.000	1866.667	1733.333	12	.1714	60	.8571
8	1866.667	2133.333	2000.000	8	.1143	68	.9714
9	2133.333	2400.000	2266.667	2	.0286	70	1.0000
above	2400.000			0	.0000	70	1.0000

Mean = 1403.66 Standard Deviation = 402.385 Median = 1436.5

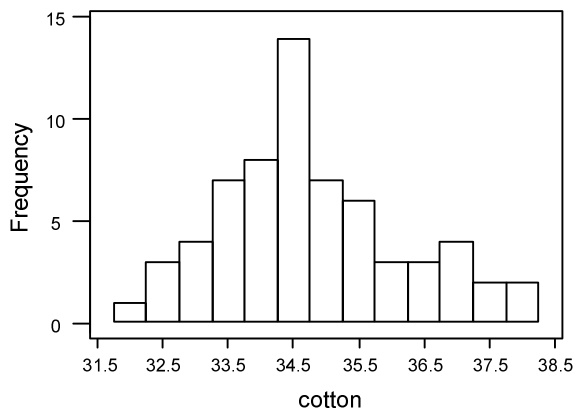


2-11.

Frequency Tabulation for Problem 2-11.Cotton

Class	Lower Limit	Upper Limit	Midpoint	Frequency	Relative Frequency	Cumulative Frequency	Cum. Rel. Frequency
at or below		31.000		0	.0000	0	.0000
1	31.000	32.000	31.500	0	.0000	0	.0000
2	32.000	33.000	32.500	6	.0938	6	.0938
3	33.000	34.000	33.500	12	.1875	18	.2813
4	34.000	35.000	34.500	22	.3438	40	.6250
5	35.000	36.000	35.500	12	.1875	52	.8125
6	36.000	37.000	36.500	7	.1094	59	.9219
7	37.000	38.000	37.500	5	.0781	64	1.0000
8	38.000	39.000	38.500	0	.0000	64	1.0000
above	39.000			0	.0000	64	1.0000

Mean = 34.7984   Standard Deviation = 1.36411   Median = 34.7

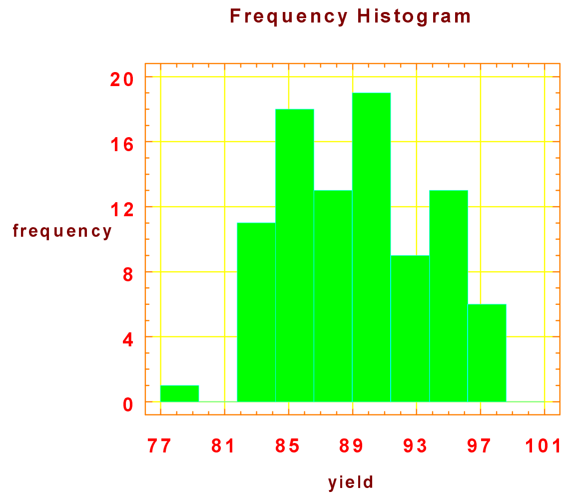


2-12.

Frequency Tabulation for Problem 2-12.Yield

Class	Lower Limit	Upper Limit	Midpoint	Frequency	Relative Frequency	Cumulative Frequency	Cum. Rel. Frequency
at or below		77.000		0	.0000	0	.0000
1	77.000	79.400	78.200	1	.0111	1	.0111
2	79.400	81.800	80.600	0	.0000	1	.0111
3	81.800	84.200	83.000	11	.1222	12	.1333
4	84.200	86.600	85.400	18	.2000	30	.3333
5	86.600	89.000	87.800	13	.1444	43	.4778
6	89.000	91.400	90.200	19	.2111	62	.6889
7	91.400	93.800	92.600	9	.1000	71	.7889
8	93.800	96.200	95.000	13	.1444	84	.9333
9	96.200	98.600	97.400	6	.0667	90	1.0000
10	98.600	101.000	99.800	0	.0000	90	1.0000
above	101.000			0	.0000	90	1.0000

Mean = 89.3756   Standard Deviation = 4.31591   Median = 89.25

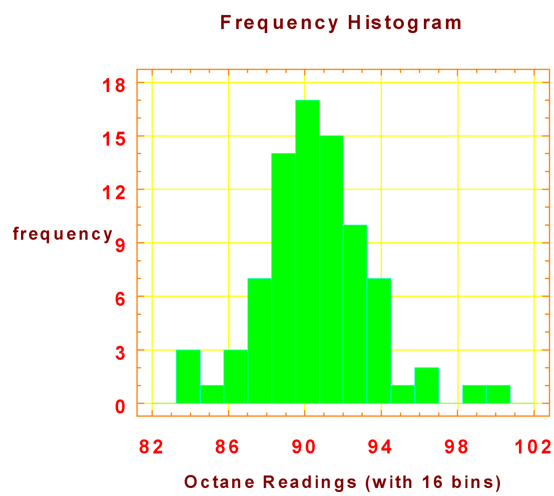


2-13.

Frequency Tabulation for Problem 2-13.Octane

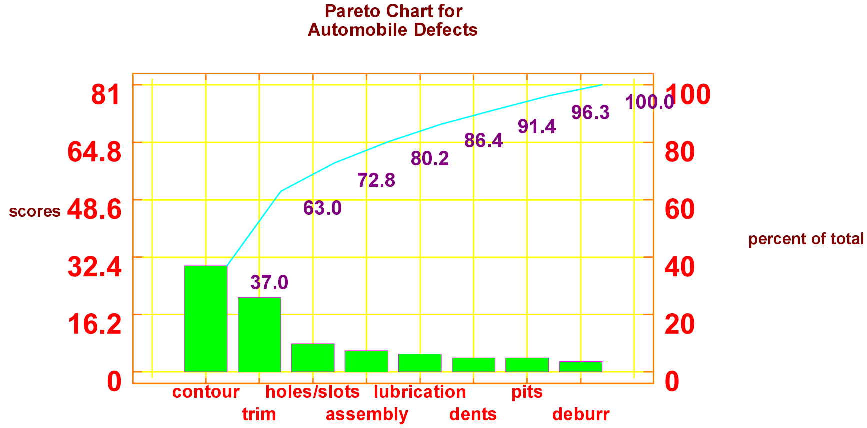
Class	Lower Limit	Upper Limit	Midpoint	Frequency	Relative Frequency	Cumulative Frequency	Cum. Rel. Frequency
at or below		82.000		0	.0000	0	.0000
1	82.000	83.250	82.625	0	.0000	0	.0000
2	83.250	84.500	83.875	3	.0366	3	.0366
3	84.500	85.750	85.125	1	.0122	4	.0488
4	85.750	87.000	86.375	3	.0366	7	.0854
5	87.000	88.250	87.625	7	.0854	14	.1707
6	88.250	89.500	88.875	14	.1707	28	.3415
7	89.500	90.750	90.125	17	.2073	45	.5488
8	90.750	92.000	91.375	15	.1829	60	.7317
9	92.000	93.250	92.625	10	.1220	70	.8537
10	93.250	94.500	93.875	7	.0854	77	.9390
11	94.500	95.750	95.125	1	.0122	78	.9512
12	95.750	97.000	96.375	2	.0244	80	.9756
13	97.000	98.250	97.625	0	.0000	80	.9756

Mean = 90.5256 Standard Deviation = 2.9052 Median = 90.4



The overall shapes appear to be the same.

2-14.



Roughly 63% of defects are described by parts out of contour and parts under trimmed.

<u>Section 2-4</u>

2-15.

Descriptive Statistics

```
Variable          N     Mean    Median   Tr Mean     StDev   SE Mean
Temperat         36    65.86     67.50     66.66     12.16      2.03
Variable        Min      Max        Q1        Q3
Temperat      31.00    84.00     58.50     75.00
```

a) Upper Quartile: 75
Lower Quartile: 58.5

b) Median: 67.5

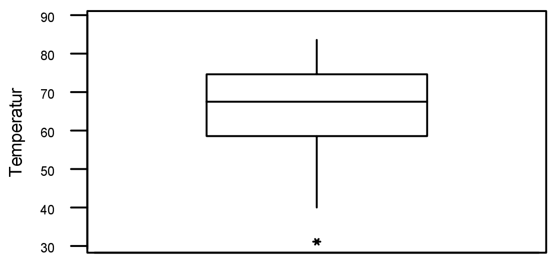
c)**Descriptive Statistics**

```
Variable          N     Mean    Median   Tr Mean     StDev   SE Mean
Temperat         35    66.86     68.00     67.35     10.74      1.82
Variable        Min      Max        Q1        Q3
Temperat      40.00    84.00     60.00     75.00
```

Upper Quartile: 75
Lower Quartile: 60
Median: 68

The lower quartile has increased while the upper quartile has remained unchanged. The median has increased slightly due to the removal of the data point. The smallest value appears quite different than the other temperature values.

d) Using the entire data set, the box plot is



The value of 31 appears to be one possible outlier.

2-16.

Descriptive Statistics

Variable	N	Mean	Median	Tr Mean	StDev	SE Mean
PMC	20	4.000	4.100	4.044	0.931	0.208

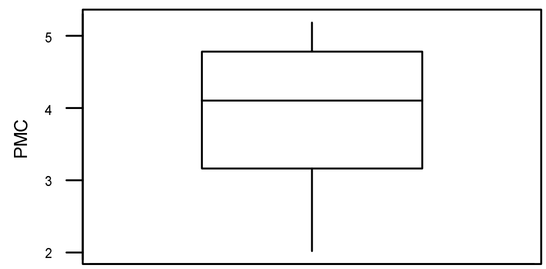
Variable	Min	Max	Q1	Q3
PMC	2.000	5.200	3.150	4.800

a) Sample Mean: 4

b) Sample Variance: 0.867

Sample Standard Deviation: 0.931

c)



2-17.

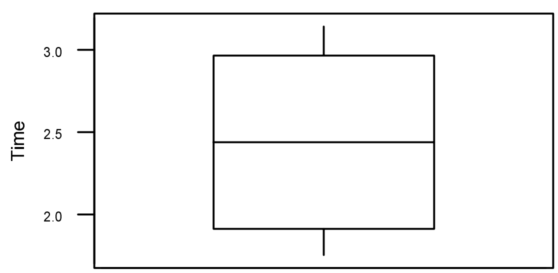
Descriptive Statistics

Variable	N	Mean	Median	Tr Mean	StDev	SE Mean
Time	8	2.415	2.440	2.415	0.534	0.189

Variable	Min	Max	Q1	Q3
Time	1.750	3.150	1.912	2.972

a) Sample Mean: 2.415 Sample Standard Deviation: 0.534

b)



2-18.

Descriptive Statistics

Variable	N	Mean	Median	Tr Mean	StDev	SE Mean
Temperat	9	952.44	953.00	952.44	3.09	1.03

Variable	Min	Max	Q1	Q3
Temperat	948.00	957.00	949.50	955.00

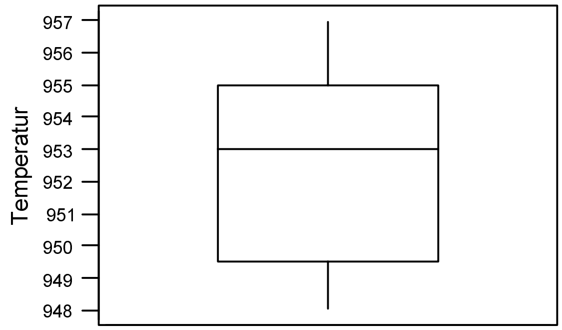
a) Sample Mean: 952.44

Sample Variance: 9.55

Sample Standard Deviation: 3.09

b) Median: 953; Any increase in the largest temperature measurement will not affect the median.

c)



2-19.

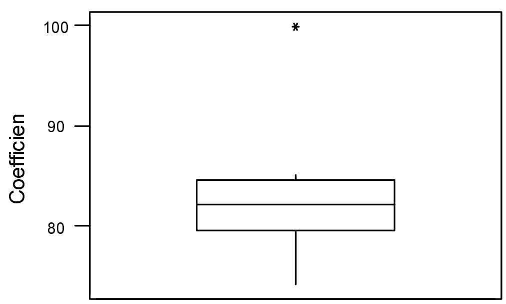
Descriptive Statistics

Variable	N	Mean	Median	Tr Mean	StDev	SE Mean
Coeffici	9	83.11	82.00	83.11	7.11	2.37

Variable	Min	Max	Q1	Q3
Coeffici	74.00	100.00	79.50	84.50

a) Upper Quartile: 84.5
Lower Quartile: 79.5

b)



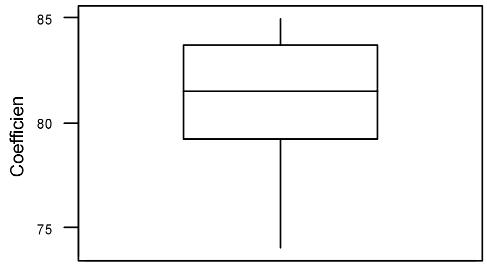
c)

Descriptive Statistics

Variable	N	Mean	Median	Tr Mean	StDev	SE Mean
Coeffici	8	81.00	81.50	81.00	3.46	1.22

Variable	Min	Max	Q1	Q3
Coeffici	74.00	85.00	79.25	83.75

Upper Quartile: 83.75
Lower Quartile: 79.25
By eliminating the largest observation, the mean has been decreased somewhat, but the variance and standard deviation have substantially decreased.



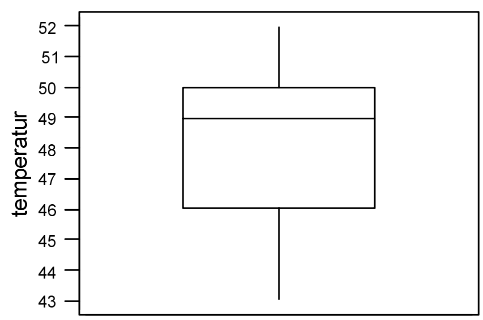
2-20.

Descriptive Statistics

Variable	N	Mean	Median	Tr Mean	StDev	SE Mean
temperat	24	48.125	49.000	48.182	2.692	0.549

Variable	Min	Max	Q1	Q3
temperat	43.000	52.000	46.000	50.000

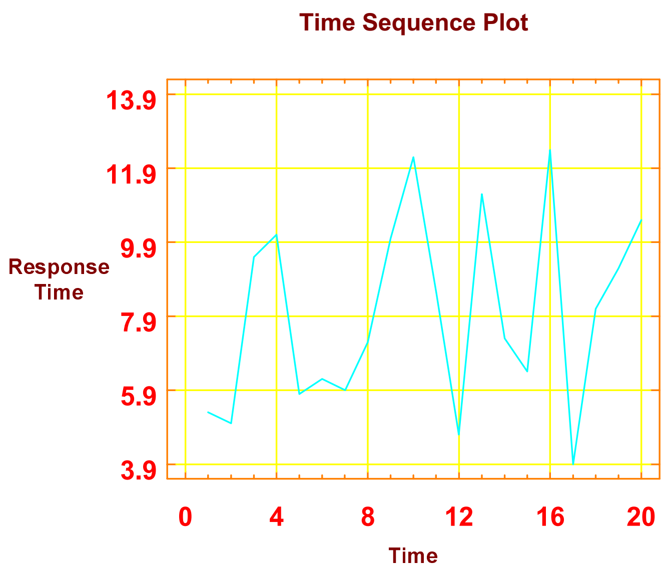
a) Sample Mean: 48.125
   Sample Median: 49
b) Sample Variance: 7.246
   Sample Standard Deviation: 2.692
c)



The data appear to be skewed.
d) 5th Percentile: 44   95th Percentile: 52

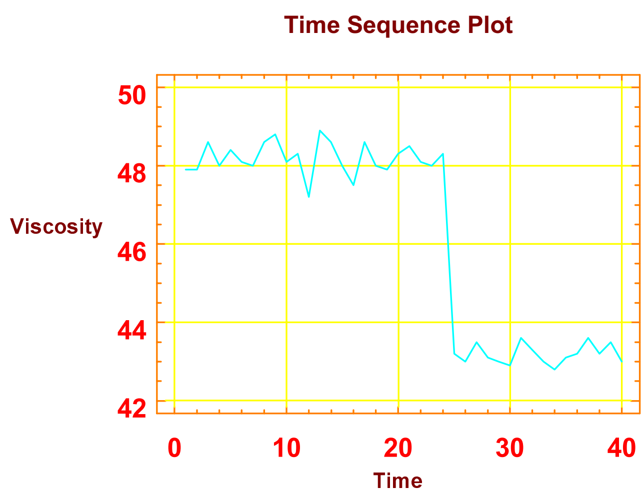
Section 2-5

2-21.



Computer response time appears random. No trends or patterns are obvious.

2-22. a)

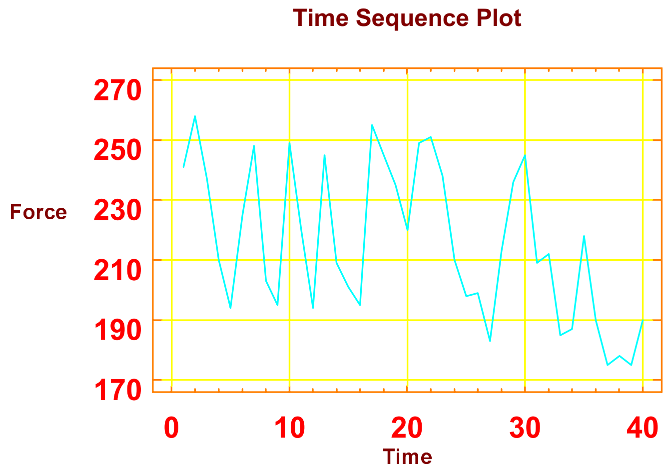


Stem-and-leaf display for Problem 2-22.Viscosity: unit = 0.1 1|2 represents 1.2

```
  2    42o|89
 12    43*|0000112223
 16    43o|5566
 16    44*|
 16    44o|
 16    45*|
 16    45o|
 16    46*|
 16    46o|
 17    47*|2
(4)    47o|5999
 19    48*|000001113334
  7    48o|5666689
```

b) The plots indicates that the process is not stable and not capable of meeting the specifications.

2-23. a)

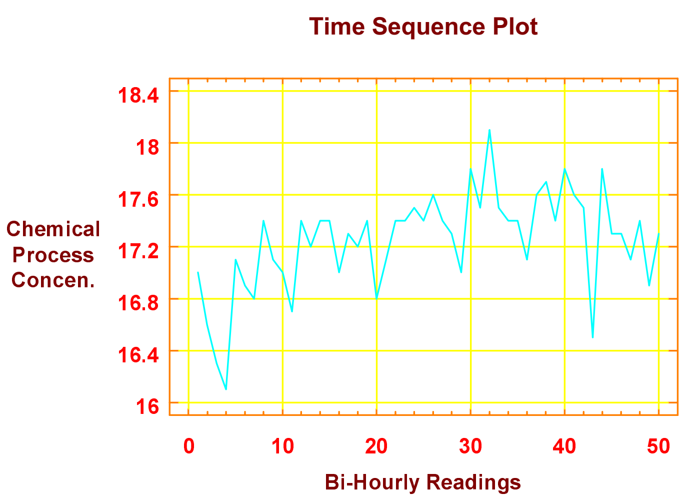


b) Stem-and-leaf display for Problem 2-23.Force: unit = 1 1|2 represents 12

```
  3    17|558
  6    18|357
 14    19|00445589
 18    20|1399
 (5)   21|00238
 17    22|005
 14    23|5678
 10    24|1555899
  3    25|158
```

In the time series plot there appears to be a downward trend beginning after time 30.

2-24.

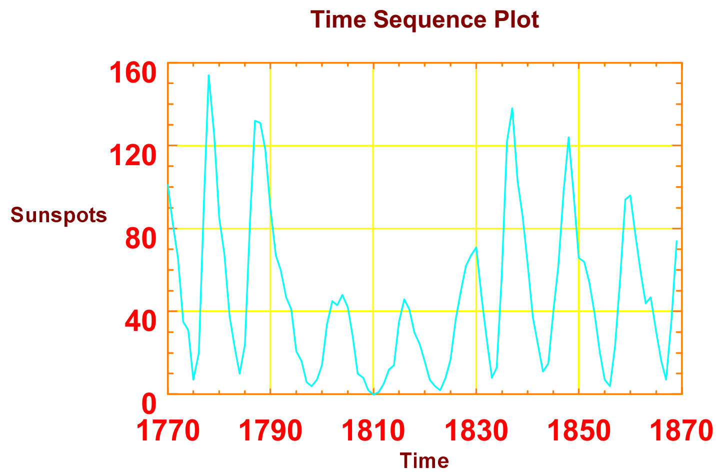


Stem-and-leaf display for Concentration: unit = 0.01   1|2 represents 0.12

```
        LO|1610,1630

   3   165|0
   4   166|0
   5   167|0
   7   168|00
   9   169|00
  13   170|0000
  18   171|00000
  20   172|00
  25   173|00000
  25   174|0000000000000
  12   175|0000
   8   176|000
   5   177|0
   4   178|000

        HI|1810
```

The data appear skewed.

2-25 a)

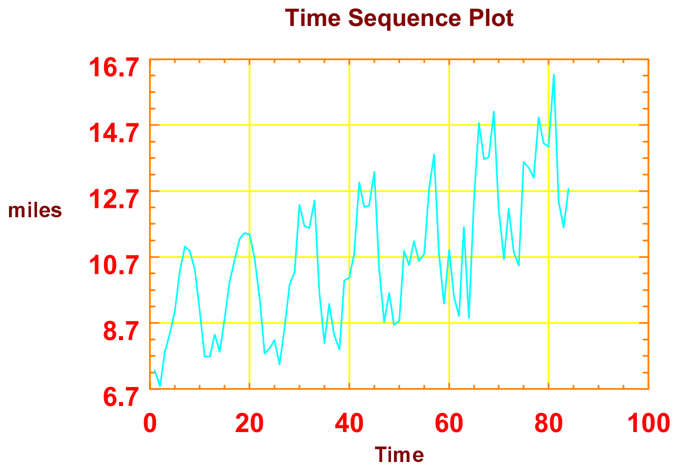


b) Stem-and-leaf display for Problem 2-25.Sunspots: unit = 1   1|2 represents 12

```
  17     0|01224445677777888
  29     1|001234456667
  39     2|0113344488
  50     3|00145567789
  50     4|011234567788
  38     5|04579
  33     6|0223466778
  23     7|147
  20     8|2356
  16     9|024668
  10    10|13
   8    11|8
   7    12|245
   4    13|128

        HI|154
```

The data appears to decrease between 1790 and 1835, the stem and leaf plot indicates skewed data.

2-26. a)



b)
Stem-and-leaf display for Problem 2-26.miles: unit = 0.1 1|2 represents 1.2

```
  1     6|7
 10     7|246678889
 22     8|013334677889
 33     9|01223466899
(18)   10|022334456667888889
 33    11|012345566
 24    12|11222345779
 13    13|1245678
  6    14|0179
  2    15|1
  1    16|2
```

There is an increasing trend in the data.

Supplemental Exercises

2-27. a) Sample Mean = 7.1838

The sample mean value is close enough to the target value to accept the solution as conforming. There is a slight difference due to inherent variability.

b) $s^2 = 0.000427$ $s = 0.02066$

A major source of variability would include measurement to measurement error.

A low variance is desirable since it may indicate consistency from measurement to measurement.

2-28. a) $\sum_{i=1}^{6} x_i^2 = 10{,}433$ $\left(\sum_{i=1}^{6} x_i\right)^2 = 62{,}001$ $n = 6$

$$s^2 = \frac{\sum_{i=1}^{6} x_i^2 - \frac{\left(\sum_{i=1}^{6} x_i\right)^2}{n}}{n-1} = \frac{10{,}433 - \frac{62{,}001}{6}}{6-1} = 19.9\Omega^2$$

$$s = \sqrt{19.9\Omega^2} = 4.46\Omega$$

b) $\overline{x} = \frac{246}{6} = 41.5$ n = 6

$$s^2 = \frac{\sum_{i=1}^{6}(x_i - \overline{x})^2}{n-1}$$

$$= \frac{99.5}{5} = 19.9\Omega^2$$

$$s = \sqrt{19.9\Omega^2} = 4.46\Omega$$

c) $s^2 = 19.9\Omega^2$ $s = 4.46\Omega$

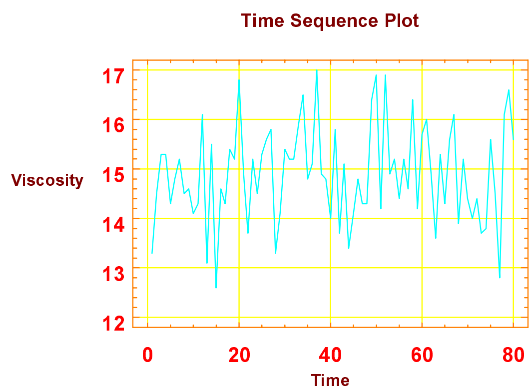
Shifting the data from the sample by a constant amount has no effect on the sample variance or standard deviation.

d) Yes, the rescaling is by a factor of 10. Therefore, s^2 and s would be rescaled by multiplying s^2 by 10^2 (resulting in $1990\Omega^2$) and s by 10 (44.6Ω).

2-29. a) Sample Range = 3.2
$s^2 = 0.8663$
s = 0.9308
b) Sample Range = 3.2
$s^2 = 0.8663$
s = 0.9308
These are the same as in part a). Any constant would produce the same results.

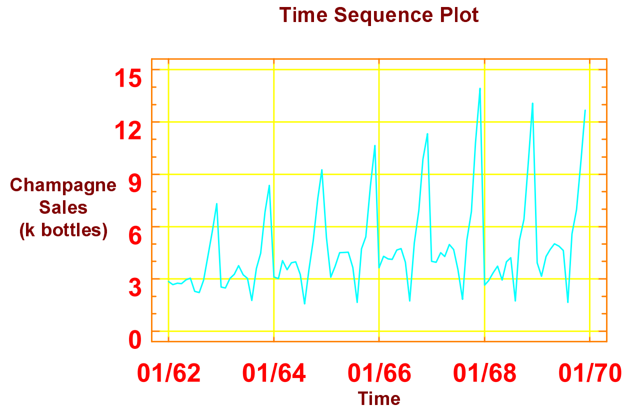
2-30. a) Sample 1 Range = 4
Sample 2 Range = 4
Yes, the two appear to exhibit the same variability
b) Sample 1 s = 1.604
Sample 2 s = 1.852
No, sample 2 has a larger standard deviation.
c) The sample range is a crude estimate of the sample variability as compared to the sample standard deviation since the standard deviation uses the information from every data point in the sample whereas the range uses the information contained in only two data points - the minimum and maximum.

2-31. a)



The data appears to vary between 12.5 and 17, with no obvious pattern.
b) The plot indicates that the two processes generate similar results. This is evident since the data appear to be centered around the same mean.
c) 1st 40 observations: Sample Mean = 14.875
Sample Variance = 0.89936
2nd 40 observations: Sample Mean = 14.9225
Sample Variance = 1.0464
The quantities indicate the processes do yield the same mean level. The variability also appears to be about the same, with the sample variance for the 2nd 40 observations being slightly larger than that for the 1st 40.

2-32. a)



There appears to be a cyclic variation in the data with the high value of the cycle generally increasing.

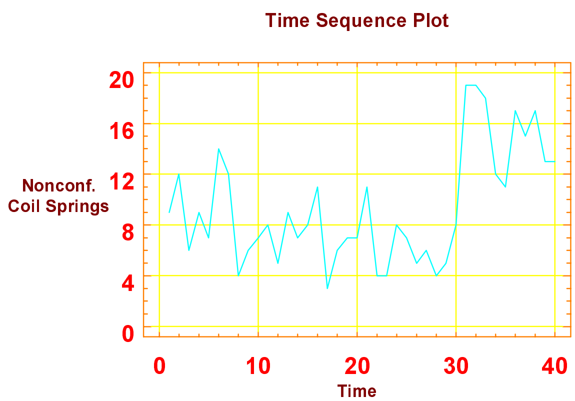
b) We might draw another cycle, with the peak similar to year 1969 at about 12.7.

2-33. a)

```
Stem-and-leaf display for Problem 2-35: unit = 1    1|2  represents 12

    1     0T|3
    8     0F|4444555
   18     0S|6666777777
  (7)     0o|8888999
   15     1*|111
   12     1T|22233
    7     1F|45
    5     1S|77
    3     1o|899
```

b) Sample Average = 9.325
Sample Standard Deviation = 4.4858

c)



The time series plot indicates there was an increase in the average number of nonconforming springs made during the 40 days. In particular, the increase occurs during the last 10 days.

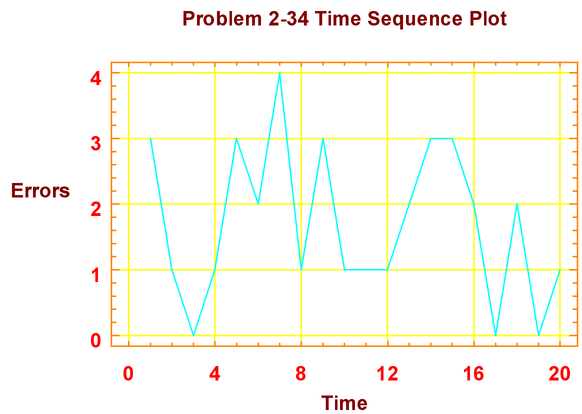
2-34. a) Stem-and-leaf display for Problem 2-36: unit = 0.1 1|2 represents 1.2

```
 3     0|000
10     1|0000000
10     2|0000
 6     3|00000
 1     4|0
```

b) Sample Average = 1.7

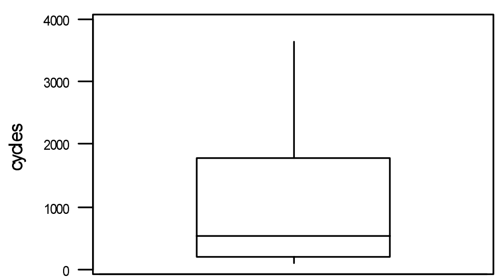
Sample Standard Deviation = 1.1743

c)

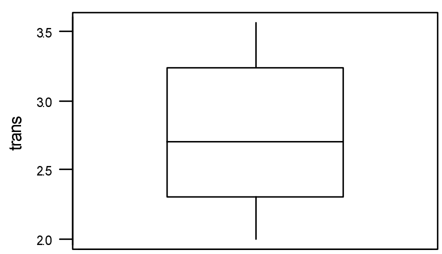


The time series plot indicates a slight decrease in the number of errors for strings 16 - 20.

2-35 a) Box plot



b) Box plot for transformed data



The log transformation made the resulting distribution much more symmetrical in both the middle 50% of the data and in the tails.

Mind-Expanding Exercises

2-36. Yes, in this case, since no upper bound on the last electronic component is available, use a measure of central location that is not dependent on this value. That measure is the median.

$$\text{Sample Median} = \frac{x_{(4)} + x_{(5)}}{2} = \frac{63+75}{2} = 69 \quad \text{hours}$$

2-37. a) $$\bar{x}_{n+1} = \frac{\sum_{i=1}^{n+1} x_i}{n+1} = \frac{\sum_{i=1}^{n} x_i + x_{n+1}}{n+1}$$

$$\bar{x}_{n+1} = \frac{n\bar{x}_n + x_{n+1}}{n+1}$$

$$\bar{x}_{n+1} = \frac{n}{n+1}\bar{x}_n + \frac{x_{n+1}}{n+1}$$

b) $$ns_{n+1}^2 = \sum_{i=1}^{n} x_i^2 + x_{n+1}^2 - \frac{\left(\sum_{i=1}^{n} x_i + x_{n+1}\right)^2}{n+1}$$

$$= \sum_{i=1}^{n} x_i^2 + x_{n+1}^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n+1} - \frac{2x_{n+1}\sum_{i=1}^{n} x_i}{n+1} - \frac{x_{n+1}^2}{n+1}$$

$$= \sum_{i=1}^{n} x_i^2 + \frac{n}{n+1}x_{n+1}^2 - \frac{n}{n+1}2x_{n+1}\bar{x}_n - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n+1}$$

$$= \sum_{i=1}^{n} x_i^2 - \frac{n}{n+1}x_{n+1}^2 - \frac{n}{n+1}2x_{n+1}\bar{x}_n + \frac{n}{n+1}\bar{x}_n - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}$$

$$= \left[\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum x_i\right)^2}{n+1}\right] + \frac{n}{n+1}\left[x_{n+1}^2 - 2x_{n+1}\bar{x}_n\right]$$

$$= \sum_{i=1}^{n} x_i^2 + \left[\frac{\left(\sum x_i\right)^2}{n} + \frac{\left(\sum x_i\right)^2}{n}\right] - \frac{\left(\sum x_i\right)^2}{n+1} + \frac{n}{n+1}\left[x_{n+1}^2 - 2x_n\bar{x}_n\right]$$

$$= \sum_{i=1}^{n} x_i^2 - \frac{\left(\sum x_i\right)^2}{n} + \frac{(n+1)\left(\sum x_i\right)^2 - n\left(\sum x_i\right)^2}{n(n+1)} + \frac{n}{n+1}\left[x_{n+1}^2 - 2x_n\bar{x}_n\right]$$

$$= (n-1)s_n^2 + \frac{\left(\sum x_i\right)^2}{n(n+1)} + \frac{n}{n+1}\left[x_{n+1}^2 - 2x_n\bar{x}_n\right]$$

$$= (n-1)s_n^2 + \frac{n\bar{x}^2}{n+1} + \frac{n}{n+1}\left[x_{n+1}^2 - 2x_n\bar{x}_n\right]$$

$$= (n-1)s_n^2 + \frac{n}{n+1}\left(x_{n+1} - 2x_n\bar{x}_n + \bar{x}_n^2\right)$$

$$= (n-1)s_n^2 + \frac{n}{n+1}\left(x_{n+1} - \bar{x}_n\right)^2$$

c) $\overline{x}_n = 41.5$ $\quad x_{n+1} = 46$

$sn^2 = 19.9$ $\quad n = 6$

$$\overline{x}_{n+1} = \frac{6(41.5)+46}{6+1}$$

$$= 42.14$$

$$s_{n+1} = \frac{\sqrt{(6-1)19.9 + \frac{6}{6+1}46 - 41.5}}{6}$$

$$= 4.4132$$

2-38. The trimmed mean is pulled toward the median by eliminating outliers.
a) 10% Trimmed Mean = 89.29
b) 20% Trimmed Mean = 89.19
Difference is very small
c) No, the differences are very small, due to a very large data set with no significant outliers.

2-39. If nT/100 is not an integer, calculate the two surrounding integer values and interpolate between the two. For example, if nT/100 = 2/3, one could calculate the mean after trimming 2 and 3 observations from each end and then interpolate between these two means.

CHAPTER 3

Section 3-1

3-1. Let "a", "b" denote a part above, below the specification
$S = \{aaa, aab, aba, abb, baa, bab, bba, bbb\}$

3-2. Let "e" denote a bit in error
Let "o" denote a bit not in error ("o" denotes okay)

$$S = \begin{Bmatrix} eeee, eoee, oeee, ooee, \\ eeeo, eoeo, oeeo, ooeo, \\ eeoe, eooe, eoeo, oooe, \\ eeoo, eooo, oeoo, oooe \end{Bmatrix}$$

3-3. Let "a" denote an acceptable power supply
Let "f" ,"m","c" denote a supply with a functional, minor, or cosmetic error, respectively.
$S = \{a, f, m, c\}$

3-4. $S = \{0,1,2,...\}$ = set of nonnegative integers

3-5. If only the number of tracks with errors is of interest, then $S = \{0,1,2,...,24\}$
If the particular tracks with errors are of interest, then S is a set of vectors with 24 components. Each vector indicates which tracks contain errors. That is, let "e" denote a track with one or more errors and let "o" denote a track without errors. Then S is the set of all vectors that contain 24 letters. Each letter is either "e", or "o" as in Exercise 3-2. The sample space contains 2^{24} vectors.

3-6. A vector with three components can describe the three digits of the ammeter. Each digit can be 0,1,2,...,9. Then S is a sample space of 1000 possible three digit integers, $S = \{000, 001, ..., 999\}$

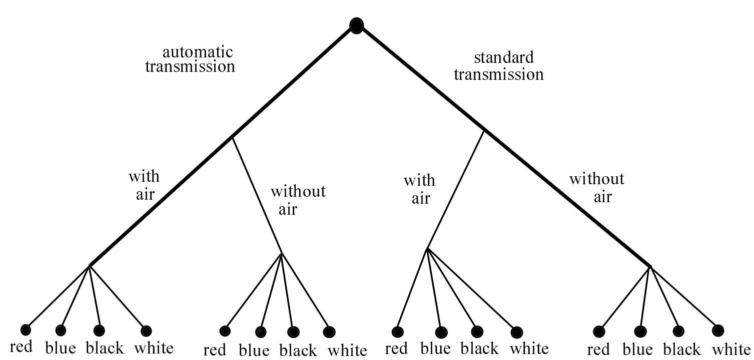
3-7. Similar to Exercise 3-6. Now, S is the sample space of 100 possible two digit integers.

3-8. Let an ordered pair of numbers, such as 43 denote the response on the first and second question. Then, S consists of the 25 ordered pairs $\{11, 12, ..., 55\}$

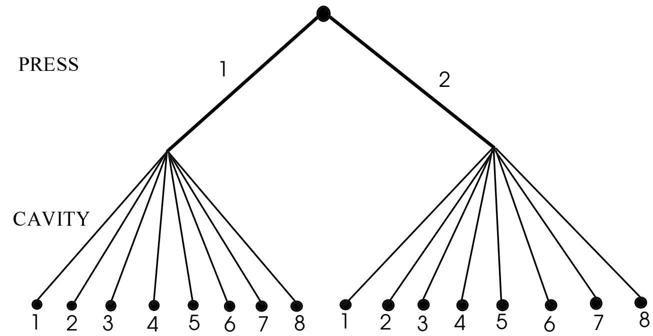
3-9. The number of small, medium, and large voids can be recorded as a vector with three components. Each component can equal any nonnegative integer. If P is the set of nonnegative integers, then S = PxPxP.

3-10. Let "a" denote an acceptable part and let "u" denote an unacceptable part.
Then, $S = \{aaa, aau, aua, auu, uaa, uau, uua, uuu\}$
Alternatively, if only the number of acceptable parts is of interest, then $S = \{0,1,2,3\}$.

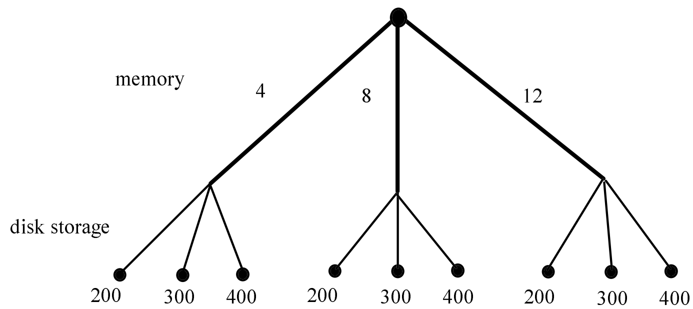
3-11.



3-12.



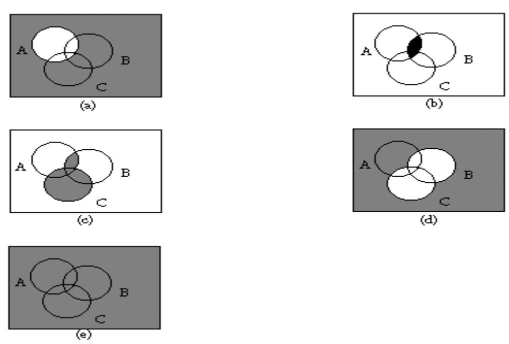
3-13.

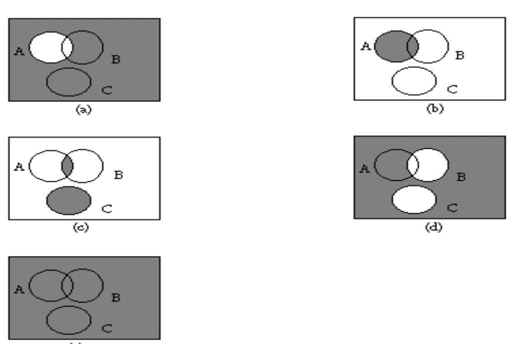


3-14. Let "b" denote a busy signal and let "c" denote a connected call $S = \{c, bc, bbc, bbbc, \ldots\}$

3-15. $S = \{s, fs, ffs, fffS, fffFS, fffFFS, fffFFFA\}$

3-16. & 3-17.

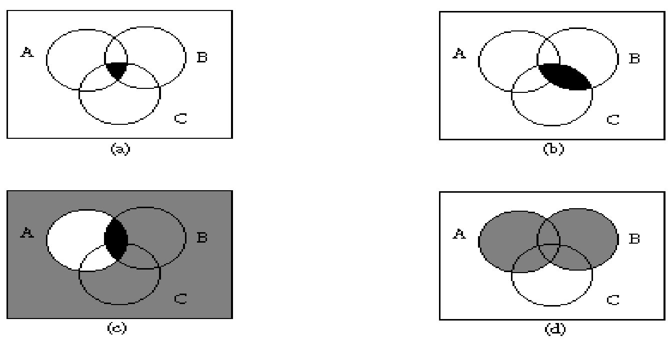




3-18. a) S = nonnegative integers from 0 to the largest integer that can be displayed by the scale.
Let X represent weight.
A is the event that $X > 11$
B is the event that $X \le 15$
C is the event that $8 \le X < 12$

b) nonnegative integers from 0 to the largest integer that can be displayed by the scale.
c) $11 < X \le 15$
d) $X \le 11$
e) nonnegative integers from 0 to the largest integer that can be displayed by the scale.
f) $A \cup C$ would contain the values of X such that: $X \ge 8$
Thus $(A \cup C)'$ would contain the values of X such that: $X < 8$.
g) $11 \le X \le 12$
h) B' would contain the values of X such that $X > 15$. Therefore, $B' \cap C$ would be the empty set. They have no outcomes in common.
i) $B \cap C$ is the event $8 \le X < 12$. Therefore, $A \cup (B \cap C)$ is the event $X \ge 8$

3-19.



3-20. Let "d" denoted a distorted bit and let "o" denote a bit that is not distorted.

a) $S = \left\{ \begin{matrix} dddd, dodd, oddd, oodd, \\ dddo, dodo, oddo, oodo, \\ ddod, dood, odod, oood, \\ ddoo, dooo, odoo, oooo \end{matrix} \right\}$

b) No, for example $A_1 \cap A_2 = \{dddd, dddo, ddod, ddoo\}$

c)

$$A_1 = \left\{ \begin{matrix} dddd, dodd, \\ dddo, dodo \\ ddod, dood \\ ddoo, dooo \end{matrix} \right\}$$

$$A_1' = \left\{ \begin{matrix} oddd, oodd, \\ oddo, oodo, \\ odod, oood, \\ odoo, oooo \end{matrix} \right\}$$

$$A_1 \cap A_2 \cap A_3 \cap A_4 = dddd$$

$$(A_1 \cap A_2) \cup (A_3 \cap A_4) = \left\{ \begin{matrix} dddd, dodd, \\ dddo, oddd, \\ ddod, oodd, \\ ddoo \end{matrix} \right\}$$

3-21. Let "d" denote a defective calculator and let "o" denote an acceptable calculator

a) $S = \{ddd, odd, ddo, odo, dod, ood, doo, ooo\}$

b) $A = \{ddd, ddo, dod, doo\}$

c) $B = \{ddd, ddo, odd, odo\}$

d) $A \cap B = \{ddd, ddo\}$

e) $B \cup C = \{ddd, ddo, odd, odo, dod, ood\}$

3-22. number of samples in $A \cap B = 80$
number of samples in A' = 14
number of samples in $A \cup B = 95$

3-23.

	number of samples
$A' \cap B$	10
B'	15
$A \cup B$	92

3-24.

	number of samples
$A' \cap B$	17
B'	5
$A \cup B$	27

3-25. a) $A' = \{x \mid x \geq 72.5\}$
b) $B' = \{x \mid x \leq 52.5\}$
c) $A \cap B = \{x \mid 52.5 < x < 72.5\}$
d) $A \cup B = \{x \mid x > 0\}$

Section 3-2

3-27. All outcomes are equally likely
a) $P(A) = 2/5$
b) $P(B) = 3/5$
c) $P(A') = 3/5$
d) $P(A \cup B) = 1$
e) $P(A \cap B) = P(\varnothing) = 0$

3-28. a) $P(A) = 0.2$
b) $P(B) = 0.8$
c) $P(A') = 0.8$
d) $P(A \cup B) = 1$
e) $P(A \cap B) = P(\varnothing) = 0$

3-29. a) $S = \{1, 2, 3, 4, 5, 6\}$
b) 1/6
c) 2/6
d) 5/6

3-30. a) $S = \{1,2,3,4,5,6,7,8\}$
b) 2/8
c) 6/8

3-31. $\frac{x}{20} = 0.3, x = 6$

3-32. a) 0.5 + 0.2 = 0.7
b) 0.3 + 0.5 = 0.8

3-33. a) 1/10
b) 5/10

3-34. a) 0.25
b) 0.75

3-35. a) $P(A) = 86/100$
b) $P(B) = 89/100$
c) $P(A') = 14/100$
d) $P(A \cap B) = 80/100$
e) $P(A \cup B) = (80+9+6)/100 = 0.95$
f) $P(A' \cup B) = (80+9+5)/100 = 0.94$

3-36. a) $P(A) = 82/100$
b) $P(B) = 85/100$
c) $P(A') = 18/100$
d) $P(A \cap B) = 75/100$
e) $P(A \cup B) = (75+7+10)/100 = 0.92$
f) $P(A' \cup B) = (75+10+8)/100 = 0.93$

3-37. a) $P(A) = 20/40$
b) $P(B) = 35/40$
c) $P(A') = 20/40$
d) $P(A \cap B) = 18/40$
e) $P(A \cup B) = (18+17+2)/40 = 37/40$
f) $P(A' \cup B) = (17+18+3)/40 = 38/40$

3-38. a) Because E and E' are mutually exclusive events and $E \cup E' = S$

$1 = P(S) = P(E \cup E') = P(E) + P(E')$. Therefore, $P(E') = 1 - P(E)$

b) Because S and $\varnothing$ are mutually exclusive events with $S = S \cup \varnothing$

$P(S) = P(S) + P(\varnothing)$. Therefore, $P(\varnothing) = 0$

c) Now, $B = A \cup (A' \cap B)$ and the events A and $A' \cap B$ are mutually exclusive. Therefore,

$P(B) = P(A) + P(A' \cap B)$. Because $P(A' \cap B) \geq 0$, $P(B) \geq P(A)$.

Section 3-3

3-39. a) $P(A') = 1 - P(A) = 0.7$

b) $P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.3+0.2 - 0.1 = 0.4$

c) $P(A' \cap B) + P(A \cap B) = P(B)$. Therefore, $P(A' \cap B) = 0.2 - 0.1 = 0.1$

d) $P(A) = P(A \cap B) + P(A \cap B')$. Therefore, $P(A \cap B') = 0.3 - 0.1 = 0.2$

e) $P((A \cup B)') = 1 - P(A \cup B) = 1 - 0.4 = 0.6$ from part b.

f) $P(A' \cup B) = P(A') + P(B) - P(A' \cap B) = 0.7 + 0.2 - 0.1 = 0.8$ from part c.

3-40. a) $P(A \cup B \cup C) = P(A) + P(B) + P(C)$, because the events are mutually exclusive. Therefore,

$P(A \cup B \cup C) = 0.2+0.3+0.4 = 0.9$

b) $P(A \cap B \cap C) = 0$, because $A \cap B \cap C = \varnothing$

c) $P(A \cap B) = 0$, because $A \cap B = \varnothing$

d) $P((A \cup B) \cap C) = 0$, because $(A \cup B) \cap C = (A \cap C) \cup (B \cap C) = \varnothing$

3-41. If A,B,C are mutually exclusive, then $P(A \cup B \cup C) = P(A) + P(B) + P(C) = 0.3 + 0.4 + 0.5 =$ 1.2, which greater than 1. Therefore, P(A), P(B),and P(C) cannot equal the given values.

3-42. Let A denote the event that a sample has high shock resistance and let B denote the event that a sample has high scratch resistance.

a) $P(A \cap B) = 80/100$

b) $P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{86}{100} + \frac{89}{100} - \frac{80}{100} = \frac{95}{100}$

c) Because $A \cap B$ does not equal $\varnothing$, A and B are not mutually exclusive.

3-43. a) 350/370

b) $\frac{345+5+12}{370} = \frac{362}{370}$

c) $\frac{345+5+8}{370} = \frac{358}{370}$

d) 345/370

3-44. Total number of shafts = 370

a) $\frac{201+149+4+8+3}{370} = \frac{364}{370}$

b) $\frac{201+149+2+6+8}{370} = \frac{366}{370}$

c) The number of shafts from Tool 2 = 163. The number of shafts from Tool 1 that conform to surface finish and roundness requirements = 200. Therefore, the answer is $\frac{163+200}{370} = \frac{363}{370}$

d) The number of shafts from Tool 2 = 163. The number of shafts from Tool 1 that conform to surface finish requirements = 201. Therefore, the answer is $\frac{163+201}{370} = \frac{364}{370}$

Section 3-4

3-45. a) P(A) = 86/100 b) P(B) = 89/100

c) $P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{80/100}{89/100} = \frac{80}{89}$

d) $P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{80/100}{86/100} = \frac{80}{86}$

3-46. a) P(A) = 82/100 b) P(B) = 85/100

c) $P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{75/100}{85/100} = \frac{75}{85}$

d) $P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{75/100}{82/100} = \frac{75}{82}$

e) $P(B|A)$ = 75/82

f) $P(A|B') = \frac{P(A \cap B')}{P(B')} = \frac{7/100}{15/100} = \frac{7}{15}$

3-47. a) 345/357 b) 5/13

3-48. a) 12/100 b) 12/28 c) 34/122

3-49. Need data from example
a) 0.15
b) 11/72
c) 0.72
d) 11/15
e) 0.11
f) 0.76

3-50. a) 20/100
b) 19/99
c) If the chips were replaced, the probability in part b. would be 20/100

3-51. a) P(A) = 15/40

b) $P(B|A)$ = 14/39

c) $P(A \cap B)$ = P(A) P(B/A) = (15/40) (14/39) = 0.135

d) $P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{15}{40} + \frac{14}{39} - \left(\frac{15}{40}\right)\left(\frac{14}{39}\right) = 0.599$

e) $P(A \cap B \cap C) = \left(\frac{15}{40}\right)\left(\frac{14}{39}\right)\left(\frac{13}{38}\right) = 0.046$

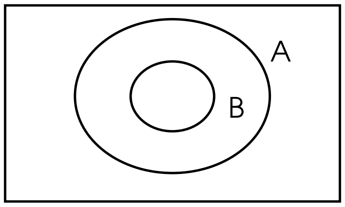
f) $P(A \cap B \cap C') = \left(\frac{15}{40}\right)\left(\frac{14}{39}\right)\left(\frac{25}{38}\right) = 0.089$

3-52. a) 4/499 = 0.0080
b) (5/500) (4/499) = 0.00008
c) (495/500) (494/499) = 0.980

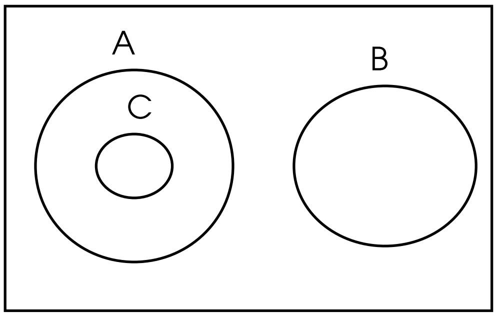
3-53. a) 3/498 = 0.0060
b) 4/498 = 0.0080

c) $\left(\frac{5}{500}\right)\left(\frac{4}{499}\right)\left(\frac{3}{498}\right) = 4.82x10^{-7}$

3-54. No, if $B \subset A$, then P(A/B) = $\frac{P(A \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1$



3-55.



Section 3-5

3-56. a) $P(A \cap B) = P(A|B)P(B) = (0.4)(0.5) = 0.20$

b) $P(A' \cap B) = P(A'|B)P(B) = (0.6)(0.5) = 0.30$

3-57.

$$\begin{aligned} P(A) &= P(A \cap B) + P(A \cap B') \\ &= P(A|B)P(B) + P(A|B')P(B') \\ &= (0.2)(0.8) + (0.3)(0.2) \\ &= 0.16 + 0.06 = 0.22 \end{aligned}$$

3-58. Let F denote the event that a connector fails.
Let W denote the event that a connector is wet.

$$\begin{aligned} P(F) &= P(F|W)P(W) + P(F|W')P(W') \\ &= (0.05)(0.10) + (0.01)(0.90) = 0.014 \end{aligned}$$

3-59. Let F denote the event that a roll contains a flaw.
Let C denote the event that a roll is cotton.

$$\begin{aligned} P(F) &= P(F|C)P(C) + P(F|C')P(C') \\ &= (0.02)(0.70) + (0.03)(0.30) = 0.023 \end{aligned}$$

3-60. a) P(A) = 0.03
b) P(A') = 0.97
c) P(B|A) = 0.40
d) P(B|A') = 0.05
e) $P(A \cap B) = P(B|A)P(A) = (0.40)(0.03) = 0.012$
f) $P(A \cap B') = P(B'|A)P(A) = (0.60)(0.03) = 0.018$
g) $P(B) = P(B|A)P(A) + P(B|A')P(A') = (0.40)(0.03) + (0.05)(0.97) = 0.0605$

3-61. Let R denote the event that a product exhibits surface roughness. Let N,A, and W denote the events that the blades are new, average, and worn, respectively. Then,

P(R)= P(R|N)P(N) + P(R|A)P(A) + P(R|W)P(W)
= (0.01)(0.25) + (0.03) (0.60) + (0.05)(0.15)
= 0.028

3-62. Let B denote the event that a glass breaks.
Let L denote the event that large packaging is used.
P(B)= P(B|L)P(L) + P(B|L')P(L')
= 0.01(0.60) + 0.02(0.40) = 0.014

3-63. Let A denote a event that the first part selected has excessive shrinkage.
Let B denote the event that the second part selected has excessive shrinkage.

a) P(B)= P($B|A$)P(A) + P($B|A'$)P(A')
= (4/24)(5/25) + (5/24)(20/25) = 0.20

b) Let C denote the event that the second part selected has excessive shrinkage.

c)

$$P(C) = P(C|A \cap B)P(A \cap B) + P(C|A \cap B')P(A \cap B')$$

$$+P(C|A' \cap B)P(A' \cap B) + P(C|A' \cap B')P(A' \cap B')$$

$$= \frac{3}{23}\left(\frac{2}{24}\right)\left(\frac{5}{25}\right) + \frac{4}{23}\left(\frac{20}{24}\right)\left(\frac{5}{25}\right) + \frac{4}{23}\left(\frac{5}{24}\right)\left(\frac{20}{25}\right) + \frac{5}{23}\left(\frac{19}{24}\right)\left(\frac{20}{25}\right)$$

$$= 0.20$$

3-64. Let A and B denote the events that the first and second chips selected are defective, respectively.

a) P(B) = P(B|A)P(A) + P(B|A')P(A') = (19/99)(20/100) + (20/99)(80/100) = 0.2

b) Let A and B denote the events that the first and second containers selected are defective, respectively.
P(B) = P(B|A)P(A) + P(B|A')P(A') = (4/499)(5/500) + (5/499)(495/500) = 0.01

<u>Section 3-6</u>

3-65. Because P($A|B$) $\neq$ P(A), the events are not independent.

3-66. P(A') = 1 - P(A) = 0.7 and P($A'|B$) = 1 - P($A|B$) = 0.7
Therefore, A' and B are independent events.

3-67. P($A \cap B$) = 80/100, P(A) = 86/100, P(B) = 89/100.
Then, P($A \cap B$) $\neq$ P(A)P(B), so A and B are <u>not</u> independent.

3-68. P($A \cap B$) = 75/100, P(A) = 82/100, P(B) = 85/100.
Then, P($A \cap B$) $\neq$ P(A)P(B), so A and B are <u>not</u> independent.

3-69. P($A \cap B$)= 80/126, P(A) = 20/126, P(B) = 65/126, Then P($A \cap B$)$\neq$ P(A)P(B), therefore, A and B are independent.

3-70. If A and B are mutually exclusive, then P($A \cap B$) = 0 and P(A)P(B) = 0.04.
Therefore, A and B are <u>not</u> independent.

3-71. a) P($A \cap B$) = 80/126, P(A) = 84/126, P(B) = 120/126. Then, P($A \cap B$) = P(A)P(B), so A and B are independent.

b) P($B|A'$) = 40/42 and P(B) = 120/126, so A' and B are independent. In general, if A and B are independent, then P($A|B$) = P(A). Also, P($A'|B$) = 1 - P($A|B$) and P(A') = 1 - P(A). Therefore, P($A'|B$) = P(A') and A' and B are independent.

3-72. It is useful to work one of these exercises with care to illustrate the laws of probability. Let H_i denote the event that the ith sample contains high levels of contamination.

a) $P(H_1' \cap H_2' \cap H_3' \cap H_4' \cap H_5') = P(H_1')P(H_2')P(H_3')P(H_4')P(H_5')$

by independence. Also, $P(H_i') = 0.9$. Therefore, the answer is $0.9^5 = 0.59$

b) $A_1 = (H_1 \cap H_2' \cap H_3' \cap H_4' \cap H_5')$

$A_2 = (H_1' \cap H_2 \cap H_3' \cap H_4' \cap H_5')$

$A_3 = (H_1' \cap H_2' \cap H_3 \cap H_4' \cap H_5')$

$A_4 = (H_1' \cap H_2' \cap H_3' \cap H_4 \cap H_5')$

$A_5 = (H_1' \cap H_2' \cap H_3' \cap H_4' \cap H_5)$

The requested probability is the probability of the union $A_1 \cup A_2 \cup A_3 \cup A_4 \cup A_5$ and these events are mutually exclusive. Also, by independence $P(A_i) = 0.9^4(0.1) = 0.0656$. Therefore, the answer is 5(0.0656) = 0.328.

c) Let B denote the event that no sample contains high levels of contamination. The requested probability is P(B') = 1 - P(B). From part (a), P(B') = 1 - 0.59 = 0.41.

3-73. Let A_i denote the event that the ith bit is a one.

a) By independence $P(A_1 \cap A_2 \cap \ldots \cap A_{10}) = P(A_1)P(A_2)\ldots P(A_{10}) = (\frac{1}{2})^{10} = 0.000976$

b) By independence, $P(A_1' \cap A_2' \cap \ldots \cap A_{10}') = P(A_1')P(A_2')\ldots P(A_{10}^c) = (\frac{1}{2})^{10} = 0.000976$

c) The probability of the following sequence is

$P(A_1' \cap A_2' \cap A_3' \cap A_4' \cap A_5' \cap A_6 \cap A_7 \cap A_8 \cap A_9 \cap A_{10}) = (\frac{1}{2})^{10}$, by independence. The number of sequences consisting of five "1"'s, and five "0"'s is $\binom{10}{5} = \frac{10!}{5!5!} = 252$. The answer is $252\left(\frac{1}{2}\right)^{10} = 0.246$

3-74. Let A denote the event that a sample is produced in cavity one of the mold.

a) By independence, $P(A_1 \cap A_2 \cap A_3 \cap A_4 \cap A_5) = (\frac{1}{8})^5 = 0.00003$

b) Let B_i be the event that all five samples are produced in cavity i. Because the B's are mutually exclusive, $P(B_1 \cup B_2 \cup \ldots \cup B_8) = P(B_1) + P(B_2) + \ldots + P(B_8)$

From part a., $P(B_i) = (\frac{1}{8})^5$. Therefore, the answer is $8(\frac{1}{8})^5 = 0.00024$

c) By independence, $P(A_1 \cap A_2 \cap A_3 \cap A_4 \cap A_5') = (\frac{1}{8})^4(\frac{7}{8})$. The number of sequences in which four out of five samples are from cavity one is 5. Therefore, the answer is $5(\frac{1}{8})^4(\frac{7}{8}) = 0.00107$.

3-75. Let A denote the upper devices function. Let B denote the lower devices function.
P(A) = (0.9)(0.9)(0.8) = 0.648
P(B) = (0.95)(0.95)(0.9) = 0.81225
P(A∩B) = (0.648)(0.81225) = 0.52634
Therefore, the probability that the circuit operates = P(A∪B) = P(A) +P(B) − P(A∩B) = 0.93391

3-76. Let A denote the first device functions, Let B denote the second device functions. Let C denote the third device functions.
P(A) = 0.9 + 0.9 − (0.9)(0.9) = 0.99
P(B) = 0.9 + 0.95 − (0.9)(0.95) = 0.995
P(C) = 0.8 + 0.9 − (0.8)(0.9) = 0.98
Therefore, the probability that the circuit operates = P(A∩B∩C) = P(A)P(B)P(C)=0.9653

3-77. Let A_i denote the event that the ith readback is successful. By independence,

$$P(A_1' \cap A_2' \cap A_3') = P(A_1')P(A_2')P(A_3') = (0.02)^3 = 0.000008.$$

3-78. a) $P(B|A) = 4/499$ and

$$\begin{aligned} P(B) &= P(B|A)P(A) + P(B|A')P(A') \\ &= (4/99)(5/500) + (5/499)(495/500) \\ &= 5/500 \end{aligned}$$

Therefore, A and B are not independent.

b) A and B are independent.

Section 3-7

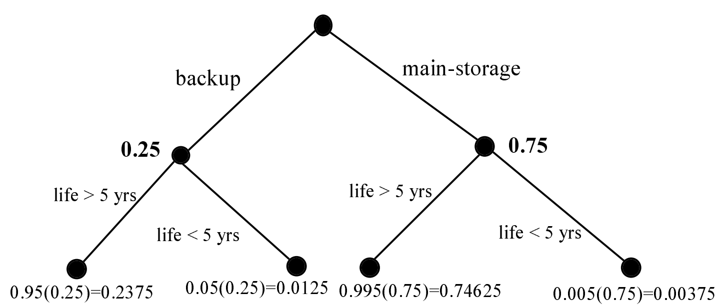
3-79. Because, $P(A|B)P(B) = P(A \cap B) = P(B|A)P(A)$,

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)} = \frac{0.8(0.2)}{0.5} = 0.08$$

3-80. Let F denote a fraudulent user and let T denote a user that originates calls from two or more metropolitan areas in a day. Then,

$$\begin{aligned} P(F|T) &= \frac{P(T|F)P(F)}{P(T|F)P(F) + P(T|F')P(F')} \\ &= \frac{0.30(0.0001)}{0.30(0.0001) + 0.01(.9999)} \\ &= 0.003 \end{aligned}$$

3-81.



a) $P(B) = 0.25$

b) $P(A|B) = 0.95$

c) $P(A|B') = 0.995$

d) $P(A \cap B) = P(A|B)P(B) = 0.95(0.25) = 0.2375$

e) $P(A \cap B') = P(A|B')P(B') = 0.995(0.75) = 0.74625$

f) $P(A) = P(A \cap B) + P(A \cap B') = 0.95(0.25) + 0.995(0.75) = 0.98375$

g) $0.95(0.25) + 0.995(0.75) = 0.98375$.

h)

$$\begin{aligned} P(B|A') &= \frac{P(A'|B)P(B)}{P(A'|B)P(B) + P(A'|B')P(B')} \\ &= \frac{0.05(0.25)}{0.05(0.25) + 0.005(0.75)} \\ &= 0.769 \end{aligned}$$

3-82. Let G denote a product that received a good review. Let H, M, and P denote products that were high, moderate, and poor performers, respectively.

a)

$$P(G) = P(G|H)P(H) + P(G|M)P(M) + P(G|P)P(P)$$
$$= 0.95(0.40) + 0.60(0.35) + 0.10(0.25)$$
$$= 0.615$$

b) Using the result from part a.,

$$P(H|G) = \frac{P(G|H)P(H)}{P(G)} = \frac{0.95(0.40)}{0.615} = 0.618$$

c) $P(H|G') = \frac{P(G'|H)P(H)}{P(G')} = \frac{0.05(0.40)}{1-0.615} = 0.052$

For problems 3-83 through 3-91: Continuous: 3-83, 3-85, 3-86, 3-88, 3-90, 3-91; Discrete: 3-84, 3-87, and 3-89

Supplemental Exercises

3-92. Let D_i denote the event that the primary failure mode is type i and let A denote the event that a board passes the test.

The sample space S = $\{A, D_1, D_2, D_3, D_4, D_5\}$ or S = $\{A, A'D_1, A'D_2, A'D_3, A'D_4, A'D_5\}$.

3-93. Let M denote a part with a moderate edge condition and let B denote a part with a below target bore.

a) $P(M \cap B) = 20/200 = 0.10$

b) $P(M \cup B) = P(M) + P(B) - P(M \cap B) = \frac{45}{200} + \frac{100}{200} - \frac{20}{200} = \frac{125}{200} = \frac{5}{8}$

c) $P(M' \cup B') = P(M') + P(B') - P(M' \cap B') = \frac{155}{200} + \frac{100}{200} - \frac{75}{200} = \frac{180}{200} = 0.9$

Also, $P(M' \cup B') = 1 - P(M \cap B) = 1 - 0.1 = 0.9$

3-94. Let A_i denote the event that the ith order is shipped on time.

a) By independence, $P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2)P(A_3) = (0.95)^3 = 0.0857$

b) Let

$$B_1 = A_1' \cap A_2 \cap A_3$$
$$B_2 = A_1 \cap A_2' \cap A_3$$
$$B_3 = A_1 \cap A_2 \cap A_3'$$

Then, because the B's are mutually exclusive,

$$P(B_1 \cup B_2 \cup B_3) = P(B_1) + P(B_2) + P(B_3)$$
$$= 3(0.95)^2(0.05)$$
$$= 0.135$$

c) Let

$$B_1 = A_1' \cap A_2' \cap A_3$$
$$B_2 = A_1' \cap A_2 \cap A_3'$$
$$B_3 = A_1 \cap A_2' \cap A_3'$$
$$B_4 = A_1' \cap A_2' \cap A_3'$$

Because the B's are mutually exclusive,

$$P(B_1 \cup B_2 \cup B_3 \cup B_4) = P(B_1) + P(B_2) + P(B_3) + P(B_4)$$
$$= 3(0.05)^2(0.95) + (0.05)^3$$
$$= 0.00725$$

3-95. Let A_i denote the event that the ith bolt selected is not torqued to the proper limit.

a) Then,

$$P(A_1 \cap A_2 \cap A_3 \cap A_4) = P(A_4 | A_1 \cap A_2 \cap A_3) P(A_1 \cap A_2 \cap A_3)$$
$$= P(A_4 | A_1 \cap A_2 \cap A_3) P(A_3 | A_1 \cap A_2) P(A_2 | A_1) P(A_1)$$
$$= \left(\frac{2}{17}\right)\left(\frac{3}{18}\right)\left(\frac{4}{19}\right)\left(\frac{5}{20}\right) = 0.282$$

b) Let B denote the event that at least one of the selected bolts are not properly torqued. Thus, B' is the event that all bolts are properly torqued. Then,

$$P(B) = 1 - P(B') = 1 - \left(\frac{15}{20}\right)\left(\frac{14}{19}\right)\left(\frac{13}{18}\right)\left(\frac{12}{17}\right) = 0.718$$

3-96. Let A,B denote the event that the first, second portion of the circuit operates. Then, P(A) = (0.99)(0.99)+0.9-(0.99)(0.99)(0.9) = 0.998

P(B) = 0.9+0.9-(0.9)(0.9) = 0.99 and

$P(A \cap B) = P(A)\,P(B) = (0.998)\,(0.99) = 0.988$

3-97. Let D denote the event that a container is incorrectly filled and let H denote the event that a container is filled under high-speed operation. Then,

a) $P(D) = P(D|H)P(H) + P(D|H')P(H') = 0.01(0.30) + 0.001(0.70) = 0.0037$

b) $$P(H|D) = \frac{P(D|H)P(D)}{P(D)} = \frac{0.001(0.70)}{0.0037} = 0.1892$$

3-98. The tool fails if any component fails. Let F denote the event that the tool fails. Then, $P(F') = 0.99^{10}$ by independence and $P(F) = 1 - 0.99^{10} = 0.0956$

3-99. a) By independence, $0.15^5 = 7.59 \times 10^{-5}$

b) Let A_i denote the events that the machine is idle at the time of your ith request. Using independence, the requested probability is

$$P(A_1A_2A_3A_4A_5' \text{ or } A_1A_2A_3A_4'A_5 \text{ or } A_1A_2A_3'A_4A_5 \text{ or } A_1A_2'A_3A_4A_5 \text{ or } A_1'A_2A_3A_4A_5)$$
$$= 0.15^4(0.85) + 0.15^4(0.85) + 0.15^4(0.85) + 0.15^4(0.85) + 0.15^4(0.85)$$
$$= 5(0.15^4)(0.85)$$
$$= 0.0022$$

c) As in part b, the requested probability is

$$P(A_1A_2A_3A_4'A_5' \text{ or } A_1A_2A_3'A_4A_5' \text{ or } A_1A_2A_3'A_4'A_5 \text{ or } A_1A_2'A_3A_4A_5' \text{ or } A_1A_2'A_3A_4'A_5 \text{ or}$$
$$A_1A_2'A_3'A_4A_5 \text{ or } A_1'A_2A_3A_4A_5' \text{ or } A_1'A_2A_3A_4'A_5 \text{ or } A_1'A_2A_3'A_4A_5 \text{ or } A_1'A_2'A_3A_4A_5)$$
$$= 10(0.15^3)(0.85^2)$$
$$= 0.0244$$

3-100. Let A_i denote the event that the ith washer selected is thicker than target.

a) $$P(A_3 | A_1'A_2') = \frac{30}{48} = 0.625$$

b)

$$P(A_1A_2A_3) = P(A_3|A_1A_2)P(A_1A_2)$$
$$= P(A_3|A_1A_2)P(A_2|A_1)P(A_1)$$
$$= \frac{28}{48}\frac{29}{49}\frac{30}{50}$$
$$= 0.2071$$

c) The requested probability can be written in terms of whether or not the first and second washer selected are thicker than the target. That is,

$$
\begin{aligned}
P(A_3) &= P(A_1A_2A_3 \text{ or } A_1A_2'A_3 \text{ or } A_1'A_2A_3 \text{ or } A_1'A_2'A_3) \\
&= P(A_3|A_1A_2)P(A_1A_2) + P(A_3|A_1A_2')P(A_1A_2') \\
&\quad + P(A_3|A_1'A_2)P(A_1'A_2) + P(A_3|A_1'A_2')P(A_1'A_2') \\
&= P(A_3|A_1A_2)P(A_2|A_1)P(A_1) + P(A_3|A_1A_2')P(A_2'|A_1)P(A_1) \\
&\quad + P(A_3|A_1'A_2)P(A_2|A_1')P(A_1') + P(A_3|A_1'A_2')P(A_2'|A_1')P(A_1') \\
&= \frac{28}{48}\left(\frac{30}{50}\frac{29}{49}\right) + \frac{29}{48}\left(\frac{20}{50}\frac{30}{49}\right) + \frac{29}{48}\left(\frac{20}{50}\frac{30}{49}\right) + \frac{30}{48}\left(\frac{20}{50}\frac{19}{49}\right) \\
&= 0.60
\end{aligned}
$$

3-101. a) If n washers are selected, then the probability they are all less than the target is $\frac{20}{50} \cdot \frac{19}{49} \cdots \frac{20-n+1}{50-n+1}$.

n	probability all selected washers are less than target
1	20/50 = 0.4
2	(20/50)(19/49) = 0.155
3	(20/50)(19/49)(18/48) = 0.058

Therefore, the answer is n = 3

b) Then event E that one or more washers is thicker than target is the complement of the event that all are less than target. Therefore, P(E) equals one minus the probability in part a. Therefore, n = 3.

3-102.

a) $P(A \cup B) = \frac{112+68+246}{940} = 0.453$

b) $P(A \cap B) = \frac{246}{940} = 0.261$

c) $P(A' \cup B) = \frac{514+68+246}{940} = 0.881$

d) $P(A' \cap B') = \frac{514}{940} = 0.547$

. e) $P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{246/940}{314/940} = 0.783$

f) $P(B|A) = \frac{P(B \cap A)}{P(A)} = \frac{246/940}{358/940} = 0.687$

3-103. Let E denote a read error and let S,O,P denote skewed, off-center, and proper alignments, respectively. Then,

a) $P(E) = P(E|S)\,P(S) + P(E|O)\,P(O) + P(E|P)\,P(P)$
$= 0.01(0.10) + 0.02(0.05) + 0.001(0.85)$
$= 0.00285$

b) $P(S|E) = \frac{P(E|S)P(S)}{P(E)} = \frac{0.01(0.10)}{0.00285} = 0.351$

3-104. Let A_i denote the event that the ith row operates. Then,

$P(A_1) = 0.98, P(A_2) = (0.99)(0.99) = 0.9801, P(A_3) = 0.9801, \quad P(A_4) = 0.98.$

The probability the circuit operates is

$1 - P(A_1')P(A_2')P(A_3')P(A_4') = 1 - (0.02)(0.0199)(0.0199)(0.02)$
$= 0.9999998$

Mind-Expanding Exercises

3-105. Let E denote a read error and let S, O, B, P denote skewed, off-center, both, and proper alignments, respectively.
P(E) = P(E|S)P(S) + P(E|O)P(O) + P(E|B)P(B) + P(E|P)P(P)
= 0.01(0.10) + 0.02(0.05) + 0.06(0.01) + 0.001(0.84) = 0.00344

3-106. Let n denote the number of washers selected.
a) The probability that all are less than the target is 0.4^n , by independence.

n	0.4^n
1	0.4
2	0.16
3	0.064

Therefore, n = 3
b) The requested probability is the complement of the probability requested in part a. Therefore, n = 3

3-107. Let x denote the number of kits produced.

Revenue at each demand

	0	50	100	200
$0 \le x \le 50$	-5x	100x	100x	100x
Mean profit = 100x(0.95)-5x(0.05)-20x				
$50 \le x \le 100$	-5x	100(50)-5(x-50)	100x	100x
Mean profit = [100(50)-5(x-50)](0.4) + 100x(0.55)-5x(0.05)-20x				
$100 \le x \le 200$	-5x	100(50)-5(x-50)	100(100)-5(x-100)	100x
Mean profit = [100(50)-5(x-50)](0.4) + [100(100)-5(x-100)](0.3) + 100x(0.25) - 5x(0.05) - 20x				

	Mean Profit	Maximum Profit
$0 \le x \le 50$	74.75 x	\$ 3737.50 at x=50
$50 \le x \le 100$	32.75 x + 2100	\$ 5375 at x=100
$100 \le x \le 200$	1.25 x + 5250	\$ 5500 at x=200

Therefore, profit is maximized at 200 kits. However, the difference in profit over 100 kits is small.

3-108. Let E denote the probability that none of the bolts are identified as incorrectly torqued. The requested probability is P(E'). Let X denote the number of bolts in the sample that are incorrect. Then,
P(E) = P(E|X=0)P(X=0) + P(E|X=1) P(X=1) + P(E|X=2) P(X=2) + P(E|X=3) P(X=3) + P(E|X=4)P(X=4)
and P(X=0) = (15/20)(14/19)(13/18)(12/17) = 0.2817. The remaining probability for x can be determined from the counting methods in Appendix B-1. Then,

$$P(X=1) = \frac{\binom{5}{1}\binom{15}{3}}{\binom{20}{4}} = \frac{\left(\frac{5!}{4!1!}\right)\left(\frac{15!}{3!12!}\right)}{\left(\frac{20!}{4!16!}\right)} = \frac{5!15!4!16!}{4!3!12!20!} = 0.4696$$

$$P(X=2) = \frac{\binom{5}{2}\binom{15}{2}}{\binom{20}{4}} = \frac{\left(\frac{5!}{3!2!}\right)\left(\frac{15!}{2!13!}\right)}{\left(\frac{20!}{4!16!}\right)} = 0.2167$$

$$P(X=3) = \frac{\binom{5}{3}\binom{15}{1}}{\binom{20}{4}} = \frac{\left(\frac{5!}{3!2!}\right)\left(\frac{15!}{1!14!}\right)}{\left(\frac{20!}{4!16!}\right)} = 0.0309$$

P(X=4) = (5/20)(4/19)(3/18)(2/17) = 0.0010 and P(E|X=0) = 1, P(E|X=1) = 0.05, P(E|X=2) = $0.05^2 = 0.0025$, P(E|X=3) = $0.05^3 = 1.25 \times 10^{-4}$, P(E|X=4) = $0.05^4 = 6.25 \times 10^{-6}$. Then,

$$P(E) = 1(0.2817) + 0.05(0.4696) + 0.0025(0.2167) + 1.25 \times 10^{-4}(0.0309)$$
$$+6.25 \times 10^{-6}(0.0010)$$
$$= 0.306$$

and P(E') = 0.694

3-109.

$$P(A'\cap B') = 1 - P([A'\cap B']') = 1 - P(A \cup B)$$
$$= 1 - [P(A) + P(B) - P(A \cap B)]$$
$$= 1 - P(A) - P(B) + P(A)P(B)$$
$$= [1 - P(A)][1 - P(B)]$$
$$= P(A')P(B')$$

3-110. The total sample size is ka + a + kb + b = (k + 1)a + (k + 1)b.

$$P(A) = \frac{k(a+b)}{(k+1)a + (k+1)b} \quad , P(B) = \frac{ka + a}{(k+1)a + (k+1)b}$$

and

$$P(A \cap B) = \frac{ka}{(k+1)a + (k+1)b} = \frac{ka}{(k+1)(a+b)}$$

Then,

$$P(A)P(B) = \frac{k(a+b)(ka+a)}{[(k+1)a + (k+1)b]^2} = \frac{k(a+b)(k+1)a}{(k+1)^2(a+b)^2} = \frac{ka}{(k+1)(a+b)} = P(A \cap B)$$

CHAPTER 4

Section 4-1

4-1. The range of X is $\{0,1,2,\ldots,1000\}$

4-2. The range of X is $\{0,1,2,\ldots,50\}$

4-3. The range of X is $\{0,1,2,\ldots,99999\}$

4-4. The range of X is $\{0,1,2,3,4,5\}$

4-5. The range of X is $\{0,1,2,\ldots,491\}$. Because 490 parts are conforming, a nonconforming part must be selected in 491 selections.

4-6. The range of X is $\{0,1,2,\ldots,100\}$. Although the range actually obtained from lots typically might not exceed 10%.

4-7. The range of X is conveniently modeled as all nonnegative integers. That is, the range of X is $\{0,1,2,\ldots\}$

4-8. The range of X is conveniently modeled as all nonnegative integers. That is, the range of X is $\{0,1,2,\ldots\}$

4-9. The range of X is $\{0,1,2,\ldots,15\}$

4-10. The possible totals for two orders are 1/8 + 1/8 = 1/4, 1/8 + 1/4 = 3/8, 1/8 + 3/8 = 1/2, 1/4 + 1/4 = 1/2, 1/4 + 3/8 = 5/8, 3/8 + 3/8 = 6/8.
Therefore the range of X is $\left\{\frac{1}{4},\frac{3}{8},\frac{1}{2},\frac{5}{8},\frac{6}{8}\right\}$

Section 4-2

4-11.

$$f_X(0) = P(X = 0) = 1/6 + 1/6 = 1/3$$
$$f_X(1.5) = P(X = 1.5) = 1/3$$
$$f_X(2) = 1/6$$
$$f_X(3) = 1/6$$

4-12. a) P(X=1.5) = 1/3
b) P(0.5< X < 2.7) = P(X=1.5) + P(X=2) = 1/6 + 1/3 = 1/3
c) P(X > 3) = 0
d) P (0 ≤ X < 2) = P(X=0) + P(X=1.5) = 1/3 + 1/3 = 2/3
e) P(X=0 or X=2) = 1/3 + 1/6 = 1/2

4-13. All probabilities are greater than or equal to zero and sum to one.
a) 1
b) 0
c) 6/8 = 3/4
d) 1/8 + 2/8 +1/8 = 4/8 = 1/2

4-14. Probabilities are nonnegative and sum to one.
a) P(X ≤ 1) = (8/7)(1/2) = 8/14 = 4/7
b) P(X > 1) = (8/7)(1/4 + 1/8) = 3/7

4-15. Probabilities are nonnegative and sum to one.
a) P(X = 4) = 9/25
b) P(X ≤ 1) = 1/25 + 3/25 = 4/25
c) P(2 ≤ X < 4) = 5/25 + 7/25 = 12/25
d) P(X > −10) = 1

4-16. Probabilities are nonnegative and sum to one.

a) $P(X = 2) = 3/4(1/4)^2 = 3/64$

b) $P(X \le 2) = 3/4[1+1/4+(1/4)^2] = 63/64$

c) $P(X > 2) = 1 - P(X \le 2) = 1/64$

d) $P(X \ge 1) = 1 - P(X \le 0) = 1 - (3/4) = 1/4$

4-17. P(X = 10 million) = 0.3, P(X = 5 million) = 0.6, P(X = 1 million) = 0.1

4-18. Let C_i denote the event that part I is correctly classified

$P(X = 0) = 0.02^3 = 8 \times 10^{-6}$

$P(X = 1) = 3[0.98(0.02)(0.02)]=0.0012$

$P(X = 2) = 3[0.98(0.98)(0.02)]=0.0576$

$P(X = 3) = 0.98^3 = 0.9412$

4-19. X = number of wafers that pass

a) $P(X=3) = (0.8)^3 = 0.512$

b) $P(X=0) = (0.2)^3 = 0.008$

$P(X=1) = 3(0.2)^2(0.8) = 0.096$

$P(X=2) = 3(0.2)(0.8)^2 = 0.384$

$P(X=3) = (0.8)^3 = 0.512$

4-20. X = number of components that meet specifications

a) $P(X=2) = (0.95)(0.98) = 0.931$

b) $P(X=0) = (0.05)(0.02) = 0.001$

$P(X=1) = (0.05)(0.98) + (0.95)(0.02) = 0.068$

$P(X=2) = (0.95)(0.98) = 0.931$

4-21. X = number of components that meet specifications

a) $P(X=3) = (0.95)(0.98)(0.99) = 0.92169$

b) $P(X=0) = (0.05)(0.02)(0.01) = 0.0001$

$P(X=1) = (0.95)(0.02)(0.01) + (0.05)(0.98)(0.01) = 0.00167$

$P(X=2) = (0.95)(0.98)(0.01) + (0.95)(0.02)(0.99) + (0.05)(0.98)(0.99) = 0.07663$

$P(X=3) = (0.95)(0.98)(0.99) = 0.92169$

4-22. Using the rule for a sum of a geometric series:

$$\sum_{x=0}^{\infty} r^n (a_0) = \frac{a_0}{1-r} \quad \text{where} \quad |r| < 1$$

Then,

$$\sum_{x=1}^{\infty} (0.99)^{x-1}(0.01) = \sum_{x=0}^{\infty} (0.99)^x (0.01) = \frac{0.01}{1-0.99} = 1$$

<u>Section 4-3</u>

4-23.

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{3}, & 0 \le x < 1.5 \\ \frac{2}{3}, & 1.5 \le x < 2 \\ 1, & 2 \le x \end{cases}$$

4-24.

$$F(x) = \begin{cases} 0, & x < -2 \\ 1/8 & -2 \le x < -1 \\ 3/8 & -1 \le x < 0 \\ 5/8 & 0 \le x < 1 \\ 7/8 & 1 \le x < 2 \\ 1 & 2 \le x \end{cases}$$

a) $P(X \le 1.25) = 7/8$
b) $P(X \le 2.2) = 1$
c) $P(-1.1 < X \le 1) = 7/8 - 1/8 = 3/4$
d) $P(X > 0) = 1 - P(X \le 0) = 1 - 5/8 = 3/8$

4-25.

$$F(x) = \begin{cases} 0, & x < 0 \\ 1/25 & 0 \le x < 1 \\ 4/25 & 1 \le x < 2 \\ 9/25 & 2 \le x < 3 \\ 16/25 & 3 \le x < 4 \\ 1 & 4 \le x \end{cases}$$

a) $P(X < 1.5) = 4/25$
b) $P(X \le 3) = 16/25$
c) $P(X > 2) = 1 - P(X \le 2) = 1 - 9/25 = 16/25$
d) $P(1 < X \le 2) = P(X \le 2) - P(X \le 1) = 9/25 - 4/25 = 5/25 = 1/5$

4-26.

$$F(x) = \begin{cases} 0, & x < 1 \\ 0.1, & 1 \le x < 5 \\ 0.7, & 5 \le x < 10 \\ 1, & 10 \le x \end{cases}$$

4-27.

$$F(x) = \begin{cases} 0, & x < 0 \\ 0.008, & 0 \le x < 1 \\ 0.104, & 1 \le x < 2 \\ 0.488, & 2 \le x < 3 \\ 1, & 3 \le x \end{cases}$$

4-28. a) $P(X \le 3) = 1$
b) $P(X \le 2) = 0.5$
c) $P(1 \le X \le 2) = P(X=1) = 0.5$
d) $P(X>2) = 1 - P(X\le 2) = 0.5$

4-29. a) $P(X\le 0.5) = 1$
b) $P(X\le 0.4) = 0.75$
c) $P(0.4 \le X \le 0.6) = P(X=0.5)=0.25$
d) $P(X<0) = 0.25$
e) $P(0\le X<0.1) = 0$
f) $P(-0.1<X<0.1) = 0$

4-30. a) $P(X\le 1/8) = 0.2$
b) $P(X\le 1/4) = 0.9$
c) $P(X\le 5/16) = 0.9$
d) $P(X>1/4) = 0.1$
e) $P(X\le 1/2) = 1$

Section 4-4

4-31. $E(X) = 2.3469$, $V(X) = 0.0243$

4-32. $E(X) = 1.33$, $V(X) = 1.139$

4-33. $E(X) = 0$, $V(X) = 1.5$

4-34. $E(X) = 2.8$, $V(X) = 0.52$

4-35. $E(X) = 6.1$, $V(X) = 7.89$

4-36. $E(X) = 2.4$, $V(X) = 0.48$

4-37. $x = 30$

Section 4-5

4-38. $E(X) = (0+100)/2 = 50$, $V(X) = [(100-0+1)^2-1]/12 = 850$

4-39. $E(X) = (3+1)/2 = 2$, $V(X) = [(3-1+1)^2 -1]/12 = 0.667$

4-40. $E(X) = \frac{1}{8}\left(\frac{1}{3}\right) + \frac{1}{4}\left(\frac{1}{3}\right) + \frac{3}{8}\left(\frac{1}{3}\right) = \frac{1}{4}$, $V(X) = \left(\frac{1}{8}\right)^2\left(\frac{1}{3}\right) + \left(\frac{1}{4}\right)^2\left(\frac{1}{3}\right) + \left(\frac{3}{8}\right)^2\left(\frac{1}{3}\right) - \left(\frac{1}{4}\right)^2 = 0.0104$

4-41. $X=(1/100)Y$, $Y = 15, 16, 17, 18, 19$.

$E(X) = (1/100)E(Y) = \frac{1}{100}\left(\frac{15+19}{2}\right) = 0.17$ $\quad V(X) = \left(\frac{1}{100}\right)^2\left[\frac{(19-15+1)^2-1}{12}\right] = 0.0002$

4-42. $X = 590 + 0.1Y$, $Y = 0, 1, 2, ..., 9$

$E(X) = 590 + 0.1\left(\frac{0+9}{2}\right) = 590.45$, $V(X) = (0.1)^2\left[\frac{(9-0+1)^2-1}{12}\right] = 0.0825$

4-43. The range of Y is 0, 5, 10, ..., 45, $E(X) = (0+9)/2 = 4.5$
$E(Y) = 0(1/10)+5(1/10)+...+45(1/10)$
$= 5[0(0.1) +1(0.1)+ ... +9(0.1)]$
$= 5E(X)$
$= 5(4.5)$
$= 22.5$
$V(X) = 25$, $V(Y) = 25$, $\sigma_Y = 14.36$

4-44. $E(cX) = \sum_x cxf(x) = c\sum_x xf(x) = cE(X)$,

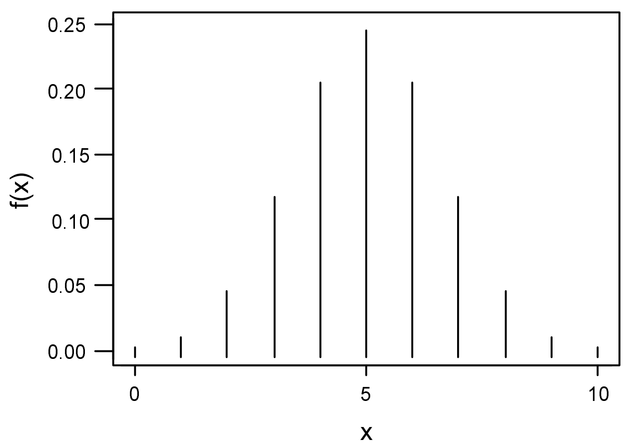
$V(cX) = \sum_x (cx - c\mu)^2 f(x) = c^2\sum_x (x-\mu)^2 f(x) = cV(X)$

Section 4-6

4-45. A binomial distribution is based on independent trials with two outcomes and a constant probability of success on each trial.

a) reasonable
b) independence assumption not reasonable
c) The probability that the second component fails depends on the failure time of the first component. The binomial distribution is not reasonable.
d) not independent trials with constant probability
e) probability of a correct answer not constant.
f) reasonable
g) probability of finding a defect not constant.
h) if the fills are independent with a constant probability of an underfill, then the binomial distribution for the number packages underfilled is reasonable.
i) because of the bursts, each trial (that consists of sending a bit) is not independent
j) not independent trials with constant probability

4-46.



a) The value of x that appears to be most likely is 5.
b) The values that are least likely appear to be 0 and 10, the extreme values.

4-47. a) $P(X = 5) = \binom{10}{5} 0.5^5 (0.5)^5 = 0.246$

b) $P(X \le 2) = \binom{10}{0} 0.5^0 0.5^{10} + \binom{10}{1} 0.5^1 0.5^9 + \binom{10}{2} 0.5^2 0.5^8$

$= 0.5^{10} + 10(0.5)^{10} + 45(0.5)^{10}$

$= 0.055$

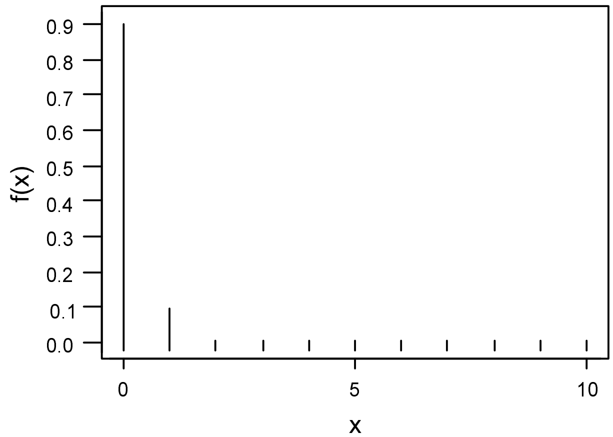
c) $P(X \ge 9) = \binom{10}{9} 0.5^9 (0.5)^1 + \binom{10}{10} 0.5^{10} (0.5)^0 = 0.011$

d) $P(3 \le X < 5) = \binom{10}{3} 0.5^3 0.5^7 + \binom{10}{4} 0.5^4 0.5^6$

$= 120(0.5)^{10} + 210(0.5)^{10}$

$= 0.322$

4-48.



a) The value of X that appears to be most likely is 0.
b) The value of X that appears to be least likely is 10, although the probabilities for values of x greater than 1 are very small..

4-49. a) $P(X = 5) = \binom{10}{5} 0.01^5 (0.99)^5 = 2.40 \times 10^{-8}$

$$b) P(X \le 2) = \binom{10}{0} 0.01^0 (0.99)^{10} + \binom{10}{1} 0.01^1 (0.99)^9 + \binom{10}{2} 0.01^2 (0.99)^8$$
$$= 0.9999$$

$$c) P(X \ge 9) = \binom{10}{9} 0.01^9 (0.99)^1 + \binom{10}{10} 0.01^{10} (0.99)^0$$
$$= 9.91 \times 10^{-18}$$

$$d) P(3 \le X < 5) = \binom{10}{3} 0.01^3 (0.99)^7 + \binom{10}{4} 0.01^4 (0.99)^6$$
$$= 1.138 \times 10^{-4}$$

4-50.

$$F(x) = \begin{cases} 0, & x < 0 \\ 0.125, & 0 \le x < 1 \\ 0.5, & 1 \le x < 2 \\ 0.875, & 2 \le x < 3 \\ 1, & 3 \le x \end{cases}$$

4-51.

$$F(x) = \begin{cases} 0, & x < 0 \\ 0.4219, & 0 \le x < 1 \\ 0.8438, & 1 \le x < 2 \\ 0.9844, & 2 \le x < 3 \\ 1, & 3 \le x \end{cases}$$

4-52. Let X denote the number of defective circuits. Then, X has a binomial distribution with n = 40 and p = 0.01. Then, $P(X = 0) = \binom{40}{0} 0.01^0 0.99^{40} = 0.6690$.

4-53. a) $P(X = 1) = \binom{1000}{1} 0.001^1 (0.999)^{999} = 0.368$

$$\text{b)}\, P(X \ge 1) = 1 - P(X = 0) = 1 - \binom{1000}{1} 0.001^1 (0.999)^{999} = 0.632$$

$$\text{c)}\, P(X \le 2) = \binom{1000}{0} 0.001^0 (0.999)^{1000} + \binom{1000}{1} 0.001^1 (0.999)^{999} + \binom{1000}{2} 0.001^2 0.999^{998}$$
$$= 0.920$$

d) $E(X) = 1000(0.001) = 1$

$V(X) = 1000(0.001)(0.999) = 0.999$

4-54. Let X denote the number of times the line is occupied. Then, X has a binomial distribution with n = 10 and p = 0.4

a) $P(X = 3) = \binom{10}{3} 0.4^3 (0.6)^7 = 0.215$

$$\text{b)}\, P(X \ge 1) = 1 - P(X = 0)$$
$$= 1 - \binom{10}{0} 0.4^0 (0.6)^{10}$$
$$= 0.994$$

c) $E(X) = 10(0.4) = 4$

4-55. a) n = 50, p = 5/50 = 0.1, since E(X) = 5 = np.

$$\text{b)}\, P(X \le 2) = \binom{50}{0} 0.1^0 (0.9)^{50} + \binom{50}{1} 0.1^1 (0.9)^{49} + \binom{50}{2} 0.1^2 (0.9)^{48} = 0.112$$

$$\text{c)}\, P(X \ge 49) = \binom{50}{49} 0.1^{49} (0.9)^1 + \binom{50}{50} 0.1^{50} (0.9)^0 = 4.51 \times 10^{-48}$$

4-56. E(X) = 20 (0.01) = 0.2

V(X) = 20 (0.01) (0.99) = 0.198

$$\mu_X + 3\sigma_X = 0.2 + 3\sqrt{0.198} = 1.53$$

a) $P(X > 1.53) = P(X \ge 2) = 1 - P(X \le 1)$

$$= 1 - \left[\binom{20}{0} 0.01^0 0.99^{20} + \binom{20}{1} 0.01^1 0.99^{19}\right]$$
$$= 0.0169$$

b) X is binomial with n = 20 and p = 0.04

$$P(X > 1) = 1 - P(X \le 1)$$
$$= 1 - \left[\binom{20}{0} 0.04^0 (0.96)^{20} + \binom{20}{1} 0.04^1 (0.96)^{19}\right]$$
$$= 0.1897$$

c) Let Y denote the number of times X exceeds 1 in the next five samples. Then, Y is binomial with n = 5 and p = 0.190 from part b.

$$P(Y \ge 1) = 1 - P(Y = 0) = 1 - \left[\binom{5}{0} 0.190^0 (0.810)^5\right] = 0.651$$

The probability is 0.651 that at least one sample from the next five will contain more than one defective.

4-57. Let X denote the passengers with tickets that do not show up for the flight. Then, X is binomial with n = 125 and p = 0.1.

a) $P(X \ge 5) = 1 - P(X \le 4)$

$$= 1 - \left[\begin{array}{l}\binom{125}{0}0.1^0(0.9)^{125} + \binom{125}{1}0.1^1(0.9)^{124} + \binom{125}{2}0.1^2(0.9)^{123} \\ +\binom{125}{3}0.1^3(0.9)^{122} + \binom{125}{4}0.1^4(0.9)^{121}\end{array}\right]$$

$= 0.996$

b) $P(X > 5) = 1 - P(X \le 5) = 0.989$

c) $E(X) = np = 125(0.1) = 12.5$

$V(X) = \sqrt{np(1-p)} = \sqrt{125(0.1)(0.9)} = 3.354$

4-58. Let X denote the number of defective components among those stocked.

a) $P(X = 0) = \binom{100}{0}0.02^0(0.98)^{100} = 0.133$

b) $P(X \le 2) = \binom{102}{0}0.02^0 0.98^{102} + \binom{102}{1}0.02^1 0.98^{101} + \binom{102}{2}0.02^2 0.98^{100} = 0.666$

c) $P(X \le 5) = 0.981$

4-59. Let X denote the number of questions answered correctly. Then, X is binomial with n = 25 and p = 0.25.

a) $P(X \ge 20) = \binom{25}{20}0.25^{20}(0.75)^5 + \binom{25}{21}0.25^{21}(0.75)^4 + \binom{25}{22}0.25^{22}(0.75)^3$

$+\binom{25}{23}0.25^{23}(0.75)^2 + \binom{25}{24}0.25^{24}(0.75)^1 + \binom{25}{25}0.25^{25}(0.75)^0 = 0.0000$

b) $P(X < 5) = \binom{25}{0}0.25^0(0.75)^{25} + \binom{25}{1}0.25^1(0.75)^{24} + \binom{25}{2}0.25^2(0.75)^{23}$

$+\binom{25}{3}0.25^3(0.75)^{22} + \binom{25}{4}0.25^4(0.75)^{21} = 0.2137$

c) $E(X) = np = 25(0.25) = 6.25$ $\qquad V(X) = \sqrt{np(1-p)} = \sqrt{25(0.25)(0.75)} = 2.165$

4-60. Let X denote the number of mornings the light is green.

a) $P(X = 1) = \binom{5}{1}0.2^1(0.8)^4 = 0.410$

b) $P(X = 4) = \binom{20}{4}0.2^4(0.8)^{16} = 0.218$

c) $P(X > 4) = 1 - P(X \le 4) = 1 - 0.630 = 0.370$

Section 4-7

4-61. a. 0.5 b. $0.5^4 = 0.0625$ c. $0.5^8 = 0.0039$ d. $P(X \le 2) = P(X=1) + P(X=2) = 0.75$

4-62. $E(X) = 2.5 = 1/p$ giving $p = 0.4$

a) 0.4 b) 0.0864 c) 0.0518 d) 0.7840 e) 0.2160

4-63. Let X denote the umber of trials to obtain the first successful alignment. Then X is a geometric random variable with p = 0.8

a) $0.2^3(0.8) = 0.0064$

b) $P(X \le 4) = P(X=1) + P(X = 2) + P(X =3) + P(X = 4) = 0.9984$

c) $P(X \ge 4) = 1 - P(X \le 3) = 1 - 0.992 = 0.008$

4-64. Let X denote the number of calls needed to obtain a connection. Then, X is a geometric random variable with p = 0.02

a) $P(X=10) = (0.98)^9 0.002 = 0.0163$

b) $P(X>5) = 1 - P(X \le 4) = 1 - [P(X=1) + P(X = 2) + P(X =3) + P(X = 4)] = 1 - 0.0776 = 0.9224$

c) $E(X) = 1/0.02 = 50$

4-65. Let X denote the number of mornings needed to obtain a green light. Then X is a geometric random variable with p = 0.20.

a) $P(X = 4) = 0.1024$

b) By independence, $(0.8)^{10} = 0.1074$. (Also, $P(X > 10) = 0.1074$)

4-66. let Y denote the number of samples needed to exceed 1 in Exercise 4-56. Then Y has a geometric distribution with p = 0.0169.

a) $P(Y = 10) = (1 - 0.0169)^9(0.0169) = 0.0145$

b) Y is a geometric random variable with p = 0.1897 from Exercise 4-56.
$P(Y = 10) = (1 - 0.0169)^9(0.0169) = 0.0286$

c) $E(Y) = 1/0.1897 = 5.27$

4-67. Let X denote the number of trials to obtain the first success.

a) $E(X) = 1/0.2 = 5$

b) Because of the lack of memory property, the expected value is still 5.

4-68. Negative binomial random variable: $f(x; p, r) = \binom{x-1}{r-1}(1-p)^{x-r}p^r$. When r = 1, this reduces to f(x; p, r) = $(1-p)^{x-1}p$, which is the pdf of a geometric random variable. Also, $E(X) = r/p$ and $V(X) = [r(1-p)]/p^2$ reduce to $E(X) = 1/p$ and $V(X) = (1-p)/p^2$, respectively.

4-69. a) $E(X) = 4/0.2 = 20$

b) $P(X=20) = \binom{19}{3}(0.80)^{16}0.2^4 = 0.0436$

c) $P(X=19) = \binom{18}{3}(0.80)^{15}0.2^4 = 0.0459$

d) $P(X=21) = \binom{20}{3}(0.80)^{17}0.2^4 = 0.0411$

e) The most likely value for X should be near μ_X. By trying several cases, the most likely value is X = 19.

4-70. Let X denote the number of attempts needed to obtain a calibration that conforms to specifications. Then, X is geometric with p = 0.6.

$P(X \le 3) = P(X=1) + P(X=2) + P(X=3) = 0.6 + 0.4(0.6) + 0.4^2(0.6) = 0.936.$

4-71. Let X denote the number of fills needed to detect three underweight packages. Then X is a negative binomial random variable with p = 0.001 and r = 3.

a) $E(X) = 3/0.001 = 3000$

b) $V(X) = [3(0.999)/0.001^2] = 2997000$. Therefore, $\sigma_X = 1731.18$

4-72. Let X denote the number of transactions until all computers have failed. Then, X is negative binomial random variable with $p = 10^{-8}$ and r = 3.

a) $E(X) = 3 \times 10^{-8}$

b) $V(X) = [3(1-10^{-8}0]/(10^{-16}) = 2.99 \times 10^{16}$

4-73. Let X denote a geometric random variable with parameter p. Let q = 1-p.

$$E(X) = \sum_{x=1}^{\infty} x(1-p)^{x-1}p$$
$$= p\sum_{x=1}^{\infty} xq^{x-1} = p\frac{d}{dq}\left[\sum_{x=0}^{\infty} q^x\right] = p\frac{d}{dq}\left[\frac{1}{1-q}\right]$$
$$= \frac{p}{(1-q)^2} = \frac{p}{p^2} = \frac{1}{p}$$

$$V(X) = \sum_{x=1}^{\infty} (1-\tfrac{1}{p})^2(1-p)^{x-1}p = \sum_{x=1}^{\infty}\left(px^2 - 2x + \tfrac{1}{p}\right)(1-p)^{x-1}$$
$$= p\sum_{x=1}^{\infty} x^2q^{x-1} - 2\sum_{x=1}^{\infty} xq^{x-1} + \frac{1}{p}\sum_{x=1}^{\infty} q^{x-1}$$
$$= p\sum_{x=1}^{\infty} x^2q^{x-1} - \frac{2}{p^2} + \frac{1}{p^2}$$
$$= p\sum_{x=1}^{\infty} x^2q^{x-1} - \frac{1}{p^2}$$
$$= p\frac{d}{dq}\left[q + 2q^2 + 3q^3 + \ldots\right] - \frac{1}{p^2}$$
$$= p\frac{d}{dq}\left[q(1 + 2q + 3q^2 + \ldots)\right] - \frac{1}{p^2}$$
$$= p\frac{d}{dq}\left[\frac{q}{(1-q)^2}\right] - \frac{1}{p^2} = 2pq(1-q)^{-3} + p(1-q)^{-2} - \frac{1}{p^2}$$
$$= \frac{[2(1-p) + p - 1]}{p^2} = \frac{(1-p)}{p^2} = \frac{q}{p^2}$$

Section 4-8

4-74. a) $P(X=1) = \dfrac{\binom{20}{1}\binom{80}{3}}{\binom{100}{4}} = \dfrac{(20\times 80\times 79\times 78)/6}{(100\times 99\times 98\times 97)/24} = 0.4191$

b) P(X = 6) = 0, the sample size is only 4.

c) $P(X=4) = \dfrac{\binom{20}{4}\binom{80}{0}}{\binom{100}{4}} = \dfrac{(20\times 19\times 18\times 17)/24}{(100\times 99\times 98\times 97)/24} = 0.0012$

d) E(X) = 4(20/100) = 0.8
V(X) = 4(0.2)(0.8)(96/99) = 0.6206

4-75. a) $P(X=1)=\frac{\binom{4}{1}\binom{16}{3}}{\binom{20}{4}}=\frac{(4\times 16\times 15\times 14)/6}{(20\times 19\times 18\times 17)/24}=0.4623$

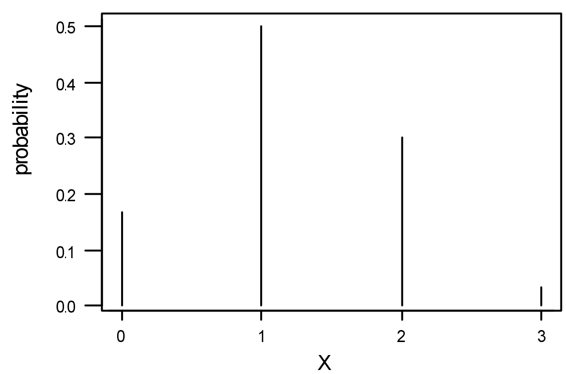
b) $P(X=4)=\frac{\binom{4}{4}\binom{16}{0}}{\binom{20}{4}}=\frac{1}{(20\times 19\times 18\times 17)/24}=0.00021$

c)

$$P(X\le 2)=P(X=0)+P(X=1)+P(X=2)$$
$$=\frac{\binom{4}{0}\binom{16}{4}}{\binom{20}{4}}+\frac{\binom{4}{1}\binom{16}{3}}{\binom{20}{4}}+\frac{\binom{4}{2}\binom{16}{2}}{\binom{20}{4}}$$
$$=\frac{\left(\frac{16\times 15\times 14\times 13}{24}+\frac{4\times 16\times 15\times 14}{6}+\frac{6\times 16\times 15}{2}\right)}{\left(\frac{20\times 19\times 18\times 17}{4}\right)}=0.1644$$

d) E(X) = 4(4/20) = 0.8
V(X) = 4(0.2)(0.8)(16/19) = 0.539

4-76.



$$P(X=0)=\frac{\binom{4}{0}\binom{6}{3}}{\binom{10}{3}}=\frac{\frac{6.5\times 4}{6}}{\frac{10\times 9\times 8}{6}}=\frac{20}{120}=\frac{1}{6}$$

$$P(X=1)=\frac{\binom{4}{1}\binom{6}{2}}{\binom{10}{3}}=\frac{\frac{4\times 6\times 5}{2}}{\frac{10\times 9\times 8}{6}}=\frac{1}{2}$$

$$P(X=2)=\frac{\binom{4}{2}\binom{6}{1}}{\binom{10}{3}}=\frac{6\times 6}{120}=\frac{3}{10}$$

$$P(X=3)=\frac{\binom{4}{3}\binom{6}{0}}{\binom{10}{3}}=\frac{4}{120}=\frac{1}{30}$$

4-77.

$$F(x) = \begin{cases} 0, & x < 0 \\ 1/6, & 0 \le x < 1 \\ 2/3, & 1 \le x < 2 \\ 29/30, & 2 \le x < 3 \\ 1, & 3 \le x \end{cases}$$

4-78. Let X denote the number of unacceptable washers in the sample of 10.

a. $P(X=0) = \dfrac{\binom{5}{0}\binom{70}{10}}{\binom{75}{10}} = \dfrac{\frac{70!}{10!60!}}{\frac{75!}{10!65!}} = \dfrac{65 \times 64 \times 6 \times 62 \times 61}{75 \times 74 \times 73 \times 72 \times 71} = 0.4786$

b. $P(X \ge 1) = 1 - P(X=0) = 0.5214$

c. $P(X=1) = \dfrac{\binom{5}{1}\binom{70}{9}}{\binom{75}{10}} = \dfrac{\frac{5!70!}{9!61!}}{\frac{75!}{10!65!}} = \dfrac{5 \times 65 \times 64 \times 63 \times 62 \times 10}{75 \times 74 \times 73 \times 72 \times 71} = 0.3923$

d. $E(X) = 10(5/75) = 2/3$

4-79. Let X denote the number of cards in the sample that are defective.

a)

$P(X \ge 1) = 1 - P(X=0)$

$P(X=0) = \dfrac{\binom{20}{0}\binom{120}{20}}{\binom{140}{20}} = \dfrac{\frac{120!}{20!100!}}{\frac{140!}{20!120!}} = \dfrac{120!120!}{100!140!} = 0.0424$

$P(X \ge 1) = 1 - 0.0424 = 0.9576$

b)

$P(X \ge 1) = 1 - P(X=0)$

$P(X=0) = \dfrac{\binom{5}{0}\binom{135}{20}}{\binom{140}{20}} = \dfrac{\frac{135!}{20!115!}}{\frac{140!}{20!120!}} = \dfrac{135!120!}{115!140!} = 0.4571$

$P(X \ge 1) = 1 - 0.4571 = 0.5429$

4-80. Let X denote the number of blades in the sample that are dull.

a)

$P(X \ge 1) = 1 - P(X=0)$

$P(X=0) = \dfrac{\binom{10}{0}\binom{38}{5}}{\binom{48}{5}} = \dfrac{\frac{38!}{5!33!}}{\frac{48!}{5!43!}} = \dfrac{38!43!}{48!33!} = 0.2931$

$P(X \ge 1) = 1 - P(X=0) = 0.7069$

b) Let Y denote the number of days needed to replace the assembly.

$P(Y=3) = 0.2931^2(0.7069) = 0.0607$

c) On the first day, $P(X=0)=\dfrac{\binom{2}{0}\binom{46}{5}}{\binom{48}{5}}=\dfrac{\frac{46!}{5!41!}}{\frac{48!}{5!43!}}=\dfrac{46!43!}{48!41!}=0.8005$

On the second day, $P(X=0)=\dfrac{\binom{6}{0}\binom{42}{5}}{\binom{48}{5}}=\dfrac{\frac{42!}{5!37!}}{\frac{48!}{5!43!}}=\dfrac{42!43!}{48!37!}=0.4968$

On the third day, P(X = 0) = 0.2931 from part a. Therefore,
P(Y = 3) = 0.8005(0.4968)(1-0.2931) = 0.2811.

4-81. Let X denote the count of the numbers in the state's sample that match those in the player's sample. Then, X has a hypergeometric distribution with N = 48, n = 6, and K = 6.

a) $P(X=6)=\dfrac{\binom{6}{6}\binom{42}{0}}{\binom{48}{6}}=\left(\dfrac{48!}{6!42!}\right)^{-1}=\dfrac{1}{12{,}271{,}512}=8.15\times10^{-8}$

b) $P(X=5)=\dfrac{\binom{6}{5}\binom{42}{1}}{\binom{48}{6}}=\dfrac{6\times42}{\binom{48}{6}}=2.05\times10^{-5}$

c) $P(X=4)=\dfrac{\binom{6}{4}\binom{42}{2}}{\binom{48}{6}}=\dfrac{\frac{15\times42\times41}{2}}{\binom{48}{6}}=0.00105$

d) Let Y denote the number of weeks needed to match all six numbers. Then, Y has a geometric distribution with p = $\dfrac{1}{12{,}271{,}512}$ and E(Y) = 1/p = 12,271,512 weeks. This is more than 2359 centuries!

4-82. a) For Exercise 4-74, the finite population correction is 96/99.
For Exercise 4-75, the finite population correction is 16/19.
Because the finite population correction for Exercise 4-74 is closer to one, the binomial approximation to the distribution of X should be better in Exercise 4-74.

b) Assuming X has a binomial distribution with n = 4 and p = 0.2,

$P(X=1)=\binom{4}{1}0.2^{1}0.8^{3}=0.4096$

$P(X=4)=\binom{4}{4}0.2^{4}0.8^{0}=0.0016$

The results from the binomial approximation are close to the probabilities obtained in Exercise 4-74.

c) Assume X has a binomial distribution with n = 4 and p = 0.2. Consequently, P(X = 1) and P(X = 4) are the same as computed in part b. of this exercise. This binomial approximation is not as close to the true answer as the results obtained in part b. of this exercise.

4-83. From Exercise 4-79, X is approximately binomial with n = 10 and p = 5/75 = 1/15.

a) $P(X=0)=\binom{10}{0}\left(\frac{1}{15}\right)^{0}\left(\frac{14}{15}\right)^{10}=0.5016$

b) $P(X\geq1)=1-P(X=0)=0.4984$

c) $P(X=1)=\binom{10}{1}\left(\frac{1}{15}\right)^{1}\left(\frac{14}{15}\right)^{9}=0.3583$

d) E(X) = 10(1/15) = 2/3.

Section 4-9

4-84. a) $P(X=0)=\frac{e^{-4}4^0}{0!}=e^{-4}=0.0183$

b) $P(X\le 2)=P(X=0)+P(X=1)+P(X=2)$

$$=e^{-4}+\frac{e^{-4}4^1}{1!}+\frac{e^{-4}4^2}{2!}$$
$$=0.2381$$

c) $P(X=4)=\frac{e^{-4}4^4}{4!}=0.1954$

d) $P(X=8)=\frac{e^{-4}4^8}{8!}=0.0298$

4-85. a) $P(X=0)=e^{-0.4}=0.6703$

b) $P(X\le 2)=e^{-0.4}+\frac{e^{-0.4}(0.4)}{1!}+\frac{e^{-0.4}(0.4)^2}{2!}=0.9921$

c) $P(X=4)=\frac{e^{-0.4}(0.4)^4}{4!}=0.000715$

d) $P(X=8)=\frac{e^{-0.4}(0.4)^8}{8!}=1.09\times 10^{-8}$

4-86. $P(X=0)=e^{-\lambda}=0.05$. Therefore, $\lambda=-\ln(0.05)=2.996$.
Consequently, $E(X)=V(X)=2.996$.

4-87. a) Let X denote the number of calls in one hour. Then, X is a Poisson random variable with $\lambda = 10$.
$P(X=5)=\frac{e^{-10}10^5}{5!}=0.0378$.

b) $P(X\le 3)=e^{-10}+\frac{e^{-10}10}{1!}+\frac{e^{-10}10^2}{2!}+\frac{e^{-10}10^3}{3!}=0.0103$

c) Let Y denote the number of calls in two hours. Then, Y is a Poisson random variable with $\lambda=20$. $P(Y=15)=\frac{e^{-20}20^{15}}{15!}=0.0516$

d) Let W denote the number of calls in 30 minutes. Then W is a Poisson random variable with $\lambda=5$. $P(W=5)=\frac{e^{-5}5^5}{5!}=0.1755$

4-88. a) Let X denote the number of flaws in one square meter of cloth. Then, X is a Poisson random variable with $\lambda=0.1$. $P(X=2)=\frac{e^{-0.1}(0.1)^2}{2!}=0.0045$

b) Let Y denote the number of flaws in 10 square meters of cloth. Then, Y is a Poisson random variable with $\lambda=1$. $P(Y=1)=\frac{e^{-1}1^1}{1!}=e^{-1}=0.3679$

c) Let W denote the number of flaws in 20 square meters of cloth. Then, W is a Poisson random variable with $\lambda=2$. $P(W=0)=e^{-2}=0.1353$

d) $P(Y\ge 2)=1-P(Y\le 1)=1-P(Y=0)-P(Y=1)$
$$=1-e^{-1}-e^{-1}$$
$$=0.2642$$

4-89. a) Let X denote the number of cracks in 5 miles of highway. Then, X is a Poisson random variable with $\lambda = 10$. $P(X = 0) = e^{-10} = 4.54 \times 10^{-5}$

b) Let Y denote the number of cracks in a half mile of highway. Then, Y is a Poisson random variable with $\lambda = 1$. $P(Y \ge 1) = 1 - P(Y = 0) = 1 - e^{-1} = 0.6321$

c) The assumptions of a Poisson process require that the probability of a count is constant for all intervals. If the probability of a count depends on traffic load and the load varies, then the assumptions of a Poisson process are not valid. Separate Poisson random variables might be appropriate for the heavy and light load sections of the highway.

4-90. a) Let X denote the number of flaws in 10 square feet of plastic panel. Then, X is a Poisson random variable with $\lambda = 0.5$. $P(X = 0) = e^{-0.5} = 0.6065$

b) Let Y denote the number of cars with no flaws,

$$P(Y = 0) = \binom{10}{0}(0.6065)^0(0.3935)^{10} = 0.000089$$

c) Let W denote the number of cars with surface flaws. Because the number of flaws has a Poisson distribution, the occurrences of surface flaws in cars are independent events with constant probability. From part a., the probability a car contains surface flaws is 1−0.6065 = 0.3935. Consequently, W is binomial with n = 10 and p = 0.3935.

$$P(W \le 1) = P(W = 0) + P(W = 1)$$

$$= \binom{10}{0}0.3935^0(0.6065)^{10} + \binom{10}{1}0.3935^1(0.6065)^9$$

$$= 0.0504$$

4-91. a) Let X denote the failures in 8 hours. Then, X has a Poisson distribution with $\lambda = 0.16$.

$P(X = 0) = e^{-0.16} = 0.8521$

b) Let Y denote the number of failure in 24 hours. Then, Y has a Poisson distribution with $\lambda = 0.48$. $P(Y \ge 1) = 1 - P(Y = 0) = 1 - e^{-48} = 0.3812$

Supplemental Exercises

4-92. Let X denote the number of totes in the sample that do not conform to purity requirements. Then, X has a hypergeometric distribution with N = 15, n = 3, and K = 2.

$$P(X \ge 1) = 1 - P(X = 0) = 1 - \frac{\binom{2}{0}\binom{13}{3}}{\binom{15}{3}} = 1 - \frac{13!12!}{10!15!} = 0.3714$$

4-93. Let X denote the number of calls that are answered in 30 seconds or less. Then, X is a binomial random variable with p = 0.75.

a) $P(X = 9) = \binom{10}{9}(0.75)^9(0.25)^1 = 0.1877$

b) $P(X \ge 16) = P(X=16) + P(X=17) + P(X=18) + P(X=19) + P(X=20)$

$$= \binom{20}{16}(0.75)^{16}(0.25)^4 + \binom{20}{17}(0.75)^{17}(0.25)^3 + \binom{20}{18}(0.75)^{18}(0.25)^2$$

$$+ \binom{20}{19}(0.75)^{19}(0.25)^1 + \binom{20}{20}(0.75)^{20}(0.25)^0$$

$$= 0.4148$$

c) $E(X) = 20(0.75) = 15$

4-94. Let Y denote the number of calls needed to obtain an answer in less than 30 seconds.
a) $P(Y=4) = 0.25^3(0.75) = 0.0117$
b) $E(Y) = 1/p = 1/0.75 = 4/3$

4-95. Let W denote the number of calls needed to obtain two answers in less than 30 seconds. Then, W has a negative binomial distribution with p = 0.75.
a) $P(W=6) = \binom{5}{1}(0.25)^4(0.75)^2 = 0.0110$
b) $E(W) = r/p = 2/0.75 = 8/3$

4-96. a) Let X denote the number of messages sent in one hour. $P(X=5) = \frac{e^{-5}5^5}{5!} = 0.1755$
b) Let Y denote the number of messages sent in 1.5 hours. Then, Y is a Poisson random variable with $\lambda = 7.5$. $P(Y=10) = \frac{e^{-7.5}(7.5)^{10}}{10!} = 0.0858$
c) Let W denote the number of messages sent in one-half hour. Then, W is a Poisson random variable with $\lambda = 2.5$. $P(W<2) = P(W=0) + P(W=1) = 0.2873$

4-97. $X \sim \text{Poisson}(\lambda = 0.01)$, $X \sim \text{Poisson}(\lambda = 1)$
$P(X_{100} \leq 3) = 0.981012$

4-98. Let X denote the number of individuals that recover in one week. Assume the individuals are independent. Then, X is a binomial random variable with n = 20 and p = 0.1. $P(X \geq 4) = 1 - P(X \leq 3) = 1 - 0.8670 = 0.1330$.

4-99. Let X denote the number of assemblies needed to obtain a defective. Then, X is a geometric random variable with p = 0.01.
a) $E(X) = 1/p = 100$.
b) $V(X) = 0.99/0.01^2 = 9900$ and $\sigma_X = 99.50$

4-100. If n assemblies are checked, then let X denote the number of defective assemblies. If $P(X \geq 1) \geq 0.95$, then $P(X=0) \leq 0.05$. Now,
$P(X=0) = \binom{n}{0}(0.01)^0(0.25)^n = 99^n$ and $0.99^n \leq 0.05$. Therefore,
$$n(\ln(0.99)) \leq \ln(0.05)$$
$$n \geq \frac{\ln(0.05)}{\ln(0.95)} = 298.07$$
This would require n = 299.

4-101. Require f(1) + f(2) + f(3) + f(4) = 1. Therefore, c(1+2+3+4) = 1. Therefore, c = 0.1.

4-102. Let X denote the number of products that fail during the warranty period. Assume the units are independent. Then, X is a binomial random variable with n = 500 and p = 0.02.
a) $P(X = 0) = 4.1 \times 10^{-5}$
b) $E(X) = 500(0.02) = 10$
c) $P(X > 2) = 1 - P(X \leq 1) = 0.9995$

4-103. a) $P(X \leq 3) = 0.2 + 0.4 = 0.6$
b) $P(X > 2.5) = 0.4 + 0.3 + 0.1 = 0.8$
c) $P(2.7 < X < 5.1) = 0.4 + 0.3 = 0.7$
d) $E(X) = 2(0.2) + 3(0.4) + 5(0.3) + 8(0.1) = 3.9$
e) $V(X) = 2^2(0.2) + 3^2(0.4) + 5^2(0.3) + 8^2(0.1) - (3.9)^2 = 3.09$

4-104.

x	2	5.7	6.5	8.5
f(x)	0.2	0.3	0.3	0.2

4-105. Let X denote the number of bolts in the sample from supplier 1 and let Y denote the number of bolts in the sample from supplier 2. Then, x is a hypergeometric random variable with N = 100, n = 4, and K = 30. Also, Y is a hypergeometric random variable with N = 100, n = 4, and K = 70.

a) P(X=4 or Y=4) = P(X = 4) + P(Y = 4)

$$= \frac{\binom{30}{4}\binom{70}{0}}{\binom{100}{4}} + \frac{\binom{30}{0}\binom{70}{4}}{\binom{100}{4}}$$

$$= 0.2408$$

b) P[(X=3 and Y=1) or (Y=3 and X = 1)]= $= \dfrac{\binom{30}{3}\binom{70}{1} + \binom{30}{1}\binom{70}{3}}{\binom{100}{4}} = 0.4913$

4-106. Let X denote the number of errors in a sector. Then, X is a Poisson random variable with $\lambda = 0.32768$.

a) $P(X>1) = 1 - P(X\le1) = 1 - e^{-0.32768} - e^{-0.32768}(0.32768) = 0.0433$

b) Let Y denote the number of sectors until an error is found. Then, Y is a geometric random variable and $P = P(X \ge 1) = 1 - P(X=0) = 1 - e^{-0.32768} = 0.2794$

$E(Y) = 1/p = 3.58$

4-107. Let X denote the number of orders placed in a week in a city of 800,000 people. Then X is a Poisson random variable with $\lambda = 0.25(8) = 2$.

a) $P(X \ge 3) = 1 - P(X \le 2) = 1 - [e^{-2} + e^{-2}(2) + (e^{-2}2^2)/2!] = 1 - 0.6767 = 0.3233$.

b) Let Y denote the number of orders in 2 weeks. Then, Y is a Poisson random variable with $\lambda = 4$, and $P(Y<2) = P(Y \le 1) = e^{-4} + (e^{-4}4^1)/1! = 0.092$.

4-108. Let X denote the number of totes in the sample that exceed the moisture content. Then X is a binomial random variable with n = 30. We are to determine p.

If $P(X \ge 1) = 0.9$, then $P(X = 0) = 0.1$. Then $\binom{30}{0}(p)^0(1-p)^{30} = 0.1$, giving $30\ln(1-p)=\ln(0.1)$, which results in p = 0.0739.

4-109. Let t denote an interval of time in hours and let X denote the number of messages that arrive in time t. Then, X is a Poisson random variable with $\lambda = 10t$.

Then, P(X=0) = 0.9 and $e^{-10t} = 0.9$, resulting in t = 0.0105 hours = 0.63 seconds.

4-110. a) Let X denote the number of flaws in 50 panels. Then, X is a Poisson random variable with $\lambda = 50(0.02) = 1$. $P(X = 0) = e^{-1} = 0.3679$.

b) Let Y denote the number of flaws in one panel, then $P(Y \ge 1) = 1 - P(Y=0) = 1 - e^{-0.02} = 0.0198$. Let W denote the number of panels that need to be inspected before a flaw is found. Then W is a geometric random variable with p = 0.0198 and $E(W) = 1/0.0198 = 50.51$.

Mind-Expanding Exercises

4-112.

$$E(X) = [(a + (a+1) + \ldots + b](b - a + 1)$$

$$= \left[\sum_{i=1}^{b} i - \sum_{i=1}^{a-1} i\right] \Big/ (b - a + 1) = \left[\frac{b(b+1)}{2} - \frac{(a-1)a}{2}\right](b - a + 1)$$

$$= \left[\frac{(b^2 - a^2 + b + a)}{2}\right](b - a + 1) = \left[\frac{(b+a)(b-a+1)}{2}\right](b - a + 1)$$

$$= \frac{(b+a)}{2}$$

$$V(X) = \frac{\sum_{i=a}^{b}\left[i - \frac{b+a}{2}\right]^2}{b + a - 1} = \frac{\left[\sum_{i=a}^{b} i^2 - (b+a)\sum_{i=a}^{b} i + \frac{(b+a-1)(b+a)^2}{4}\right]}{b + a - 1}$$

4-113. Let X denote the number of nonconforming products in the sample. Then, X is approximately binomial with p = 0.01 and n is to be determined.
If $P(X \geq 1) \geq 0.90$, then $P(X = 0) \leq 0.10$.
Now, P(X = 0) = $\binom{n}{0} p^0 (1-p)^n = (1-p)^n$. Consequently, $(1-p)^n \leq 0.10$, and
$n \leq \frac{\ln 0.10}{\ln(1-p)} = 229.11$. Therefore, n = 230 is required.

4-114. If the lot size is small, 10% of the lot might be insufficient to detect nonconforming product. For example, if the lot size is 10, then a sample of size one has a probability of only 0.2 of detecting a nonconforming product in a lot that is 20% nonconforming.
If the lot size is large, 10% of the lot might be a larger sample size than is practical or necessary. For example, if the lot size is 5000, then a sample of 500 is required. Furthermore, the binomial approximation to the hypergeometric distribution can be used to show the following. If 5% of the lot of size 5000 is nonconforming, then the probability of zero nonconforming product in the sample is approximately 7×10^{-12} . Using a sample of 100, the same probability is still only 0.0059. The sample of size 500 might be much larger than is needed.

4-115. Let X denote the number of panels with flaws. Then, X is a binomial random variable with n =100 and p is the probability of one or more flaws in a panel. That is, p = $1 - e^{-0.1}$ = 0.095.

$$P(X < 5) = P(X \leq 4) = P(X=0) + P(X=1) + P(X=2) + P(X=3) + P(X=4)$$

$$= \binom{100}{0} p^0 (1-p)^{100} + \binom{100}{1} p^1 (1-p)^{99} + \binom{100}{2} p^2 (1-p)^{98}$$

$$+ \binom{100}{3} p^3 (1-p)^{97} + \binom{100}{4} p^4 (1-p)^{96}$$

$$= 0.034$$

4-116.

Let X denote the number of rolls produced.

Revenue at each demand				
	0	1000	2000	3000
$0 \le x \le 1000$	0.05x	0.3x	0.3x	0.3x
mean profit = 0.05x(0.3) + 0.3x(0.7) - 0.1x				
$1000 \le x \le 2000$	0.05x	0.3(1000) + 0.05(x-1000)	0.3x	0.3x
mean profit = 0.05x(0.3) + [0.3(1000) + 0.05(x-1000)](0.2) + 0.3x(0.5) - 0.1x				
$2000 \le x \le 3000$	0.05x	0.3(1000) + 0.05(x-1000)	0.3(2000) + 0.05(x-2000)	0.3x
mean profit = 0.05x(0.3) + [0.3(1000)+0.05(x-1000)](0.2) + [0.3(2000) + 0.05(x-2000)](0.3) + 0.3x(0.2) - 0.1x				
$3000 \le x$	0.05x	0.3(1000) + 0.05(x-1000)	0.3(2000) + 0.05(x-2000)	0.3(3000)+ 0.05(x-3000)
mean profit = 0.05x(0.3) + [0.3(1000)+0.05(x-1000)](0.2) + [0.3(2000)+0.05(x-2000)]0.3 + [0.3(3000)+0.05(x-3000)]0.2 - 0.1x				

	Profit	Max. profit
$0 \le x \le 1000$	0.125 x	\$ 125 at x = 1000
$1000 \le x \le 2000$	0.075 x + 50	\$ 200 at x = 2000
$2000 \le x \le 3000$	200	\$200 for all $2000 \le x \le 3000$
$3000 \le x$	-0.05 x + 350	

Therefore, the profit is maximized for any x from [2000, 3000]

4-117. Let X denote the number of acceptable components. Then, X has a binomial distribution with p = 0.98 and n is to be determined such that $P(X \ge 100) \ge 0.95$.

n	$P(X \ge 100)$
102	0.666
103	0.848
104	0.942
105	0.981

Therefore, 105 components are needed.

CHAPTER 5

Section 5-2

5-1. a) $P(1 < X) = \int_1^\infty e^{-x}dx = (-e^{-x})\Big|_1^\infty = e^{-1} = 0.3679$

b) $P(1 < X < 2.5) = \int_1^{2.5} e^{-x}dx = (-e^{-x})\Big|_1^{2.5} = e^{-1} - e^{-2.5} = 0.2858$

c) $P(X = 3) = \int_3^3 e^{-x}dx = 0$

d) $P(X < 4) = \int_0^4 e^{-x}dx = (-e^{-x})\Big|_0^4 = 1 - e^{-4} = 0.9817$

e) $P(3 \le X) = \int_3^\infty e^{-x}dx = (-e^{-x})\Big|_3^\infty = e^{-3} = 0.0498$

5-2. a) $P(x < X) = \int_x^\infty e^{-x}dx = (-e^{-x})\Big|_x^\infty = e^{-x} = 0.10$.

Then, $x = -\ln(0.10) = 2.3$

b) $P(X \le x) = \int_0^x e^{-x}dx = (-e^{-x})\Big|_0^x = 1 - e^{-x} = 0.10$.

Then, $x = -\ln(0.9) = 0.1054$

5-3. a) $P(1 < X) = \int_4^\infty e^{-(x-4)}dx = -e^{-(x-4)}\Big|_4^\infty = 1$, because $f_X(x) = 0$ for $x < 4$. This can also be obtained from the fact that $f_X(x)$ is a probability density function for $4 < x$.

b) $P(2 \le X \le 5) = \int_4^5 e^{-(x-4)}dx = -e^{-(x-4)}\Big|_4^5 = 1 - e^{-1} = 0.6321$

c) $P(5 < X) = 1 - P(X \le 5)$. From part b., $P(X \le 5) = 0.6321$. Therefore, $P(5 < X) = 0.3679$.

d) $P(8 < X < 12) = \int_8^{12} e^{-(x-4)}dx = -e^{-(x-4)}\Big|_8^{12} = e^{-4} - e^{-8} = 0.0180$

e) $P(X < x) = \int_4^x e^{-(x-4)}dx = -e^{-(x-4)}\Big|_4^x = 1 - e^{-(x-4)} = 0.90$.

Then, $x = 4 - \ln(0.10) = 6.303$

5-4. a) $P(0 < X) = 0.5$, by symmetry.

b) $P(0.5 < X) = \int_{0.5}^1 1.5x^2dx = 0.5x^3\Big|_{0.5}^1 = 0.5 - 0.0625 = 0.4375$

c) $P(-0.5 \le X \le 0.5) = \int_{-0.5}^{0.5} 1.5x^2dx = 0.5x^3\Big|_{-0.5}^{0.5} = 0.125$

d) $P(X < -2) = 0$

e) $P(X < 0 \text{ or } X > -0.5) = 1$

f) $P(x < X) = \int_x^1 1.5x^2dx = 0.5x^3\Big|_x^1 = 0.5 - 0.5x^3 = 0.05$

Then, $x = 0.9655$

5-5. a) $P(X > 3000) = \int_{3000}^{\infty} \frac{e^{\frac{-x}{1000}}}{1000} dx = -e^{\frac{-x}{1000}} \Big|_{3000}^{\infty} = e^{-3} = 0.05$

b) $P(1000 < X < 2000) = \int_{1000}^{2000} \frac{e^{\frac{-x}{1000}}}{1000} dx = -e^{\frac{-x}{1000}} \Big|_{1000}^{2000} = e^{-1} - e^{-2} = 0.233$

c) $P(X < 1000) = \int_{0}^{1000} \frac{e^{\frac{-x}{1000}}}{1000} dx = -e^{\frac{-x}{1000}} \Big|_{0}^{1000} = 1 - e^{-1} = 0.632$

d) $P(X < x) = \int_{0}^{x} \frac{e^{\frac{-x}{1000}}}{1000} dx = -e^{\frac{-x}{1000}} \Big|_{0}^{x} = 1 - e^{-x/1000} = 0.10$.

Then, $e^{-x/1000} = 0.9$, and $x = -1000 \ln 0.9 = 105.36$.

5-6. a) $P(X > 50) = \int_{50}^{50.25} 2.0 dx = 2x \Big|_{50}^{50.25} = 1/2$

b) $P(X > x) = 0.90 = \int_{x}^{50.25} 2.0 dx = 2x \Big|_{x}^{50.25} = 100.5 - 2x$

Then, $2x = 99.6$ and $x = 49.8$.

c) $E(X) = \int_{49.75}^{50.25} x 2.0 dx = x^2 \Big|_{49.75}^{50.25} = 2525.06 - 2475.06 = 50$

$V(X) = \int_{49.75}^{50.25} (x-50)^2 2.0 dx = \frac{2(x-50)^3}{3} \Big|_{49.75}^{50.25} = 0.01042 - (-0.01042) = 0.02084$

5-7. a) $P(X < 74.8) = \int_{74.6}^{74.8} 1.25 dx = 1.25x \Big|_{74.6}^{74.8} = 0.25$

b) $P(X < 74.8 \text{ or } X > 75.2) = P(X < 74.8) + P(X > 75.2)$ because the two events are mutually exclusive. The result is $0.25 + 0.25 = 0.50$.

c) $P(74.7 < X < 75.3) = \int_{74.7}^{75.3} 1.25 dx = 1.25x \Big|_{74.7}^{75.3} = 1.25(0.6) = 0.750$

d) $E(X) = \int_{74.6}^{75.4} x 1.25 dx = \frac{1.25x^2}{2} \Big|_{74.6}^{75.4} = 3553.225 - 3478.225 = 75$

$V(X) = \int_{74.6}^{75.4} (x-75)^2 1.25 dx = \frac{1.25(x-75)^3}{3} \Big|_{74.6}^{75.4} = 0.0267 - (0.0267) = 0.0533$

5-8. a) $P(X < 2.25 \text{ or } X > 2.75) = P(X < 2.25) + P(X > 2.75)$ because the two events are mutually exclusive. Then, $P(X < 2.25) = 0$ and

$P(X > 2.75) = \int_{2.75}^{2.8} 2 dx = 2(0.05) = 0.10$.

b) If the probability density function is centered at 2.5 meters, then $f_X(x) = 2$ for $2.25 < x < 2.75$ and all rods will meet specifications.

c) $E(X) = \int_{2.3}^{2.8} 2x dx = x^2 \Big|_{2.3}^{2.8} = 7.84 - 5.29 = 2.55$

$V(X) = \int_{2.3}^{2.8} (x-2.55)^2 2 dx = \frac{2(x-2.55)^3}{3} \Big|_{2.3}^{2.8} = 0.0104 - (-0.01041) = 0.02082$

5-9. Because the integral $\int_{x_1}^{x_2} f(x)dx$ is not changed whether or not any of the endpoints x_1 and x_2 are included in the integral, all the probabilities listed are equal.

Section 5-3

5-10. a) $P(X<2.8) = P(X \le 2.8)$ because X is a continuous random variable.
Then, $P(X<2.8) = F(2.8) = 0.2(2.8) = 0.56$.
b) $P(X > 1.5) = 1 - P(X \le 1.5) = 1 - 0.2(1.5) = 0.7$
c) $P(X < -2) = F_X(-2) = 0$
d) $P(X > 6) = 1 - F_X(6) = 0$

5-11. a) $P(X < 1.8) = P(X \le 1.8) = F_X(1.8)$ because X is a continuous random variable. Then,
$F_X(1.8) = 0.25(1.8) + 0.5 = 0.95$
b) $P(X > -1.5) = 1 - P(X \le -1.5) = 1/8$
c) $P(X < -2) = 0$
d) $P(-1 < X < 1) = P(-1 < X \le 1) = F_X(1) - F_X(-1) = \frac{3}{4} - \frac{1}{4} = \frac{1}{2}$

5-12. Now, $f_X(x) = e^{-x}$ for $0 < x$ and $F_X(x) = \int_0^x e^{-x}dx = -e^{-x}\Big|_0^x = 1 - e^{-x}$
for $0 < x$. Then, $F_X(x) = \begin{cases} 0, & x \le 0 \\ 1 - e^{-x}, & x > 0 \end{cases}$

5-13. Now, $f_X(x) = 1.5x^2$ for $-1 < x < 1$ and $F_X(x) = \int_{-1}^x 1.5x^2dx = 0.5x^3\Big|_{-1}^x = 0.5x^3 + 0.5$ for $-1 < x < 1$.

Then, $F(x) = \begin{cases} 0, & x < -1 \\ 0.5x^3 + 0.5, & -1 \le x < 1 \\ 1, & 1 \le x \end{cases}$

5-14. Now, $f(x) = \frac{e^{-x/1000}}{1000}$ for $x > 0$ and $F(x) = \int_0^x \frac{e^{-x/1000}}{1000}dx = -e^{-x/1000}\Big|_0^x$
for $0 < x$.
$F(x) = \begin{cases} 0, & x \le 0 \\ 1 - e^{-x/1000}, & x > 0 \end{cases}$
$P(X > 3000) = 1 - P(X \le 3000) = 1 - F(3000) = e^{-3}$

5-15. Now, f(x) = 1.25 for 74.6 < x < 75.4 and $F(x) = \int_{74.6}^{x} 1.25dx = 1.25x - 93.25$

for 74.6 < x < 75.4. Then,

$$F(x) = \begin{cases} 0, & x < 74.6 \\ 1.25x - 93.25, & 74.6 \le x < 75.4 \\ 1, & 75.4 \le x \end{cases}$$

$P(X > 75) = 1 - P(X \le 75) = 1 - F(75) = 1 - 0.5 = 0.5$ because X is a continuous random variable.

5-16. $f(x) = 2e^{-2x}, \quad x > 0$

5-17. $f(x) = \begin{cases} 0.2, & 0 < x < 4 \\ 0.4, & 4 \le x < 9 \end{cases}$

5-18. $f(x) = \begin{cases} 0.25, & -2 < x < 1 \\ 0.5, & 1 \le x < 1.5 \end{cases}$

5-19. $F(x) = \int_0^x 0.5xdx = \left.\frac{0.5x^2}{2}\right|_0^x = 0.25x^2$ for 0 < x < 2. Then,

$$F(x) = \begin{cases} 0, & x < 0 \\ 0.25x^2, & 0 \le x < 2 \\ 1, & 2 \le x \end{cases}$$

Section 5-4

5-20. $E(X) = \int_0^4 0.25xdx = 0.25\left.\frac{x^2}{2}\right|_0^4 = 2$

$$V(X) = \int_0^4 0.25(x-2)^2 dx = 0.25\left.\frac{(x-2)^3}{3}\right|_0^4 = \frac{2}{3} + \frac{2}{3} = \frac{4}{3}$$

5-21. $E(X) = \int_0^4 0.125x^2dx = 0.125\left.\frac{x^3}{3}\right|_0^4 = 8/3$

$$V(X) = \int_0^4 0.125x(x - \tfrac{8}{3})^2 dx = 0.125\int_0^4 (x^3 - \tfrac{16}{3}x^2 + \tfrac{64}{3}x)dx$$

$$= 0.125\left(\frac{x^4}{4} - \frac{16}{3}\frac{x^3}{3} + \frac{64}{3}\frac{1}{2}x^2\right)\Bigg|_0^4 = 15.11$$

5-22. $E(X) = \int_{-1}^{1} 1.5x^3 dx = 1.5 \left.\frac{x^4}{4}\right|_{-1}^{1} = \frac{3}{4}$

$$V(X) = \int_{-1}^{1} 1.5x^2 (x - \frac{3}{4})^2 dx = 1.5 \int_{-1}^{1} \left(x^4 - \frac{3x^3}{2} + \frac{9x^2}{16} \right) dx$$

$$= 1.5 \left(\frac{x^5}{5} - \frac{3x^4}{8} + \frac{3x^3}{16} \right) \Bigg|_{-1}^{1} = \frac{93}{80} = 1.1625$$

5-23. $E(X) = \int_{49.75}^{50.25} 2x dx = x^2 \Big|_{49.75}^{50.25} = 50$

$$V(X) = \int_{49.75}^{50.25} 2(x - 50)^2 dx = 2 \int_{49.75}^{50.25} (x^2 - 100x + 2500) dx$$

$$= 2\left(\frac{x^3}{3} - 100\frac{x^2}{2} + 2500x\right)\Bigg|_{49.75}^{50.25}$$

$$= 0.0208$$

5-24. $E(X) = \int_{100}^{120} x \frac{600}{x^2} dx = 600 \ln x \Big|_{100}^{120} = 109.39$

$$V(X) = \int_{100}^{120} (x - 109.39)^2 \frac{600}{x^2} dx = 600 \int_{100}^{120} 1 - \frac{2(109.39)}{x} + \frac{(109.39)^2}{x^2} dx$$

$$= 600(x - 218.78 \ln x - 109.39^2 x^{-1}) \Big|_{100}^{120} = 33.19$$

5-25. $E(X) = \int_{1}^{\infty} x 2x^{-3} dx = -2x^{-1} \Big|_{1}^{\infty} = 2$

5-26. a) $E(X) = \int_{5}^{\infty} x 10e^{-(x-5)} dx$.

Using integration by parts with $u = x$ and $dv = 10e^{-(x-5)} dx$, we obtain

$$E(X) = -xe^{-10(x-5)} \Big|_{5}^{\infty} + \int_{5}^{\infty} e^{-10(x-5)} dx = 5 - \frac{e^{-10(x-5)}}{10} \Bigg|_{5}^{\infty} = 5.1$$

Now, $V(X) = \int_{5}^{\infty} (x - 5.1)^2 10e^{-10(x-5)} dx$. Using the integration by parts with $u = (x - 5.1)^2$ and $dv = 10e^{-10(x-5)}$, we obtain $V(X) = -(x - 5.1)^2 e^{-10(x-5)} \Big|_{5}^{\infty} + 2\int_{5}^{\infty} (x - 5.1)e^{-10(x-5)} dx$. From the definition of E(X) the integral above is recognized to equal 0.

Therefore, $V(X) = (5 - 5.1)^2 = 0.01$.

b) $P(X > 5.1) = \int_{5.1}^{\infty} 10e^{-(x-5)} dx = -e^{-10(x-5)} \Big|_{5.1}^{\infty} = e^{-10(5.1-5)} = 0.3679$

5-27. a)

$$E(X) = \int_{1200}^{1210} x0.1dx = 0.05x^2 \Big|_{1200}^{1210} = 1205$$

$$V(X) = \int_{1200}^{1210} (x-1205)^2 0.1dx = 0.1\frac{(x-1205)^3}{3}\Big|_{1200}^{1210} = 8.333$$

Therefore, $\sigma_X = \sqrt{V(X)} = 2.887$

b) Clearly, centering the process at the center of the specifications results in the greatest proportion of cables within specifications.

Section 5-5

5-28. a) E(X) = (5.5+1.5)/2 = 3.5,

$$V(X) = \frac{(5.5-1.5)^2}{12} = 3/4, \text{ and } \sigma_X = \sqrt{\frac{3}{4}} = 0.866 .$$

b) $P(X < 2.5) = \int_{1.5}^{2.5} 0.25dx = 0.25x\Big|_{1.5}^{2.5} = 0.25$

5-29. a) E(X) = (-1+1)/2 = 0,

$$V(X) = \frac{(1-(-1))^2}{12} = 1/3, \text{ and } \sigma_X = 0.577$$

b) $P(-x < X < x) = \int_{-x}^{x} \frac{1}{2}dt = 0.5t\Big|_{-x}^{x} = 0.5(2x) = x$

Therefore, x should equal 0.90.

5-30. a) f(x)= 2.0 for 49.75 < x < 50.25. E(X) = (50.25 + 49.75)/2 = 50.0,

$$V(X) = \frac{(50.25-49.75)^2}{12} = 0.0208, \text{ and } \sigma_X = 0.144 .$$

b) $F(x) = \int_{49.75}^{x} 2.0dt$ for 49.75 < x < 50.25. Therefore,

$$F(x) = \begin{cases} 0, & x < 49.75 \\ 2x - 99.5, & 49.75 \le x < 50.25 \\ 1, & 50.25 \le x \end{cases}$$

c) $P(X < 50.1) = F(50.1) = 2(50.1) - 99.5 = 0.7$

5-31. a) The distribution of X is f(x) = 10 for 0.95 < x < 1.05. Now,

$$F(x) = \begin{cases} 0, & x < 0.95 \\ 10x - 9.5, & 0.95 \le x < 1.05 \\ 1, & 1.05 \le x \end{cases}$$

b) $P(X > 1.02) = 1 - P(X \le 1.02) = 1 - F(1.02) = 0.3$

c) If P(X > x)=0.90, then 1 – F(X) = 0.90 and F(X) = 0.10. Therefore, 10x - 9.5 = 0.10 and x = 0.96.

d) E(X) = (1.05 + 0.95)/2 = 1.00 and V(X) = $\frac{(1.05-0.95)^2}{12} = 0.00083$

5-32. a) The distribution of X is f(x) = 100 for 0.2050 < x < 0.2150. Therefore,

$$F(x) = \begin{cases} 0, & x < 0.2050 \\ 100x - 20.50, & 0.2050 \le x < 0.2150 \\ 1, & 0.2150 \le x \end{cases}$$

b) $P(X > 0.2125) = 1 - F(0.2125) = 1 - [100(0.2125) - 20.50] = 0.25$

c) If P(X > x)=0.10, then 1 – F(X) = 0.10 and F(X) = 0.90. Therefore, 100x - 20.50 = 0.90 and x = 0.2140.

d) E(X) = (0.2050 + 0.2150)/2 = 0.2100 and V(X) = $\frac{(0.2150 - 0.2050)^2}{12} = 8.33 \times 10^{-6}$

5-33. a) $P(X > 35) = \int_{35}^{40} 0.1dx = 0.1x\Big|_{35}^{40} = 0.5$

b) P(X > x)=0.90 and $P(X > x) = \int_{x}^{40} 0.1dt = 0.1(40 - x)$.

Now, 0.1(40-x) = 0.90 and x = 31

c) E(X) = (30 + 40)/2 = 35 and V(X) = $\frac{(40 - 30)^2}{12} = 8.33$

Section 5-6

5-34. a) 0.90658
b) 0.99865
c) 1 – 0.92647 = 0.07353
d) P(Z > –2.15) = p(Z < 2.15) = 0.98422
e) P(–2.34 < Z < 1.76) = P(Z<1.76) – P(Z > 2.34) = 0.95116

5-35. a) P(–1 < Z < 1) = P(Z < 1) – P(Z > 1)
= 0.84134 – (1 – 0.84134)
= 0.68268
b) P(–2 < Z < 2) = P(Z < 2) – [1 – P(Z < 2)]
= 0.9545
c) P(–3 < Z < 3) = P(Z < 3) – [1 – P(Z < 3)]
= 0.9973
d) P(Z > 3) = 1 – P(Z < 3)
= 0.00135
e) P(0 < Z < 1) = P(Z < 1) – P(Z < 0)
= 0.84134 – 0.5
= 0.34134

5-36. a) P(Z < 1.28) = 0.90
b) P(Z < 0) = 0.5
c) If P(Z > z) = 0.1, then P(Z < z) = 0.90 and z = 1.28
d) If P(Z > z) = 0.9, then P(Z < z) = 0.10 and z = –1.28
e) P(–1.24 < Z < z) = P(Z < z) – P(Z < –1.24)
= P(Z < z) – 0.10749.
Therefore, P(Z < z) = 0.8 + 0.10749 = 0.90749 and z = 1.33

5-37. a) Because of the symmetry of the normal distribution, the area in each tail of the distribution must equal 0.025. Therefore the value in Table II that corresponds to 0.975 is 1.96. Thus, z = 1.96.
b) Find the value in Table II corresponding to 0.995. z = 2.58.
c) Find the value in Table II corresponding to 0.84. z = 1.0
d) Find the value in Table II corresponding to 0.99865. z = 3.0.

5-38. a) $P(X < 13) = P(Z < (13-10)/2)$
$= P(Z < 1.5)$
$= 0.93319$

b) $P(X > 9) = 1 - P(X < 9)$
$= 1 - P(Z < (9-10)/2)$
$= 1 - P(Z < -0.5)$
$= 1 - [1- P(Z < 0.5)]$
$= P(Z < 0.5)$
$= 0.69146.$

c) $P(6 < X < 14) = P\left(\frac{6-10}{2} < Z < \frac{14-10}{2}\right)$
$= P(-2 < Z < 2)$
$= P(Z < 2) - P(Z < -2)]$
$= 0.9545.$

d) $P(2 < X < 4) = P\left(\frac{2-10}{2} < Z < \frac{4-10}{2}\right)$
$= P(-4 < Z < -3)$
$= P(Z < -3) - P(Z < -4)$
$= 0.00135$

e) $P(-2 < X < 8) = P(X < 8) - P(X < -2)$
$= P\left(Z < \frac{8-10}{2}\right) - P\left(Z < \frac{-2-10}{2}\right)$
$= P(Z < -1) - P(Z < -6)$
$= 0.15866.$

5-39. a) $P(X > x) = P\left(Z > \frac{x-10}{2}\right) = 0.5$. Therefore, $\frac{x-10}{2} = 0$ and $x = 10$.

b) $P(X > x) = P\left(Z > \frac{x-10}{2}\right)$
$= 1 - P\left(Z < \frac{x-10}{2}\right)$
$= 0.95.$

Therefore, $P\left(Z < \frac{x-10}{2}\right) = 0.05$ and $\frac{x-10}{2} = -1.64$. Consequently, $x = 6.72$.

c) $P(x < X < 10) = P\left(\frac{x-10}{2} < Z < 0\right)$
$= P(Z < 0) - P\left(Z < \frac{x-10}{2}\right)$
$= 0.5 - P\left(Z < \frac{x-10}{2}\right) = 0.2.$

Therefore, $P\left(Z < \frac{x-10}{2}\right) = 0.3$ and $\frac{x-10}{2} = -0.52$. Consequently, $x = 8.96$.

d) $P(10 - x < X < 10 + x) = P(-x/2 < Z < x/2)$
$= 0.95.$

Therefore, $x/2 = 1.96$ and $x = 3.92$

e) $P(10 - x < X < 10 + x) = P(-x/2 < Z < x/2)$
$= 0.99.$

Therefore, $x/2 = 2.58$ and $x = 5.16$

5-40. a) $P(X < 11) = P\left(Z < \frac{11-5}{4}\right)$

$= P(Z < 1.5)$
$= 0.93319$

b) $P(X > 0) = P(Z > (-5/4))$
$= P(Z > -1.25)$
$= 1 - P(Z < -1.25)$
$= 0.89435$

c) $P(3 < X < 7) = P\left(\frac{3-5}{4} < Z < \frac{7-5}{4}\right)$

$= P(-0.5 < Z < 0.5)$
$= P(Z < 0.5) - P(Z < -0.5)$
$= 0.38292$

d) $P(-2 < X < 9) = P\left(\frac{-2-5}{4} < Z < \frac{9-5}{4}\right)$

$= P(-1.75 < Z < 1)$
$= P(Z < 1) - P(Z < -1.75)]$
$= 0.80128$

e) $P(2 < X < 8) = P(-0.75 < Z < 0.75)$
$= P(Z < 0.75) - P(Z < -0.75)$
$= 0.54674$

5-41. a) $P(X > x) = P\left(Z > \frac{x-5}{4}\right) = 0.5.$

Therefore, x = 5.

b) $P(X > x) = P\left(Z > \frac{x-5}{4}\right) = 0.95.$

Therefore, $P\left(Z < \frac{x-5}{4}\right) = 0.05$

Therefore, $\frac{x-5}{4} = -1.64$, and $x = -1.56$.

c) $P(x < X < 9) = P\left(\frac{x-5}{4} < Z < 1\right) = 0.2.$

Therefore, $P(Z < 1) - P(Z < \frac{x-5}{4}) = 0.2$ where $P(Z < 1) = 0.84134$.

Thus $P(Z < \frac{x-5}{4}) = 0.64134$. Consequently, $\frac{x-5}{4} = 0.36$ and $x = 6.44$.

d) $P(3 < X < x) = P\left(\frac{3-5}{4} < Z < \frac{x-5}{4}\right) = 0.95.$

Therefore, $P\left(Z < \frac{x-5}{4}\right) - P(Z < -0.5) = 0.95$ and $P\left(Z < \frac{x-5}{4}\right) - 0.30854 = 0.95$. Consequently,

$P\left(Z < \frac{x-5}{4}\right) = 1.20854$. Because a probability can not be greater than one, there is no solution for x. In fact, $P(3 < X) = P(-0.5 < Z) = 0.69146$. Therefore, even if x is set to infinity the probability requested cannot equal 0.95.

e) $P(5 - x < X < 5 + x) = P\left(\frac{5-x-5}{4} < Z < \frac{5+x-5}{4}\right)$

$= P\left(\frac{-x}{4} < Z < \frac{x}{4}\right) = 0.99$

Therefore, $x/4 = 2.58$ and $x = 11.32$.

5-42. a) $P(X < 6250) = P\left(Z < \frac{6250 - 6000}{100}\right)$

$= P(Z < 2.5)$
$= 0.99379$

b) $P(5800 < X < 5900) = P(-2 < Z < -1)$
$= P(Z < -1) - P(Z < -2)$
$= 0.13591$

c) $P(X > x) = P\left(Z > \frac{x - 6000}{100}\right) = 0.95.$

Therefore, $\frac{x - 6000}{100} = -1.64$ and x = 5836.

5-43. a) $P(X < 40) = P\left(Z < \frac{40 - 35}{2}\right)$

$= P(Z < 2.5)$
$= 0.99379$

b) $P(X < 30) = P\left(Z < \frac{30 - 35}{2}\right)$

$= P(Z < -2.5)$
$= 0.00621$

5-44. a) $P(X > 0.62) = P\left(Z > \frac{0.62 - 0.5}{0.05}\right)$

$= P(Z > 3)$
$= 1 - P(Z < 3)$
$= 0.00135$

b) $P(0.47 < X < 0.63) = P(-0.6 < Z < 2.6)$
$= P(Z < 2.6) - P(Z < -0.6)$
$= 0.99534 - 0.27425$
$= 0.72109$

c) $P(X < x) = P\left(Z < \frac{x - 0.5}{0.05}\right) = 0.90.$

Therefore, $\frac{x - 0.5}{0.05} = 1.28$ and x = 0.564.

5-45. a) $P(X < 12) = P(Z < \frac{12 - 12.4}{0.1}) = P(Z < -4) \cong 0$

b) $P(X < 12.1) = P\left(Z < \frac{12.1 - 12.4}{0.1}\right)$

$= P(Z < -3)$
$= 0.00135$

and

$P(X > 12.6) = P\left(Z > \frac{12.6 - 12.4}{0.1}\right)$

$= P(Z > 2)$
$= 0.02275.$

Therefore, the proportion of cans scrapped is 0.00135 + 0.02275 = 0.0241.

c) $P(12.4 - x < X < 12.4 + x) = 0.99.$

Therefore, $P\left(-\frac{x}{0.1} < Z < \frac{x}{0.1}\right) = 0.99$

Consequently, $P\left(Z < \frac{x}{0.1}\right) = 0.995$ and x = 0.1(2.58) = 0.258.

The limits are (12.142, 12.658).

5-46. a) If P(X > 12) = 0.999, then $P\left(Z > \frac{12-\mu}{0.1}\right) = 0.999$.

Therefore, $\frac{12-\mu}{0.1} = -3.09$ and $\mu = 12.309$.

b) If P(X > 12) = 0.999, then $P\left(Z > \frac{12-\mu}{0.05}\right) = 0.999$.

Therefore, $\frac{12-\mu}{0.05} = -3.09$ and $\mu = 12.1545$.

5-47. a) $P(X > 0.5) = P\left(Z > \frac{0.5-0.4}{0.05}\right)$

$= P(Z > 2)$
$= 1 - 0.97725$
$= 0.02275$

b) $P(0.4 < X < 0.5) = P(0 < Z < 2)$
$= P(Z < 2) - P(Z < 0)$
$= 0.47725$

c) P(X > x) = 0.90, then $P\left(Z > \frac{x-0.4}{0.05}\right) = 0.90$.

Therefore, $\frac{x-0.4}{0.05} = -1.28$ and x = 0.336.

5-48. a) $P(90.3 < X) = P\left(\frac{90.3-90.2}{0.1} < Z\right)$

$= P(1 < Z)$
$= P(Z > 1)$
$= 1 - P(Z < 1)$
$= 1 - 0.84134$
$= 0.15866.$

$P(X < 89.7) = P\left(Z < \frac{89.7-90.2}{0.1}\right)$

$= P(Z < -5)$
$\cong 0$.

Therefore, the answer is 0.15866.

b) The process mean should be set at the center of the specifications; that is, at $\mu = 90.0$.

c) $P(89.7 < X < 90.3) = P\left(\frac{89.7-90}{0.1} < Z < \frac{90.3-90}{0.1}\right)$

$= P(-3 < Z < 3) = 0.9973.$

5-49. a) $P(89.7 < X < 90.3) = P\left(\frac{89.7-90}{0.1} < Z < \frac{90.3-90}{0.1}\right)$

$= P(-3 < Z < 3)$
$= 0.9973.$

b) P(X < 89.8) + P(X > 90.2)

$= P\left(\frac{X-90}{0.1} < \frac{89.8-90}{0.1}\right) + P\left(\frac{X-90}{0.1} > \frac{90.2-90}{0.1}\right)$

$= P(Z < -2) + P(Z > 2)$

$= P(Z < -2) + (1 - P(Z < 2))$

$= 0.0455$

5-50. a) $P(50 < X < 80) = P\left(\frac{50-100}{20} < Z < \frac{80-100}{20}\right)$

$= P(-2.5 < Z < -1)$
$= P(Z < -1) - P(Z < -2.5)$
$= 0.15245.$

b) $P(X > x) = 0.10$. Therefore, $P\left(Z > \frac{x-100}{20}\right) = 0.10$ and $\frac{x-100}{20} = 1.28$. Consequently, $x = 125.6$.

5-51. a) $P(X < 5000) = P\left(Z < \frac{5000-7000}{600}\right)$

$= P(Z < -3.33)$
$= 0.00043.$

b) $P(X > x) = 0.95$. Therefore, $P\left(Z > \frac{x-7000}{600}\right) = 0.95$ and $\frac{x-7000}{600} = -1.64$. Consequently, $x = 6016$.

5-52. a) $P(X > 0.0026) = P\left(Z > \frac{0.0026-0.002}{0.0004}\right)$

$= P(Z > 1.5)$
$= 0.06681.$

b) $P(0.0014 < X < 0.0026) = P(-1.5 < Z < 1.5)$
$= 0.86638.$

c) $P(0.0014 < X < 0.0026) = P\left(\frac{0.0014-0.002}{\sigma} < Z < \frac{0.0026-0.002}{\sigma}\right)$

$= P\left(\frac{-0.0006}{\sigma} < Z < \frac{0.0006}{\sigma}\right).$

Therefore, $P\left(Z < \frac{0.0006}{\sigma}\right) = 0.9975$. Consequently, $\frac{0.0006}{\sigma} = 2.81$ and $\sigma = 0.000214$.

5-53. a) $P(X > 13) = P\left(Z > \frac{13-12}{0.5}\right)$

$= P(Z > 2)$
$= 0.02275$

b) If $P(X < 13) = 0.999$, then $P\left(Z < \frac{13-12}{\sigma}\right) = 0.999$.

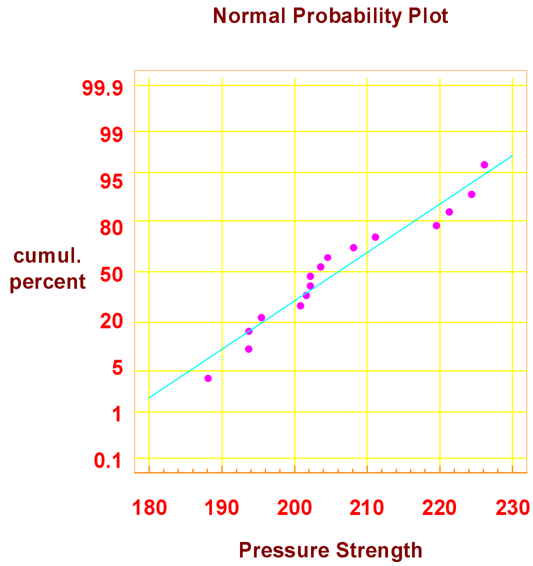
Therefore, $1/\sigma = 3.09$ and $\sigma = 1/3.09 = 0.324$.

c) If $P(X < 13) = 0.999$, then $P\left(Z < \frac{13-\mu}{0.5}\right) = 0.999$.

Therefore, $\frac{13-\mu}{0.5} = 3.09$ and $\mu = 11.455$

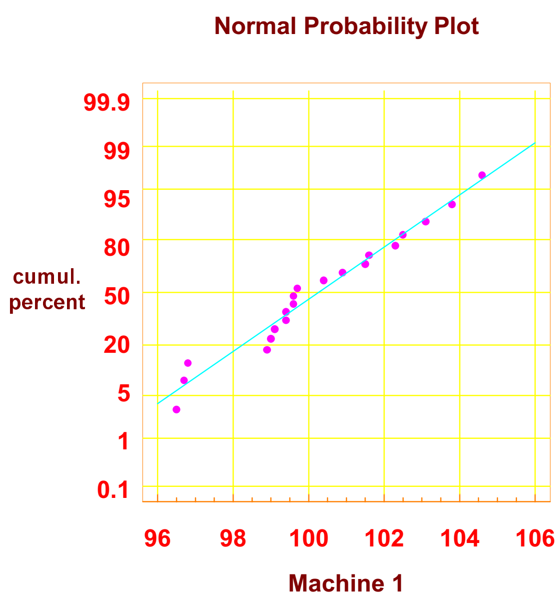
Section 5-7

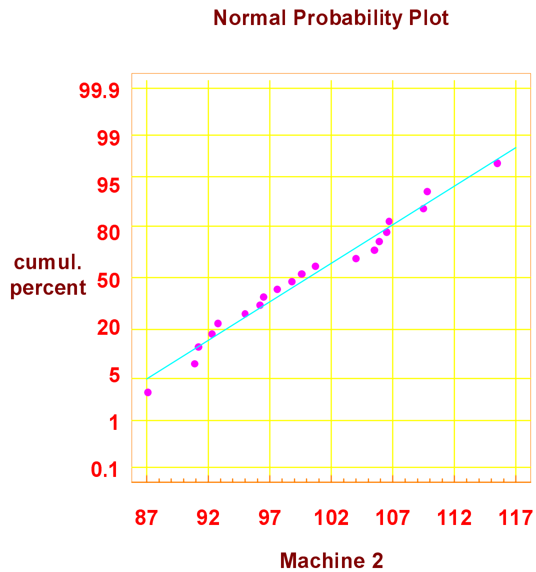
5-54.



The assumption of normality appears reasonable. This is evident by the fact that the data falls along a straight line.

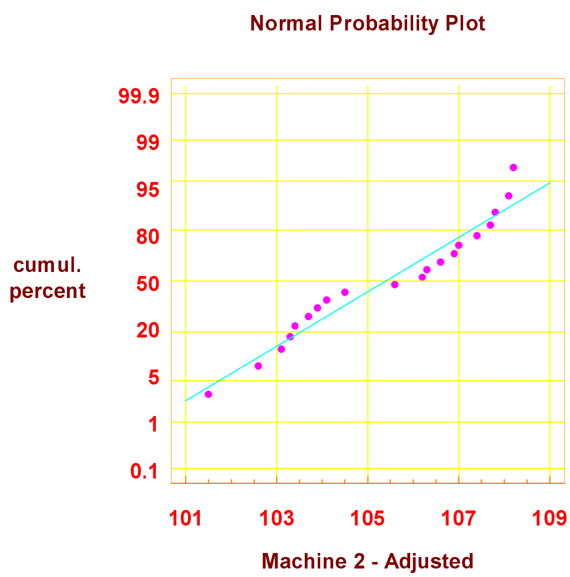
5-55.





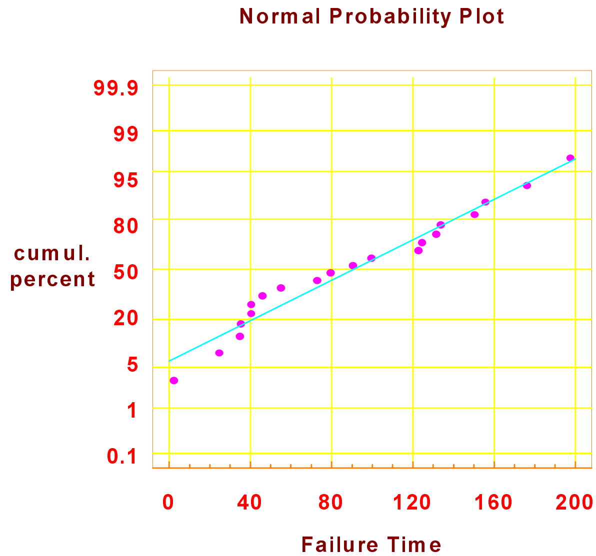
Yes, however, machine 2 has greater variability. Note the difference in horizontal scales.

5-56.



The variability between parts is greatly reduced. Now, machine 2 has lower part-to-part variability than machine 1. Normality still looks reasonable.

5-57.



The assumption of normality appears to be reasonable. This is evident by the fact that the data fall along a straight line.

Section 5-8

5-58. a) $E(X) = 200(0.4) = 80$, $V(X) = 200(0.4)(0.6) = 48$ and $\sigma_X = \sqrt{48}$.

Then, $P(X \le 70) \cong P\left(Z \le \frac{70-80}{\sqrt{48}}\right) = P(Z \le 1.44) = 0.925$

b) $P(70 < X \le 90) \cong P\left(\frac{70-80}{\sqrt{48}} < Z \le \frac{90-80}{\sqrt{48}}\right) = P(-1.44 < Z \le 1.44)$

$= 0.925 - (1 - 0.925) = 0.85$

5-59. a) $P(X < 4) = \binom{100}{0}0.1^0(0.9)^{100} + \binom{100}{1}0.1^1(0.9)^{99} + \binom{100}{2}0.1^2 0.9^{98} + \binom{100}{3}0.1^3 0.9^{97}$

$= 0.0078$

b) $E(X) = 10$, $V(X) = 100(0.1)(0.9) = 9$ and $\sigma_X = 3$.

Then, $P(X < 4) \cong P(Z < \frac{4-10}{3}) = P(Z < -2) = 0.023$

c) $P(8 < X < 12) \cong P\left(\frac{8-10}{3} < Z < \frac{12-10}{3}\right) = P(-0.67 < Z < 0.67) = 0.497$

5-60. Let X denote the number of defective chips in the lot.
Then, $E(X) = 1000(0.02) = 20$, $V(X) = 1000(0.02)(0.98) = 19.6$.

a) $P(X > 25) \cong P\left(Z > \frac{25-20}{\sqrt{19.6}}\right) = P(Z > 1.13) = 1 - P(Z \le 1.13) = 0.129$

b) $P(20 < X < 30) \cong P(0 < Z < \frac{10}{\sqrt{19.6}}) = P(0 < Z < 2.26)$

$= P(Z \le 2.26) - P(Z < 0) = 0.98809 - 0.5 = 0.488$

5-61. Let X = number of defective inspected parts

E(X) = 100(0.08) = 8

V(X) = 100(0.08)(0.92) = 7.36

a) $P(X < 6) = P(X \le 5) = \sum_{i=0}^{5} \binom{100}{i} (0.08)^i (0.92)^{100-i} = 0.1799$

b) $P(X < 6) \cong P\left(Z < \frac{6-8}{\sqrt{7.36}}\right) = P(Z < -0.74) = 0.22965$

5-62. Let X denote the number of original components that fail during the useful life of the product. Then, X is a binomial random variable with p = 0.001 and n = 200. Also, E(X) = 200 (0.001) = 0.2 and V(X) = 200(0.001)(0.999) = 0.1998.

$P(X \ge 5) \cong P\left(Z \ge \frac{5-0.2}{\sqrt{0.1998}}\right) = P(Z \ge 24) = 1 - P(Z < 24) = 1 - 1 = 0$.

5-63. Let X denote the number of particles in 10 cm^2 of dust. Then, X is a Poisson random variable with $\lambda = 10(1000) = 10{,}000$. Also, E(X) = $\lambda = 10{,}000$ and V(X) = $\lambda^2 = 10^8$.

$P(X > 10{,}000) \cong P\left(Z > \frac{10{,}000 - 10{,}000}{\sqrt{10^8}}\right) = P(Z > 0) = 0.5$

5-64. E(X) = 50(0.1) = 5 and V(X) = 50(0.1)(0.9) = 4.5

a) $P(X \le 2) = P(X \le 2.5) \cong P\left(Z \le \frac{2.5-5}{\sqrt{4.5}}\right) = P(Z \le -1.18)$

$= 1 - P(Z \le 1.18) = 1 - 0.881 = 0.119$

b) $P(X \le 2) \cong P\left(Z \le \frac{2-2.5}{\sqrt{4.5}}\right) = P(Z \le -0.82) = 1 - P(Z \le 0.82)$

$= 1 - 0.794 = 0.206$

c) $P(X \le 2) = \binom{50}{0} 0.1^0 0.9^{50} + \binom{50}{1} 0.1^1 0.9^{49} + \binom{50}{2} 0.1^2 0.9^{48} = 0.118$

The probability computed using the continuity correction is closer.

d) $P(X < 10) = P(X \le 9.5) \cong P\left(Z \le \frac{9.5-5}{\sqrt{4.5}}\right) = P(Z \le 2.12) = 0.983$

5-65. E(X) = 50(0.1) = 5 and V(X) = 50(0.1)(0.9) = 4.5

a) $P(X \ge 2) = P(X \ge 1.5) \cong P\left(Z \le \frac{1.5-5}{\sqrt{4.5}}\right) = P(Z \ge -1.65) = 0.951$

b) $P(X \ge 2) \cong P\left(Z \ge \frac{2-5}{\sqrt{4.5}}\right) = P(Z \ge -1.414) = 0.921$

c) $P(X \ge 2) = 1 - P(X < 2) = 1 - \binom{50}{0} 0.1^0 0.9^{50} - \binom{50}{1} 0.1^1 0.9^{49} = 0.966$

The probability computed using the continuity correction is closer.

d) $P(X > 6) = P(X \ge 7) = P(X \ge 6.5) \cong P(Z \ge 0.707) = 0.24$

5-66. E(X) = 50(0.1) = 5 and V(X) = 50(0.1)(0.9) = 4.5

a)

$$P(2 \le X \le 5) = P(1.5 \le X \le 5.5) \cong P(\frac{1.5-5}{\sqrt{4.5}} \le Z \le \frac{5.5-5}{\sqrt{4.5}})$$
$$= P(-1.65 \le Z \le 0.236) = P(Z \le 0.236) - P(Z \le -1.65)$$
$$= 0.592 - (1 - 0.95053) = 0.543$$

b)

$$P(2 \le X \le 5) \cong P(\frac{2-5}{\sqrt{4.5}} \le Z \le \frac{5-5}{\sqrt{4.5}}) = P(-1.414 \le Z \le 0)$$
$$= 0.5 - P(Z \le -1.414)$$
$$= 0.5 - (1 - 0.921) = 0.421$$

The exact probability is 0.582

5-67. E(X) = 50(0.1) = 5 and V(X) = 50(0.1)(0.9) = 4.5

a)

$$P(X = 10) = P(9.5 \le X \le 10.5) \cong P(\frac{9.5-5}{\sqrt{4.5}} \le Z \le \frac{10.5-5}{\sqrt{4.5}})$$
$$= P(2.121 \le Z \le 2.593) = 0.012$$

b)

$$P(X = 5) = P(4.5 \le X \le 5.5) \cong P(\frac{4.5-5}{\sqrt{4.5}} \le Z \le \frac{5.5-5}{\sqrt{4.5}})$$
$$= P(-0.236 \le Z \le 0.236) = 0.182$$

5-68. Let X denote the number of defective chips in the lot. Then, X is binomial with n = 1000 and p = 0.02. Also,
E(X) = 1000(0.02) = 20 and
V(X) = 1000(0.02)(0.98) = 19.6.

a)

$$P(20 < X < 30) = P(21 \le X \le 29)$$
$$= P(20.5 \le X \le 29.5) \cong P(\frac{20.5-20}{\sqrt{19.6}} \le Z \le \frac{29.5-20}{\sqrt{19.6}})$$
$$= P(0.113 \le Z \le 2.15) = P(Z \le 2.15) - P(Z \le 0.11)$$
$$= 0.984 - 0.545 = 0.439$$

b)

$$P(X = 20) = P(19.5 \le X \le 20.5) \cong P(\frac{19.5-20}{\sqrt{19.6}} \le Z \le \frac{20.5-20}{\sqrt{19.6}})$$
$$= P(-0.113 \le Z \le 0.113) = 0.545 - (1 - 0.545) = 0.09$$

c) Because the area under the normal probability density along an interval of fixed length is greatest when the interval is centered at the mean of the normal distribution , x = 20.

Section 5-9

5-69. a) $P(X \le 0) = \int_0^0 \lambda e^{-\lambda x} dx = 0$

b) $P(X \ge 2) = \int_2^\infty 2e^{-2x} dx = -e^{-2x}\Big|_2^\infty = e^{-4} = 0.0183$

c) $P(X \le 1) = \int_0^1 2e^{-2x} dx = -e^{-2x}\Big|_0^1 = 1 - e^{-2} = 0.8647$

d) $P(1 < X < 2) = \int_1^2 2e^{-2x} dx = -e^{-2x}\Big|_1^2 = e^{-2} - e^{-4} = 0.1170$

e) $P(X \le x) = \int_0^x 2e^{-2t} dt = -e^{-2t}\Big|_0^x = 1 - e^{-2x} = 0.05$ and $x = 0.0256$

5-70. If E(X) = 10, then $\lambda = 0.1$.

a) $P(X > 10) = \int_{10}^\infty 0.1e^{-0.1x} dx = -e^{-0.1x}\Big|_{10}^\infty = e^{-1} = 0.3679$

b) $P(X > 20) = -e^{-0.1x}\Big|_{20}^\infty = e^{-2} = 0.1353$

c) $P(X > 30) = -e^{-0.1x}\Big|_{30}^\infty = e^{-3} = 0.0498$

d) $P(X < x) = \int_0^x 0.1e^{-0.1t} dt = -e^{-0.1t}\Big|_0^x = 1 - e^{-0.1x} = 0.95$ and $x = 29.96$.

5-71. Let X denote the time until the first count. Then, X is an exponential random variable with $\lambda = 2$ counts per minute.

a) $P(X > 0.5) = \int_{0.5}^\infty 2e^{-2x} dx = -e^{-2x}\Big|_{0.5}^\infty = e^{-1} = 0.3679$

b) $P(X < \frac{10}{60}) = \int_0^{1/6} 2e^{-2x} dx = -e^{-2x}\Big|_0^{1/6} = 1 - e^{-1/3} = 0.2835$

c) $P(1 < X < 2) = -e^{-2x}\Big|_1^2 = e^{-2} - e^{-4} = 0.1170$

5-72. a) $E(X) = 1/\lambda = 1/3 = 0.333$ minutes
b) $V(X) = 1/\lambda^2 = 1/3^2 = 0.111$, $\sigma = 0.3333$
c) $P(X < x) = \int_0^x 3e^{-3t} dt = -e^{-3t}\Big|_0^x = 1 - e^{-3x} = 0.95$, $x = 0.9986$

5-73. Let X denote the time until the first call. Then, X is exponential and $\lambda = \frac{1}{E(X)} = \frac{1}{15}$ calls/minute.

a) $P(X > 30) = \int_{30}^\infty \frac{1}{15} e^{-\frac{x}{15}} dx = -e^{-\frac{x}{15}}\Big|_{30}^\infty = e^{-2} = 0.1353$

b) The probability of at least one call in a 10-minute interval equals one minus the probability of zero calls in a 10-minute interval and that is P(X > 10).

$P(X > 10) = -e^{-\frac{x}{15}}\Big|_{10}^\infty = e^{-2/3} = 0.5134$.

Therefore, the answer is 1- 0.5134 = 0.4866. Alternatively, the requested probability equals P(X < 10) = 0.4866.

c) $P(5 < X < 10) = -e^{-\frac{x}{15}}\Big|_5^{10} = e^{-1/3} - e^{-2/3} = 0.2031$

d) P(X < x) = 0.90 and $P(X < x) = -e^{-\frac{t}{15}}\Big|_0^x = 1 - e^{-x/15} = 0.90$. Therefore, x = 34.54 minutes.

5-74. Let X be the life of regulator. Then, X is an exponential random variable with $\lambda = 1/E(X) = 1/6$

a) Because the Poisson process from which the exponential distribution is derived is memoryless, this probability is

$$P(X < 6) = \int_0^6 \tfrac{1}{6} e^{-x/6} dx = -e^{-x/6}\Big|_0^6 = 1 - e^{-1} = 0.6321$$

b) Because the failure times are memoryless, the mean time until the next failure is E(X) = 6 years.

5-75. Let X denote the time until a message is received. Then, X is an exponential random variable and $\lambda = 1/E(X) = 1/2$.

a) $P(X > 2) = \int_2^\infty \tfrac{1}{2} e^{-x/2} dx = -e^{-x/2}\Big|_2^\infty = e^{-1} = 0.3679$

b) The same as part a.

c) E(X) = 2 hours.

5-76. Let X denote the time until the arrival of a taxi. Then, X is an exponential random variable with $\lambda = 1/E(X) = 0.1$ arrivals/ minute.

a) $P(X > 60) = \int_{60}^\infty 0.1e^{-0.1x} dx = -e^{-0.1x}\Big|_{60}^\infty = e^{-6} = 0.0025$

b) $P(X < 10) = \int_0^{10} 0.1e^{-0.1x} dx = -e^{-0.1x}\Big|_0^{10} = 1 - e^{-1} = 0.6321$

5-77. a) $P(X > x) = \int_x^\infty 0.1e^{-0.1t} dt = -e^{-0.1t}\Big|_x^\infty = e^{-0.1x} = 0.1$ and x = 23.03 minutes.

b) P(X < x) = 0.9 implies that P(X > x) = 0.1. Therefore, this answer is the same as part a.

c) $P(X < x) = -e^{-0.1t}\Big|_0^x = 1 - e^{-0.1x} = 0.5$ and x = 6.93 minutes.

5-78. Let X denote the distance between major cracks. Then, X is an exponential random variable with $\lambda = 1/E(X) = 0.2$ cracks/mile.

a) $P(X > 10) = \int_{10}^\infty 0.2e^{-0.2x} dx = -e^{-0.2x}\Big|_{10}^\infty = e^{-2} = 0.1353$

b) Let Y denote the number of cracks in 10 miles of highway. Because the distance between cracks is exponential, Y is a Poisson random variable with $\lambda = 10(0.2) = 2$ cracks per 10 miles. P(Y = 2) =

$\frac{e^{-2} 2^2}{2!} = 0.2707$

c) $\sigma_X = 1/\lambda = 5$ miles.

5-79. a) $P(12 < X < 15) = \int_{12}^{15} 0.2e^{-0.2x} dx = -e^{-0.2x}\Big|_{12}^{15} = e^{-2.4} - e^{-3} = 0.0409$

b) $P(X > 5) = -e^{-0.2x}\Big|_5^\infty = e^{-1} = 0.3679$. By independence of the intervals in a Poisson process, the answer is $0.3679^2 = 0.1353$. Alternatively, the answer is P(X > 10) = $e^{-2} = 0.1353$. The probability does depend on whether or not the lengths of highway are consecutive.

c) By the memoryless property, this answer is P(X > 10) = 0.1353 from part b.

5-80. Let X denote the lifetime of an assembly. Then, X is an exponential random variable with $\lambda = 1/E(X) = 1/400$ failures per hour.

a) P(X < 100) = $\int_0^{100} \frac{1}{400} e^{-x/400} dx = -e^{-x/400}\Big|_0^{100} = 1 - e^{-0.25} = 0.2212$

b) P(X > 500) = $-e^{-x/400}\Big|_{500}^{\infty} = e^{-5/4} = 0.2865$

c) From the memoryless property of the exponential, this answer is the same as part a., P(X < 100) = 0.2212.

5-81. a) Let U denote the number of assemblies out of 10 that fail before 100 hours. By the memoryless property of a Poisson process, U has a binomial distribution with n = 10 and p =0.2212 (from Exercise 5-80). Then,

$P(U \geq 1) = 1 - P(U = 0) = 1 - \binom{10}{0} 0.2212^0 (1 - 0.2212)^{10} = 0.9179$

b) Let V denote the number of assemblies out of 10 that fail before 800 hours. Then, V is a binomial random variable with n = 10 and p = P(X < 800), where X denotes the lifetime of an assembly.

Now, P(X < 800) = $\int_0^{800} \frac{1}{400} e^{-x/400} dx = -e^{-x/400}\Big|_0^{800} = 1 - e^{-2} = 0.8647$.

Therefore, P(V = 10) = $\binom{10}{10} 0.8647^{10} (1 - 0.8647)^0 = 0.2336$.

5-82. Let X denote the number of calls in 3 hours. Because the time between calls is an exponential random variable, the number of calls in 3 hours is a Poisson random variable. Now, the mean time between calls is 0.5 hours and $\lambda = 1/0.5 = 2$ calls per hour = 6 calls in 3 hours.

$$P(X \geq 4) = 1 - P(X \leq 3) = 1 - \left[\frac{e^{-6}6^0}{0!} + \frac{e^{-6}6^1}{1!} + \frac{e^{-6}6^2}{2!} + \frac{e^{-6}6^3}{3!}\right] = 0.8488$$

5-83. Let Y denote the number of arrivals in one hour. If the time between arrivals is exponential, then the count of arrivals is a Poisson random variable and $\lambda = 1$ arrival per hour.

$$P(Y > 3) = 1 - P(Y \leq 3) = 1 - \left[\frac{e^{-1}1^0}{0!} + \frac{e^{-1}1^1}{1!} + \frac{e^{-1}1^2}{2!} + \frac{e^{-1}1^3}{3!}\right] = 0.3528$$

5-84. a) From Exercise 5-83, P(Y > 3) = 0.3528. Let W denote the number of one hour intervals out of 30 that contain more than 3 arrivals. By the memoryless property of a Poisson process, W is a binomial random variable with n = 30 and p = 0.3528.

P(W = 0) = $\binom{30}{0} 0.3528^0 (1 - 0.3528)^{30} = 2.15 \times 10^{-6}$

b) Let X denote the time between arrivals. Then, X is an exponential random variable with $\lambda = 3$ arrivals per hour. P(X > x) = 0.1 and $P(X > x) = \int_x^{\infty} 3e^{-3t} dt = -e^{-3t}\Big|_x^{\infty} = e^{-3x} = 0.1$. Therefore, x = 0.768 hours.

5-85. Let X denote the number of calls in 30 minutes. Because the time between calls is an exponential random variable, X is a Poisson random variable with $\lambda = 1/E(X) = 0.1$ calls per minute = 3 calls per 30 minutes.

a) $P(X > 3) = 1 - P(X \leq 3) = 1 - \left[\frac{e^{-3}3^0}{0!} + \frac{e^{-3}3^1}{1!} + \frac{e^{-3}3^2}{2!} + \frac{e^{-3}3^3}{3!}\right] = 0.3528$

b) P(X = 0) = $\frac{e^{-3}3^0}{0!} = 0.04979$

c) Let Y denote the time between calls in minutes. Then, $P(Y \geq x) = 0.01$ and

$P(Y \geq x) = \int_x^{\infty} 0.1e^{-0.1t} dt = -e^{-0.1t}\Big|_x^{\infty} = e^{-0.1x}$. Therefore, $e^{-0.1x} = 0.01$ and x = 46.05 minutes.

5-86. a) From Exercise 5-85, P(Y > 120) = $\int_{120}^{\infty} 0.1e^{-0.1y}dy = -e^{-0.1y}\Big|_{120}^{\infty} = e^{-12} = 6.14 \times 10^{-6}$.

b) Because the calls are a Poisson process, the number of calls in disjoint intervals are independent. From Exercise 5-85 part b., the probability of no calls in one-half hour is $e^{-3} = 0.04979$. Therefore, the answer is $\left[e^{-3}\right]^4 = e^{-12} = 6.14 \times 10^{-6}$. Alternatively, the answer is the probability of no calls in two hours. From part a. of this exercise, this is e^{-12} .

c) Because a Poisson process is memoryless, probabilities do not depend on whether or not intervals are consecutive. Therefore, parts a. and b. have the same answer.

5-87. a) $P(X > \theta) = \int_{\theta}^{\infty} \frac{1}{\theta} e^{-x/\theta}dx = -e^{-x/\theta}\Big|_{\theta}^{\infty} = e^{-1} = 0.3679$

b) $P(X > 2\theta) = -e^{-x/\theta}\Big|_{2\theta}^{\infty} = e^{-2} = 0.1353$

c) $P(X > 3\theta) = -e^{-x/\theta}\Big|_{3\theta}^{\infty} = e^{-3} = 0.0498$

d) The results do not depend on θ .

5-88. X is an exponential random variable with $\lambda = 0.2$ flaws per meter.

a) E(X) = $1/\lambda = 5$ meters.

b) P(X > 10) = $\int_{10}^{\infty} 0.2e^{-0.2x}dx = -e^{-0.2x}\Big|_{10}^{\infty} = e^{-2} = 0.1353$

c) No, see Exercise 5-87 part c.

d) P(X < x) = 0.90. Then, P(X < x) = $-e^{-0.2t}\Big|_{0}^{x} = 1 - e^{-0.2x}$. Therefore, $1 - e^{-0.2x} = 0.9$ and x = 11.51.

5-89. P(X > 8) = $\int_{8}^{\infty} 0.2e^{-0.2x}dx = -e^{-8/5} = 0.2019$

The distance between successive flaws is either less than 8 meters or not. The distances are independent and P(X > 8) = 0.2019. Let Y denote the number of flaws until the distance exceeds 8 meters. Then, Y is a geometric random variable with p = 0.2019.

a) P(Y = 5) = $(1 - 0.2019)^4 0.2019 = 0.0819$.

b) E(Y) = 1/0.2019 = 4.95.

5-90. $E(X) = \int_{0}^{\infty} x\lambda e^{-\lambda x}dx.$ Use integration by parts with *u* = x and *dv* = $\lambda e^{-\lambda x}$.

Then, $E(X) = -xe^{-\lambda x}\Big|_{0}^{\infty} + \int_{0}^{\infty} e^{-\lambda x} = \frac{-e^{-\lambda x}}{\lambda}\Big|_{0}^{\infty} = 1/\lambda$

V(X) = $\int_{0}^{\infty} (x - \frac{1}{\lambda})^2 \lambda e^{-\lambda x}dx.$ Use integration by parts with *u* $= (x - \frac{1}{\lambda})^2$ and *dv* = $\lambda e^{-\lambda x}$. Then,

$$V(X) = -(x - \tfrac{1}{\lambda})^2 e^{-\lambda x}\Big|_{0}^{\infty} + 2\int_{0}^{\infty} (x - \tfrac{1}{\lambda})e^{-\lambda x}dx$$

$$= (\tfrac{1}{\lambda})^2 + \tfrac{2}{\lambda}\int_{0}^{\infty} (x - \tfrac{1}{\lambda})\lambda e^{-\lambda x}dx$$

The last integral is seen to be zero from the definition of E(X). Therefore, V(X) = $\left(\frac{1}{\lambda}\right)^2$.

Section 5-10

5-91. a) The time until the tenth call is an Erlang random variable with $\lambda = 5$ calls per minute and r = 10.
b) E(X) = 10/5 = 2 minutes. V(X) = 10/25 = 0.4 minutes.
c) Because a Poisson process is memoryless, the mean time is 10 minutes.

5-92. Let Y denote the number of calls in one minute. Then, Y is a Poisson random variable with $\lambda = 5$ calls per minute.

a) P(Y = 4) = $\frac{e^{-5}5^4}{4!} = 0.1755$

b) P(Y > 2) = 1 - $P(Y \le 2) = 1 - \frac{e^{-5}5^0}{0!} - \frac{e^{-5}5^1}{1!} - \frac{e^{-5}5^2}{2!} = 0.8754$.

Let W denote the number of one minute intervals out of 10 that contain more than 2 calls. Because the calls are a Poisson process, W is a binomial random variable with n = 10 and p = 0.8754.

Therefore, P(W = 10) = $\binom{10}{10}0.8754^{10}(1-0.8754)^0 = 0.2643$.

5-93. Let X denote the pounds of material to obtain 15 particles. Then, X has an Erlang distribution with r = 15 and $\lambda = 0.01$.

a) E(X) = $\frac{r}{\lambda} = \frac{15}{0.01} = 1500$ pounds.

b) V(X) = $\frac{15}{0.01^2} = 150{,}000$ and $\sigma_X = \sqrt{150{,}000} = 387.3$ pounds.

5-94. Let X denote the time until 5 messages arrive at a node. Then, X has an Erlang distribution with r = 5 and $\lambda = 30$ messages per minute.

a) E(X) = 5/30 = 1/6 minute = 10 seconds.

b)V(X) = $\frac{5}{30^2} = 1/180$ minute^2 = 1/3 second and $\sigma_X = 0.745$ minute = 4.472 seconds.

c) Let Y denote the number of messages that arrive in 10 seconds. Then, Y is a Poisson random variable with $\lambda = 30$ messages per minute = 5 messages per 10 seconds.

$$P(Y \ge 5) = 1 - P(Y \le 4) = 1 - \left[\frac{e^{-5}5^0}{0!} + \frac{e^{-5}5^1}{1!} + \frac{e^{-5}5^2}{2!} + \frac{e^{-5}5^3}{3!} + \frac{e^{-5}5^4}{4!}\right]$$

$$= 0.5595$$

d) Let Y denote the number of messages that arrive in 5 seconds. Then, Y is a Poisson random variable with $\lambda = 2.5$ messages per 5 seconds.

$P(Y \ge 5) = 1 - P(Y \le 4) = 1 - 0.8912 = 0.1088$

5-95. Let X denote the number of bits until five errors occur. Then, X has an Erlang distribution with r = 5 and $\lambda = 10^{-5}$ error per bit.

a) E(X) = $\frac{r}{\lambda} = 5 \times 10^5$ bits.

b) V(X) = $\frac{r}{\lambda^2} = 5 \times 10^{10}$ and $\sigma_X = \sqrt{5 \times 10^{10}} = 223607$ bits.

c) Let Y denote the number of errors in 10^5 bits. Then, Y is a Poisson random variable with

$\lambda = 1/10^5 = 10^{-5}$ error per bit = 1 error per 10^5 bits.

$$P(Y \ge 3) = 1 - P(Y \le 2) = 1 - \left[\frac{e^{-1}1^0}{0!} + \frac{e^{-1}1^1}{1!} + \frac{e^{-1}1^2}{2!}\right] = 0.0803$$

5-96. a) Let X denote the number of customers that arrive in 10 minutes. Then, X is a Poisson random variable with $\lambda = 0.2$ arrivals per minute = 2 arrivals per 10 minutes.

$$P(X > 3) = 1 - P(X \le 3) = 1 - \left[\frac{e^{-2}2^0}{0!} + \frac{e^{-2}2^1}{1!} + \frac{e^{-2}2^2}{2!} + \frac{e^{-2}2^3}{3!} \right] = 0.1429$$

b) Let Y denote the number of customers that arrive in 15 minutes. Then, Y is a Poisson random variable with $\lambda = 3$ arrivals per 15 minutes.

$$P(Y \ge 5) = 1 - P(Y \le 4)$$

$$= 1 - \left[\frac{e^{-3}3^0}{0!} + \frac{e^{-3}3^1}{1!} + \frac{e^{-3}3^2}{2!} + \frac{e^{-3}3^3}{3!} + \frac{e^{-3}3^4}{4!} \right] = 0.1847$$

5-97. Let X denote the time in days until the fourth problem. Then, X has an Erlang distribution with r = 4 and $\lambda = 1/30$ problem per day.

a) $E(X) = \dfrac{4}{30^{-1}} = 120$ days.

b) Let Y denote the number of problems in 120 days. Then, Y is a Poisson random variable with $\lambda = 4$ problems per 120 days.

$$P(Y < 4) = \left[\frac{e^{-4}4^0}{0!} + \frac{e^{-4}4^1}{1!} + \frac{e^{-4}4^2}{2!} + \frac{e^{-4}4^3}{3!} \right] = 0.4335$$

5-98. a) $\Gamma(6) = 5! = 120$

b) $\Gamma(\frac{5}{2}) = \frac{3}{2}\Gamma(\frac{3}{2}) = \frac{3}{2}\frac{1}{2}\Gamma(\frac{1}{2}) = \frac{3}{4}\pi^{1/2}$

c) $\Gamma(\frac{9}{2}) = \frac{7}{2}\Gamma(\frac{7}{2}) = \frac{7}{2}\frac{5}{2}\frac{3}{2}\frac{1}{2}\Gamma(\frac{1}{2}) = \frac{35}{16}\pi^{1/2}$

5-99. $\Gamma(r) = \int_0^\infty x^{r-1}e^{-x}dx$. Use integration by parts with $u = x^{r-1}$ and $dv = e^{-x}$. Then,

$$\Gamma(r) = -x^{r-1}e^{-x}\Big|_0^\infty + (r-1)\int_0^\infty x^{r-2}e^{-x}dx = (r-1)\Gamma(r-1).$$

5-100. $\int_0^\infty f(x;\lambda,r)dx = \int_0^\infty \frac{\lambda^r x^{r-1}e^{-\lambda x}}{\Gamma(r)}dx$. Let y = λx, then the integral is $\int_0^\infty \frac{\lambda\, y^{r-1}e^{-y}}{\Gamma(r)}\frac{dy}{\lambda}$. From the definition of $\Gamma(r)$, this integral is recognized to equal 1.

5-101. If X is a chi-square random variable, then X is a special case of a gamma random variable. Now, E(X) = $\frac{r}{\lambda} = \frac{(7/2)}{(1/2)} = 7$ and $V(X) = \frac{r}{\lambda^2} = \frac{(7/2)}{(1/2)^2} = 14$.

Section 5-11

5-102.

$$E(X) = 100\Gamma(1+\tfrac{1}{0.2}) = 100 \times 5! = 12{,}000$$

$$V(X) = 100^2\Gamma(1+\tfrac{2}{0.2}) - 100^2[\Gamma(1+\tfrac{1}{0.2})]^2 = 3.61 \times 10^{10}$$

5-103. a) $P(X < 10000) = F_X(10000) = 1 - e^{-100^{0.2}} = 1 - e^{-2.512} = 0.9189$

b) $P(X > 5000) = 1 - F_X(5000) = e^{-50^{0.2}} = 0.1123$

5-104. Let X denote lifetime of a bearing.

a) $P(X > 8000) = 1 - F_X(8000) = e^{-0.8^2} = 0.5273$

b)

$$E(X) = 10000\Gamma(1+\tfrac{1}{2}) = 10000\Gamma(1.5)$$
$$= 10000(0.5)\Gamma(0.5) = 5000\sqrt{\pi} = 8862.3$$

c) Let Y denote the number of bearings out of 10 that last at least 8000 hours. Then, Y is a binomial random variable with n = 10 and p = 0.5273.

$$P(Y = 10) = \binom{10}{10} 0.5273^{10}(1 - 0.5273)^0 = 0.00166 .$$

5-105. Let X denote the lifetime.

a) $E(X) = \delta\Gamma(1+\frac{1}{0.5}) = \delta\Gamma(3) = 2\delta = 600.$ Then $\delta = 300$. Now, P(X > 500) = $e^{-(\frac{5}{3})^{0.5}} = 0.2750$

b) P(X < 400) = $1 - e^{-(\frac{5}{4})^{0.5}} = 0.6731$

5-106. Let X denote the lifetime

a) $E(X) = 700\Gamma(1+\frac{1}{2}) = 620.4$

b)

$$V(X) = 700^2\Gamma(2) - 700^2[\Gamma(1.5)]^2$$
$$= 700^2(1) - 700^2(0.25\pi) = 105{,}154.9$$

c) P(X > 620.4) = $e^{-(\frac{620.4}{700})^2} = 0.4559$

5-107. X has an exponential distribution with $\lambda = 1/1000$. Therefore, E(X) = 1000.

<u>Supplemental Exercises</u>

5-108. a) $P(X < 2) = \int_2^2 f(x)dx = 0$

b) $P(X > 3) = \int_3^4 (0.5x - 1)dx = 0.5\frac{x^2}{2} - x\Big|_3^4 = 0.75$

c) $P(2.5 < X < 3.5) = \int_{2.5}^{3.5} (0.5x - 1)dx = 0.5\frac{x^2}{2} - x\Big|_{2.5}^{3.5} = 0.5$

5-109. $F(x) = \int_2^x (0.5t - 1)dt = 0.5\frac{t^2}{2} - t\Big|_2^x = \frac{x^2}{4} - x + 1$. Then,

$$F(x) = \begin{cases} 0, & x < 2 \\ \frac{x^2}{4} - x + 1, & 2 \le x < 4 \\ 1, & 4 \le x \end{cases}$$

5-110. $E(X) = \int_2^4 x(0.5x-1)dx = 0.5\frac{x^3}{3} - \frac{x^2}{2}\Big|_2^4 = \frac{32}{3} - 8 - (\frac{4}{3} - 2) = \frac{10}{3}$

$$V(X) = \int_2^4 (x-\tfrac{10}{3})^2(0.5x-1)dx = \int_2^4 (x^2 - \tfrac{20}{3}x + \tfrac{100}{3})(0.5x-1)dx$$

$$= \int_2^4 (0.5x^3 - \tfrac{13}{3}x^2 + \tfrac{70}{3}x - \tfrac{100}{3})dx = \tfrac{x^4}{8} - \tfrac{13}{9}x^3 + \tfrac{70}{6}x^2 - \tfrac{100}{3}x\Big|_2^4$$

$$= 22.44$$

5-111. Let X denote the time between calls. Then, $\lambda = 1/E(X) = 0.1$ calls per minute.

a) $P(X < 5) = \int_0^5 0.1e^{-0.1x}dx = -e^{-0.1x}\Big|_0^5 = 1 - e^{-0.5} = 0.3935$

b) $P(5 < X < 15) = -e^{-0.1x}\Big|_5^{15} = e^{-0.5} - e^{-1.5} = 0.3834$

c) P(X < x) = 0.9. Then, $P(X < x) = \int_0^x 0.1e^{-0.1t}dt = 1 - e^{-0.1x} = 0.9$. Now, x = 23.03 minutes.

5-112. a) This answer is the same as part a. of Exercise 5-111.

b) This is the probability that there are no calls over a period of 15 minutes. Because a Poisson process is memoryless, it does not matter whether or not the intervals are consecutive.

$$P(X > 15) = \int_{15}^{\infty} 0.1e^{-0.1x}dx = -e^{-0.1x}\Big|_{15}^{\infty} = e^{-1.5} = 0.2231$$

5-113. a) Let Y denote the number of calls in one minute. Then, Y is a Poisson random variable with $\lambda = 0.1$.

$P(Y = 4) = e^{-0.1}\frac{0.1^4}{4!} = 3.77 \times 10^{-6}$.

b) Let W denote the time until the fifth call. Then, W has an Erlang distribution with $\lambda = 0.1$ and r = 5. E(W) = 5/0.1 = 50 minutes.

5-114. Let X denote the lifetime. Then $\lambda = 1/E(X) = 1/6$.

$P(X < 3) = \int_0^3 \frac{1}{6}e^{-x/6}dx = -e^{-x/6}\Big|_0^3 = 1 - e^{-0.5} = 0.3935$.

5-115. Let W denote the number of CPUs that fail within the next three years. Then, W is a binomial random variable with n = 10 and p = 0.3935 (from Exercise 5-115). Then,

$P(W \geq 1) = 1 - P(W = 0) = 1 - \binom{10}{0}0.3935^0(1-0.3935)^{10} = 0.9933$.

5-116. Let X denote the number of fibers visible in a grid cell. Then, X has a Poisson distribution and $\lambda = 100$ fibers per cm^2 = 80,000 fibers per sample = 0.5 fibers per grid cell.

a) $P(X \geq 1) = 1 - P(X) = 1 - \frac{e^{-0.5}0.5^0}{0!} = 0.3935$.

b) Let W denote the number of grid cells examined until 10 contain fibers. If the number of fibers have a Poisson distribution, then the number of fibers in each grid cell are independent. Therefore, W has a negative binomial distribution with p = 0.3935. Consequently, E(W) = 10/0.3935 = 25.41 cells.

c) V(W) = $\frac{10(1-0.3935)}{0.3935^2}$. Therefore, $\sigma_W = 6.25$ cells.

5-117. Let X denote the height of a plant.

a) P(X>3.5) = P(Z > (3.5-2.5)/0.5) = P(Z > 2) = 1 - P(Z ≤ 2) = 0.23

b) P(2.0 < X < 2.5) = P(-1 < Z < 1) = 0.683

c) P(X > x) = 0.90 = P(Z > (x-2.5)/0.5) = 0.90 and (x-2.5)/0.5 = -1.28. Therefore, x = 1.86.

5-118. a) P(X > 3.5) from part a. of Exercise 5-117 is 0.023.
b) Yes, because the probability of a plant growing to a height of 3.5 centimeters or more without irrigation is low.

5-119. Let X denote the thickness.
a) P(X > 5.5) = P(Z > (5.5-5)/0.2) = P(Z > 2.5) = 0. 006
b) $P(4.5 < X < 5.5) = P(\frac{4.5-5.0}{0.2} < Z < \frac{5.5-5.0}{0.2}) = P(-2.5 < Z < 2.5) = 0.988$
Therefore, the proportion that do not meet specifications is 1 − P(4.5 < X < 5.5) = 0.012.
c) If P(X < x) = 0.90, then $P(Z < \frac{x-5.0}{0.2}) = 0.9$. Consequently, $\frac{x-5.0}{0.2} = 1.65$ and x = 5.33.

5-120. Let X denote the dot diameter. If P(0.0014 < X < 0.0026) = 0.9973, then
$P(\frac{0.0014-0.002}{\sigma} < Z < \frac{0.0026-0.002}{\sigma}) = P(\frac{-0.0006}{\sigma} < Z < \frac{0.0006}{\sigma}) = 0.9973$. Therefore, $\frac{0.0006}{\sigma} = 3$ and $\sigma = 0.0002$.

5-121. If P(0.002-x < X < 0.002+x), then P(-x/0.0004 < Z < x/0.0004) = 0.9973. Therefore, x/0.0004 = 3 and x = 0.0012. The specifications are from 0.0008 to 0.0032.

5-122. Let X denote the life.
a) $P(X < 5800) = P(Z < \frac{5800-7000}{600}) = P(Z < -2) = 1 - P(Z \le 2) = 0.023$
b) If P(X > x) = 0.9, then $P(Z < \frac{x-7000}{600}) = 0.9$. Consequently, $\frac{x-7000}{600} = -1.65$ and x = 6010 hours.

5-123. If P(X > 10,000) = 0.99, then $P(Z > \frac{10{,}000-\mu}{600}) = 0.99$. Therefore, $\frac{10{,}000-\mu}{600} = -2.58$ and $\mu = 11{,}158$.

5-124. The probability a product lasts more than 10000 hours is $[P(X > 10000)]^3$, by independence. If $[P(X > 10000)]^3 = 0.99$, then P(X > 10000) = 0.9967.
Then, P(X > 10000) = $P(Z > \frac{10000-\mu}{600}) = 0.9967$. Therefore, $\frac{10000-\mu}{600} = -2.72$ and $\mu = 11{,}632$ hours.

5-125. a) Using the normal approximation to the binomial with n = 50, and p = 0.0001 we have:
$$P(X=1) \cong P(0.5 \le X \le 1.5) = P\left(\frac{0.5-50(0.0001)}{\sqrt{50(0.0001)(0.9999)}} \le \frac{X-np}{\sqrt{np(1-p)}} \le \frac{1.5-50(0.0001)}{\sqrt{50(0.0001)(0.9999)}}\right)$$
= P(7 < z < 21.14) = 0
Using the binomial distribution directly, $P(X=1) = 50(0.0001)^1(0.9999)^{49} = 0.005$
b) $$P(X \ge 4) \cong P(X \ge 3.5) = P\left(\frac{X-np}{\sqrt{np(1-p)}} \ge \frac{3.5-50(0.0001)}{\sqrt{50(0.0001)(0.9999)}}\right) = P(Z \ge 49.43) = 0$$
Using the binomial distribution directly, P(X ≥ 4) = 0

5-126. Using the normal approximation to the binomial with X being the number of people who will be seated.
X ~Bin(200, 0.9).
a) $$P(X \le 185) \cong P(X \le 185.5) = P\left(\frac{X-np}{\sqrt{np(1-p)}} \ge \frac{185.5-180}{\sqrt{200(0.9)(0.1)}}\right) = P(Z \le 1.3) = 0.9032$$
b) $$P(X < 185) \cong P(X \le 184.5) = P\left(\frac{X-np}{\sqrt{np(1-p)}} \ge \frac{184.5-180}{\sqrt{200(0.9)(0.1)}}\right) = P(Z \le 1.06) = 0.85543$$
c) P(X ≤ 185) ≅ 0.95, Using a continuity correction of 0.5 and successively trying various values of n:
The number of reservations taken could be reduced to about 198.

n	Z_o	Probability $P(Z < Z_0)$
190	3.51	0.99978
195	2.39	0.99158
198	**1.73**	**0.95818**

Mind-Expanding Exercises

5-127. a) P(X > x) implies that there are r - 1 or less counts in an interval of length x. Let Y denote the number of counts in an interval of length x. Then, Y is a Poisson random variable with parameter λx. Then,

$$P(X > x) = P(Y \le r-1) = \sum_{i=0}^{r-1} e^{-\lambda x}\frac{(\lambda x)^i}{i!}.$$

b) $P(X \le x) = 1 - \sum_{i=0}^{r-1} e^{-\lambda x}\frac{(\lambda x)^i}{i!}$

c) $f_X(x) = \frac{d}{dx}F_X(x)$

5-128. Let X denote the diameter of the maximum diameter bearing. Then, P(X > 1.575) = 1 - $P(X \le 1.575)$. Also, $X \le 1.575$ if and only if all the diameters are less than 1.575. Let Y denote the diameter of a bearing. Then, by independence $P(X \le 1.575) = [P(Y \le 1.575)]^{10} = \left[P(Z \le \frac{1.575-1.5}{0.025})\right]^{10} = 0.9987^{10} = 0.9864$

Then, P(X > 1.575) = 0.0134.

5-129. a) Quality loss = $Ek(X-m)^2 = kE(X-m)^2 = k\sigma^2$, by the definition of the variance.

b)

$$\begin{aligned}\text{Quality loss} &= Ek(X-m)^2 = kE(X-\mu+\mu-m)^2\\ &= kE[(X-\mu)^2 + (\mu-m)^2 + 2(\mu-m)(X-\mu)]\\ &= kE(X-\mu)^2 + k(\mu-m)^2 + 2k(\mu-m)E(X-\mu).\end{aligned}$$

The last term equals zero by the definition of the mean. Therefore, quality loss = $k\sigma^2 + k(\mu-m)^2$.

5-130. Let X denote the event that an amplifier fails before 60,000 hours. Let A denote the event that an amplifier mean is 20,000 hours. Then A' is the event that the mean of an amplifier is 50,000 hours. Now, P(E) = P(E|A)P(A) + P(E|A')P(A') and

$$P(E|A) = \int_0^{60,000} \frac{1}{20,000}e^{-x/20,000}dx = -e^{-x/20,000}\Big|_0^{60,000} = 1-e^{-3} = 0.9502$$

$$P(E|A') = -e^{-x/50,000}\Big|_0^{60,000} = 1-e^{-6/5} = 0.6988.$$

Therefore, P(E) = 0.9502(0.10) + 0.6988(0.90) = 0.7239

5-131. $P(X < t_1+t_2 | X > t_1) = \frac{P(t_1<X<t_1+t_2)}{P(X>t_1)}$ from the definition of conditional probability. Now,

$$P(t_1 < X < t_1+t_2) = \int_{t_1}^{t_1+t_2} \lambda e^{-\lambda x}dx = -e^{-\lambda x}\Big|_{t_1}^{t_1+t_2} = e^{-\lambda t_1} - e^{-\lambda(t_1+t_2)}$$

$$P(X > t_1) = -e^{-\lambda x}\Big|_{t_1}^{\infty} = e^{-\lambda t_1}$$

Therefore, $P(X < t_1+t_2 | X > t_1) = \dfrac{e^{-\lambda t_1}(1-e^{-\lambda t_2})}{e^{-\lambda t_1}} = 1-e^{-\lambda t_2} = P(X < t_2)$

5-132. a)

$$\begin{aligned}P(\mu_0 - 6\sigma < X < \mu_0 + 6\sigma) &= P(-6 < Z < 6)\\ &= 1.8\times10^{-9} = 0.0018\text{ppm}\end{aligned}$$

b)

$$\begin{aligned}P(\mu_0 - 6\sigma < X < \mu_0 + 6\sigma) &= P(-7.5 < \tfrac{X-(\mu_0+1.5\sigma)}{\sigma} < 4.5)\\ &= 3.4\times10^{-6} = 3.4\text{ppm}\end{aligned}$$

CHAPTER 6

Section 6-1

6-1. First, $f(x,y) \geq 0$. Let R denote the range of (X,Y).

Then, $\sum_R f(x,y) = \frac{1}{8} + \frac{1}{4} + \frac{1}{2} + \frac{1}{8} = 1$

6-2. a) P(X < 2.5, Y < 3) = f(1.5,2) = 1/8
b) P(X < 2.5) = f (1.5, 2) + f (1.5, 3) = 1/8 + 1/4 = 3/8
c) P(Y < 3) = f (1.5, 2) = 1/8
d) P(X > 1.8, Y < 4.7) = f (2.5, 4) = 1/2

6-3. E(X) = 1.5(3/8) + 2.5(1/2) + 3(1/8) = 2.1875
E(Y) = 2(1/8) + 3(1/4) + 4(1/2) + 5(1/8) = 3.625

6-4. a)

x	f(x)
1.5	3/8
2.5	1/2
3	1/8

b) $f_{Y|1.5}(y) = \frac{f_{XY}(1.5,y)}{f_X(1.5)}$ and $f_X(1.5)$ = 3/8. Then,

y	$f_{Y\|1.5}(y)$
2	(1/8)/(3/8)=1/3
3	(1/4)/(3/8)=2/3

c) $f_{X|2}(x) = \frac{f_{XY}(x,2)}{f_Y(2)}$ and $f_Y(2)$ = 1/8. Then,

x	$f_{X\|2}(x)$
1.5	(1/8)/(1/8)=1

d) E(Y|X=1.5) = 2(1/3)+3(2/3) =2 1/3

6-5. Let R denote the range of (X,Y). Because $\sum_R f(x,y) = c(2+3+4+3+4+5+4+5+6) = 1$, 36c = 1, and c = 1/36

6-6. a) $P(X = 1, Y < 4) = f_{XY}(1,1) + f_{XY}(1,2) + f_{XY}(1,3) = \frac{1}{36}(2+3+4) = 1/4$

b) P(X = 1) is the same as part a. = 1/4

c) $P(Y = 2) = f_{XY}(1,2) + f_{XY}(2,2) + f_{XY}(3,2) = \frac{1}{36}(3+4+5) = 1/3$

d) $P(X < 2, Y < 2) = f_{XY}(1,1) = \frac{1}{36}(2) = 1/18$

6-7.

$$E(X) = 1[f_{XY}(1,1) + f_{XY}(1,2) + f_{XY}(1,3)] + 2[f_{XY}(2,1) + f_{XY}(2,2) + f_{XY}(2,3)]$$
$$+ 3[f_{XY}(3,1) + f_{XY}(3,2) + f_{XY}(3,3)]$$
$$= \left(1 \times \frac{9}{36}\right) + \left(2 \times \frac{12}{36}\right) + \left(3 \times \frac{15}{36}\right) = 13/6 = 2.167$$

$$V(X) = (1 - \tfrac{13}{6})^2 \frac{9}{36} + (2 - \tfrac{13}{6})^2 \frac{12}{36} + (3 - \tfrac{13}{6})^2 \frac{15}{36} = 0.639$$

$$E(Y) = 2.167$$

$$V(Y) = 0.639$$

6-8. a)

x	$f_X(x) = f_{XY}(x,1) + f_{XY}(x,2) + f_{XY}(x,3)$
1	1/4
2	1/3
3	5/12

b) $f_{Y|X}(y) = \frac{f_{XY}(1,y)}{f_X(1)}$

y	$f_{Y\|X}(y)$
1	(2/36)/(1/4)=2/9
2	(3/36)/(1/4)=1/3
3	(4/36)/(1/4)=4/9

c) $f_{X|Y}(x) = \frac{f_{XY}(x,2)}{f_Y(2)}$ and $f_Y(2) = f_{XY}(1,2) + f_{XY}(2,2) + f_{XY}(3,2) = \frac{12}{36} = 1/3$

x	$f_{X\|Y}(x)$
1	(3/36)/(1/3)=1/4
2	(4/36)/(1/3)=1/3
3	(5/36)/(1/3)=5/12

d) E(Y|X=1) = 1(2/9)+2(1/3)+3(4/9) = 20/9

6-9. $f(x,y) \geq 0$ and $\sum_R f(x,y) = 1$

6-10. a) $P(X < 0.5, Y < 1.5) = f_{XY}(-1,-2) + f_{XY}(-0.5,-1) = \frac{1}{8} + \frac{1}{4} = \frac{3}{8}$

b) $P(X < 0.5) = f_{XY}(-1,-2) + f_{XY}(-0.5,-1) = \frac{3}{8}$

c) $P(Y < 1.5) = f_{XY}(-1,-2) + f_{XY}(-0.5,-1) + f_{XY}(0.5,1) = \frac{7}{8}$

d) $P(X > 0.25, Y < 4.5) = f_{XY}(0.5,1) + f_{XY}(1,2) = \frac{5}{8}$

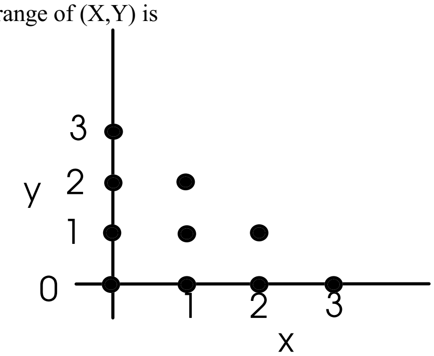
6-11.

$$E(X) = -1(\tfrac{1}{8}) - 0.5(\tfrac{1}{4}) + 0.5(\tfrac{1}{2}) + 1(\tfrac{1}{8}) = \tfrac{1}{8}$$
$$E(Y) = -2(\tfrac{1}{8}) - 1(\tfrac{1}{4}) + 1(\tfrac{1}{2}) + 2(\tfrac{1}{8}) = \tfrac{1}{8}$$

6-12. a) The range of X is {1,2,3}
b) The range of Y given that X = 1 is { 1,2,3 }

6-13. Because X and Y denote the number of printers in each category, $X \geq 0$, $Y \geq 0$ and $X + Y \leq 4$

6-14. a) The range of (X,Y) is



Let H, M, and L denote the events that a bit has high, moderate, and low distortion, respectively. Then,

x,y	$f_{XY}(x,y)$
0,0	0.85738
0,1	0.1083
0,2	0.00456
0,3	0.000064
1,0	0.027075
1,1	0.00228
1,2	0.000048
2,0	0.000285
2,1	0.000012
3,0	0.000001

b)

x	$f_x(x)$
0	0.970299
1	0.29835
2	0.000298
3	0.000001

c) $f_{Y|1}(y) = \frac{f_{XY}(1,y)}{f_X(1)}$, $f_x(1) = 0.29835$

y	$f_{Y|1}(x)$
0	0.09075
1	0.00764
2	0.000161

d) E(Y|X=1) = 0(0.09075)+1(0.00764)+2(0.000161) = 0.007962

6-15. a) The range of (X,Y) is all pairs of integers with $0 \le x \le 48$ and $0 \le y \le 48$

b) The marginal probability distribution of X(Y) is binomial with n = 48 and p = 0.02(0.03).

c) Because X and Y are independent, P(X = 0, Y = 0) = P(X = 0) P(Y = 0) =

$$\binom{48}{0}0.02^0 0.98^{48}\binom{48}{0}0.03^0 0.97^{48} = 0.0879$$

d) $P(X \le 1, Y \le 1) = P(X \le 1)P(Y \le 1)$ and

$$P(X \le 1) = \binom{48}{0}0.02^0 0.98^{48} + \binom{48}{1}0.02^1 0.98^{47} = 0.7506$$ and

$$P(Y \le 1) = 0.97^{48} + \binom{48}{1}0.03^1 0.97^{47} = 0.5758$$. Then, $P(X \le 1, Y \le 1) = 0.4322$

6-16. Because the marginal distribution of X and Y are binomial, E(X) = 48(0.02) = 0.96, V(X) = 48(0.02)(0.98) = 0.9408, E(Y) = 48(0.03) = 1.44, V(Y) = 48(0.03)(0.97) = 1.3968

6-17. Because the range of (X,Y) is not rectangular, X and Y are not independent.

6-18. $f_{XY}(1,1) = 2/36 = 1/18$ and $f_X(1) = \frac{1}{36}\left[f_{XY}(1,1) + f_{XY}(1,2) + f_{XY}(1,3)\right] = 1/4$ and

$f_Y(1) = \frac{1}{36}\left[f_{XY}(1,1) + f_{XY}(2,1) + f_{XY}(3,1)\right] = 1/4$.

Because $f_{XY}(1,1) \ne f_X(1) f_Y(1)$, X and Y are not independent.

6-19. Because the range of (X,Y) is not rectangular, X and Y are not independent.

6-20. Because the range of (X,Y) is not rectangular, X and Y are not independent.

Section 6-2

6-21. a) $P(X=1) = f_{XYZ}(1,1,1) + f_{XYZ}(1,1,2) + f_{XYZ}(1,2,1) + f_{XYZ}(1,2,2) = 0.5$

b) $P(X=1, Y=2) = f_{XYZ}(1,2,1) + f_{XYZ}(1,2,2) = 0.35$

c) $P(Z<1.5) = f_{XYZ}(1,1,1) + f_{XYZ}(1,2,2) + f_{XYZ}(2,1,1) + f_{XYZ}(2,2,1) = 0.5$

d) $P(X=1 \text{ or } Z=2) = P(X=1) + P(Z=2) - P(X=1, Z=2) = 0.5 + 0.5 - 0.3 = 0.7$

e) E(X) = 1(0.5) + 2(0.5) = 1.5

6-22. a) $P(X=1|Y=2) = \frac{P(X=1,Y=2)}{P(Y=2)} = \frac{0.15+0.2}{0.15+0.2+0.1+0.05} = 0.7$

b) $P(X=1, Y=1|Z=2) = \frac{P(X=1, Y=1, Z=2)}{P(Z=2)} = \frac{0.1}{0.1+0.2+0.15+0.05} = 0.2$

c) $P(X=1|Y=2, Z=2) = \frac{P(X=1, Y=2, Z=2)}{P(Y=2, Z=2)} = \frac{0.20}{0.20+0.05} = 0.8$

6-23. $f_{X|YZ}(x) = \frac{f_{XYZ}(x,1,2)}{f_{YZ}(1,2)}$ and $f_{YZ}(1,2) = f_{XYZ}(1,1,2) + f_{XYZ}(2,1,2) = 0.25$

x	$f_{X\|YZ}(x)$
1	0.10/0.25=0.4
2	0.15/0.25=0.6

6-24. a)The random variables X,Y, and Z have a multinomial distribution with n = 20 and

$p_1 = 0.05,\ p_2 = 0.80,\ \text{and } p_3 = 0.15$.

b) The range of (X,Y,Z) is the set of nonnegative integers with x+y+z = 20.
c) X has a binomial distribution with n = 20 and p = 0.05.
d) Because X is binomial, E(X) = 20 (0.05) = 1 and V(X) = 20(0.05)(0.95) = 0.95

6-25. a) $P(X=1, Y=17, Z=3) = \frac{20!}{1!17!3!} 0.05^1 0.80^{17} 0.15^2 = 0.0289$

b)

$$P(X \le 1, Y=17, Z=3) = P(X=0, Y=17, Z=3) + P(X=1, Y=17, Z=3)$$
$$= \frac{20!}{0!17!3!} 0.05^0 0.80^{17} 0.15^3 + 0 = 0.0866$$

Because the point $(1,17,3) \ne 20$ is not in the range of (X,Y,Z).

c) Because X is binomial, $P(X \le 1) = \binom{20}{0} 0.05^0 0.95^{20} + \binom{20}{1} 0.05^1 0.95^{19} = 0.4989$

6-26. a) $P(X=2, Z=3|Y=17) = \frac{P(X=2, Z=3, Y=17)}{P(Y=17)}$.

Because Y is binomial, $P(Y=17) = \binom{20}{17} 0.80^{17} 0.20^3 = 0.2054$.

Then, $P(X=2, Z=3, Y=17) = \frac{20!}{2!3!17!} \frac{0.05^2 0.80^{17} 0.95^3}{0.2054} = 0.1339$

b) $P(X=2|Y=17) = \frac{P(X=2, Y=17)}{P(Y=17)}$. Now, because x+y+z = 20, P(X=2, Y=17) = P(X=2, Y=17, Z=1) =

$\frac{20!}{2!17!1!} 0.05^2 0.80^{17} 0.15^1 = 0.0289$

6-27. a) The range consists of nonnegative integers with x+y+z = 4.

b) Because the samples are selected without replacement, the trials are not independent and the joint distribution is not multinomial.

c) $f_{X|2}(x) = \frac{f_{XY}(x,2)}{f_Y(2)}$. Now, $f_Y(2)$ is the probability that two of the selected printers have extra memory.

Therefore, the marginal probability distribution of Y is a hypergeometric distribution and

$P(Y = 2) = \frac{\binom{5}{2}\binom{10}{2}}{\binom{15}{4}} = 0.3297$. Then, $f_{XY}(x,2) = P(X = x, Y = 2)$ and this equals the number of subsets of size four that contain x printers with graphics enhancement, 2 printers with extra memory, and 4-x-2 = 2-x printers with both features divided by the number of subsets that contain 4 printers. Therefore,

$$f_{XY}(0,2) = \frac{\binom{4}{0}\binom{5}{2}\binom{6}{2}}{\binom{15}{4}} = 0.1099$$

$$f_{XY}(1,2) = \frac{\binom{4}{1}\binom{5}{2}\binom{6}{1}}{\binom{15}{4}} = 0.1758$$

$$f_{XY}(2,2) = \frac{\binom{4}{2}\binom{5}{2}\binom{6}{0}}{\binom{15}{4}} = 0.0440$$

Finally,

$$f_{X|2}(0) = \frac{f_{XY}(0,2)}{f_Y(2)} = 0.333$$

$$f_{X|2}(1) = \frac{f_{XY}(1,2)}{f_Y(2)} = 0.533$$

$$f_{X|2}(2) = \frac{f_{XY}(2,2)}{f_Y(2)} = 0.133$$

6-28. P(X=x, Y=y, Z=z) is the number of subsets of size 4 that contain x printers with graphics enhancements, y printers with extra memory, and z printers with both features divided by the number of subsets of size 4. From the results in Appendix I, it can be shown that

$$P(X = x, Y = y, Z = z) = \frac{\binom{4}{x}\binom{5}{y}\binom{6}{z}}{\binom{15}{4}} \quad \text{for } x+y+z = 4.$$

a) $P(X = 1, Y = 2, Z = 1) = \frac{\binom{4}{1}\binom{5}{2}\binom{6}{1}}{\binom{15}{4}} = 0.1758$

b) $P(X = 1, Y = 1) = P(X = 1, Y = 1, Z = 2) = \frac{\binom{4}{1}\binom{5}{1}\binom{6}{2}}{\binom{15}{4}} = 0.2198$

c) The marginal distribution of X is hypergeometric with N = 15, n = 4, K = 4. Therefore, E(X) = nK/N = 16/15 and V(X) = 4(4/15)(11/15)[11/14] = 0.6146.

6-29. a)

$$P(X=1,Y=2|Z=1)=P(X=1,Y=2,Z=1)/P(Z=1)$$

$$=\left[\frac{\binom{4}{1}\binom{5}{2}\binom{6}{1}}{\binom{15}{4}}\right]\Bigg/\left[\frac{\binom{6}{1}\binom{9}{3}}{\binom{15}{4}}\right]=0.4762$$

b)

$$P(X=2|Y=2)=P(X=2,Y=2)/P(Y=2)$$

$$=\left[\frac{\binom{4}{2}\binom{5}{2}\binom{6}{0}}{\binom{15}{4}}\right]\Bigg/\left[\frac{\binom{5}{2}\binom{10}{2}}{\binom{15}{4}}\right]=0.1333$$

c) Because X+Y+Z = 4, if Y = 0 and Z = 3, then X = 1. Because X must equal 1, $f_{X|YZ}(1)=1$.

6-30. Let X, Y, and Z denote the number of ovens in the sample of four with major, minor, and no defects, respectively.

a) $P(X=2,Y=2,Z=0)=\frac{4!}{2!2!0!}0.6^2 0.3^2 0.1^0=0.1944$

b) $P(X=0,Y=0,Z=4)=\frac{4!}{0!0!4!}0.6^0 0.3^0 0.1^4=0.0001$

6-31. $f_{XY}(x,y)=\sum_R f_{XYZ}(x,y,z)$ where R is the set of values for z such that x+y+z = 4. That is, R consists of the single value z = 4-x-y and $f_{XY}(x,y)=\frac{4!}{x!y!(4-x-y)!}0.6^x 0.3^y 0.1^{4-x-y}$ for $x+y\le 4$.

6-32. Let X,Y, and Z denote the number of bits with high, moderate, and low distortion. Then, the joint distribution of X, Y, and Z is multinomial with n =3 and $p_1=0.01,\ p_2=0.04,$ and $p_3=0.95$.

a)

$$P(X=2,Y=1)=P(X=2,Y=1,Z=0)$$

$$=\frac{3!}{2!1!0!}0.01^2 0.04^1 0.95^0=1.2\times 10^{-5}$$

b) $P(X=0,Y=0,Z=3)=\frac{3!}{0!0!3!}0.01^0 0.04^0 0.95^3=0.8574$

6-33. From Equation 5-12, X has a binomial distribution with n = 3 and p = 0.01. Then, E(X) = 3(0.01) = 0.03 and V(X) = 3(0.01)(0.99) = 0.0297.

Section 6-3

6-34. Determine c such that $c\int_0^3\int_0^3 xy\,dx\,dy=c\int_0^3 y\left.\frac{x^2}{2}\right|_0^3 dy=c(4.5\left.\frac{y^2}{2}\right|_0^3)=\frac{81}{4}c.$

Therefore, c = 4/81.

6-35. a) $P(X<2.5, Y<3) = \frac{4}{81}\int_0^3\int_0^{2.5} xy\,dx\,dy = \frac{4}{81}(3.125)\int_0^3 y\,dy = 0.6944$

b) P(X < 2.5) = P(X < 2.5, Y < 3) because the range of Y is from 0 to 3. Therefore, from part a., the answer is 0.6944.

c) $P(1<Y<2.5) = \frac{4}{81}\int_1^{2.5}\int_0^3 xy\,dx\,dy = \frac{4}{81}\int_1^{2.5} 4.5y\,dy = \frac{18}{81}\left.\frac{y^2}{2}\right|_1^{2.5} = 0.5833$

d) $P(X>1.8, 1<Y<2.5) = \frac{4}{81}\int_1^{2.5}\int_{1.8}^3 xy\,dx\,dy = \frac{4}{81}\int_1^{2.5} 2.88y\,dy = 0.3733$

e) $E(X) = \frac{4}{81}\int_0^3\int_0^3 x^2y\,dx\,dy = \frac{4}{81}\int_0^3 9y\,dy = \frac{4}{9}\left.\frac{y^2}{2}\right|_0^3 = 2$

f) From the symmetry between X and Y, E(Y) = E(X) =2.

6-36. a) $f_X(x) = \int_0^3 f_{XY}(x,y)dy = x\frac{4}{81}\int_0^3 y\,dy = \frac{4}{81}x(4.5) = \frac{2x}{9}$ for $0<x<3$.

b) $f_{Y|1.5}(y) = \frac{f_{XY}(1.5,y)}{f_X(1.5)} = \frac{\frac{4}{81}y(1.5)}{\frac{2}{9}(1.5)} = \frac{2}{9}y$ for $0<y<3$.

c) $E(Y|X=1.5) = \int_{x=1.5}^3 y\left(\frac{2}{9}y\right)dy = \frac{2}{9}\int_{x=1.5}^3 y^2dy = \left.\frac{2y^3}{27}\right|_{x=1.5}^3 = 2-0.25 = 1.75$

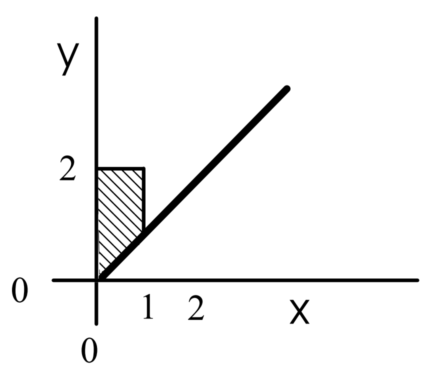
d) $f_{X|2}(x) = \frac{f_{XY}(x,2)}{f_Y(2)} = \frac{\frac{4}{81}x(2)}{\frac{2}{9}(2)} = \frac{4}{81}x$ for $0<x<3$.

6-37.

$$c\int_0^3\int_x^{x+2}(x+y)\,dy\,dx = \int_0^3 \left. xy+\frac{y^2}{2}\right|_x^{x+2} dx = c\int_0^3(4x+2)dx = 24c$$

Therefore, c = 1/24.

6-38. a) P(X < 1, Y < 2) equals the integral of $f_{XY}(x,y)$ over the following region.

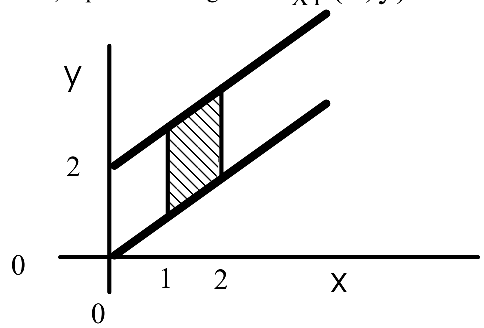


Then,

$$P(X<1, Y<2) = \frac{1}{24}\int_0^1\int_x^2 (x+y)\,dy\,dx = \frac{1}{24}\int_0^1 \left. xy+\frac{y^2}{2}\right|_x^2 dx = \frac{1}{24}\int_0^3 2x+2-\frac{3x^2}{2}dx =$$

$$\frac{1}{24}\left[\left. x^2+2x-\frac{x^3}{2}\right|_0^1\right] = 0.10417$$

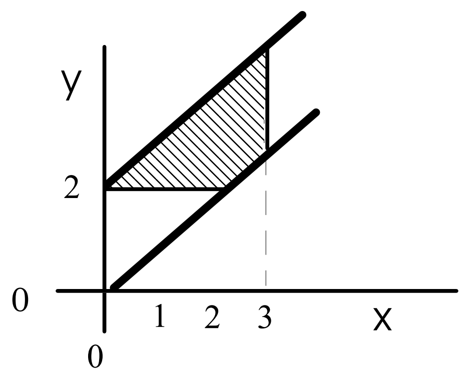
b) P(1 < X < 2) equals the integral of $f_{XY}(x, y)$ over the following region.



$$P(1 < X < 2) = \frac{1}{24}\int_{1}^{2}\int_{x}^{x+2}(x+y)\,dy\,dx = \frac{1}{24}\int_{1}^{2}\left. xy + \frac{y^2}{2}\right|_{x}^{x+2} dx$$

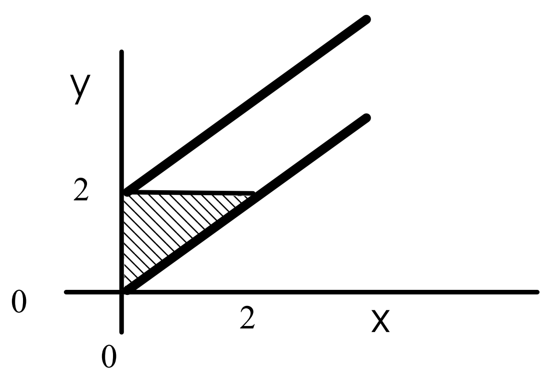
$$= \frac{1}{24}\int_{0}^{3}(3x+2)\,dx = \frac{1}{24}\left[\left.\frac{3x^2}{2} + 2x\right|_{1}^{2}\right] = \frac{1}{3}.$$

c) P(Y > 2) is the integral of $f_{XY}(x, y)$ over the following region.



$$P(Y > 2) = \frac{1}{24}\int_{0}^{2}\int_{2}^{x+2}(x+y)\,dy\,dx + \frac{1}{24}\int_{2}^{3}\int_{x}^{x+2}(x+y)\,dy\,dx = \frac{5}{6} = 0.8333$$

d) P(X < 2, Y < 2) is the integral of $f_{XY}(x, y)$ over the following region.



$$P(X < 2, Y < 2) = \frac{1}{24}\int_{0}^{2}\int_{x}^{2}(x+y)\,dy\,dx = \frac{1}{24}\int_{0}^{2}\left. xy + \frac{y^2}{2}\right|_{x}^{2} dx$$

$$= \frac{1}{24}\int_{0}^{2}\left(2x + 2 - \frac{3x^2}{2}\right)dx = \frac{1}{24}\left[\left. x^2 + 2x - \frac{x^3}{2}\right|_{0}^{2}\right] = \frac{1}{6}.$$

e)

$$E(X) = \frac{1}{24}\int_0^3 \int_x^{x+2} x(x+y)dydx = \frac{1}{24}\int_0^3 x^2y + \frac{xy^2}{2}\Bigg|_x^{x+2} dx$$

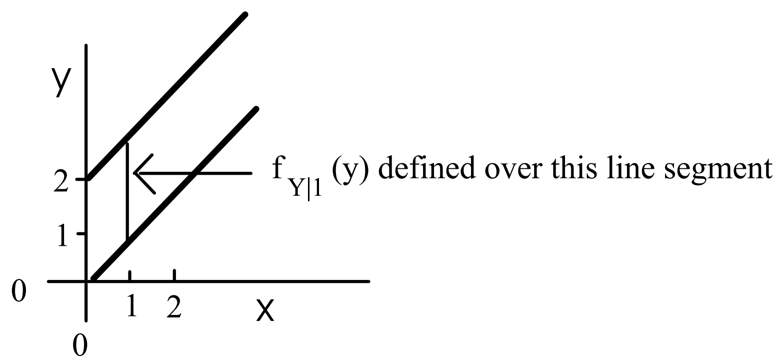
$$= \frac{1}{24}\int_0^3 (3x^2 + 2x)dx = \frac{1}{24}\left[x^3 + x^2\Big|_0^3\right] = \frac{15}{8}$$

6-39. a) $f_X(x)$ is the integral of $f_{XY}(x,y)$ over the interval from x to x+2. That is,

$$f_X(x) = \frac{1}{24}\int_x^{x+2}(x+y)dy = \frac{1}{24}\left[xy + \frac{y^2}{2}\Bigg|_x^{x+2}\right] = \frac{x}{6} + \frac{1}{12} \quad \text{for } 0 < x < 3.$$

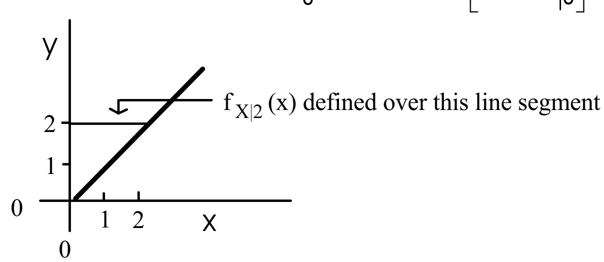
b) $f_{Y|1}(y) = \frac{f_{XY}(1,y)}{f_X(1)} = \frac{\frac{1}{24}(1+y)}{\frac{1}{6}+\frac{1}{12}} = \frac{1+y}{6}$ for $1 < y < 3$.

See the following graph,



c) $E(Y|X=1) = \int_1^3 y\left(\frac{1+y}{6}\right)dy = \int_1^3 \frac{y+y^2}{6}dy = \frac{1}{6}\int_1^3 (y+y^2)dy = \frac{1}{6}\left(\frac{y^2}{2}+\frac{y^3}{3}\right)\Bigg|_1^3 = 2.111$

d) $f_{X|2}(x) = \frac{f_{XY}(x,2)}{f_Y(2)}$ and $f_Y(y) = \frac{1}{24}\int_0^y (x+y)dx = \frac{1}{24}\left[\frac{x^2}{2} + xy\Bigg|_0^y\right] = \frac{y^2}{16}$ for $0 < y \le 2$.



Therefore, $f_Y(2) = 1/4$ and $f_{X|2}(x) = \frac{\frac{1}{24}(x+2)}{1/4} = \frac{x+2}{6}$ for $0 < x < 2$

6-40. $c\int_0^3\int_0^x xydydx = \frac{81}{8}c$. Therefore, c = 8/81

6-41. a) $P(X<1,Y<2)= \frac{8}{81}\int_0^1\int_0^x xydydx = \frac{1}{81}$.

b) $P(1<X<2) = \frac{8}{81}\int_1^2\int_0^x xydydx = \frac{5}{27}$.

c) $P(Y>2)= \frac{8}{81}\int_2^3\int_2^x xydydx = \frac{25}{81}$.

d) $P(X<2,Y<2) = \frac{8}{81}\int_0^2\int_0^2 xydydx = \frac{32}{81}$.

e) $E(X) = \frac{8}{81}\int_0^3\int_0^x x(xy)dydx = \frac{12}{5}$

f) $E(Y) = \frac{8}{81}\int_0^3\int_0^x y(xy)dydx = \frac{8}{5}$

6-42. a) $f(x) = \frac{8}{81}\int_0^x xydy = \frac{4x^3}{81}$

b) $f(Y|X=1) = \frac{f(1,y)}{f(1)} = \frac{\frac{8}{81}(1)y}{\frac{4(1)^3}{81}} = 2y \qquad 0 < x < 3$

d) $f(y) = \frac{8}{81}\int_0^3 xydx = \frac{4y}{9}$, therefore $f(X|Y=2)= \frac{f(x,2)}{f(2)} = \frac{\frac{8}{81}x(2)}{\frac{4(2)}{9}} = \frac{2x}{9} \qquad 0 < y < x$

6-43. $c\int_0^\infty\int_0^x e^{-2x-3y}dydx = \frac{1}{10}c.$, c = 1/10

6-44. a) $P(X<1,Y<2) = 10\int_0^1\int_0^x e^{-2x-3y}dydx = 0.77893$

b) $P(1<X<2) = 10\int_1^2\int_0^x e^{-2x-3y}dydx = 0.19057$

c) $P(Y > 2) = 10\int_2^\infty\int_2^x e^{-2x-3y}dydx = e^{-10}$

d) $P(X<2, Y<2)= 10\int_0^2\int_0^x e^{-2x-3y}dydx = 0.9695$

e) $E(X) = 10\int_0^\infty\int_0^x xe^{-2x-3y}dydx = \frac{7}{10}$

f) $E(Y) = 10\int_0^\infty\int_0^x ye^{-2x-3y}dydx = \frac{1}{5}$

6-45. a) $f(x)= 10\int_0^x e^{-2x-3y}dy = -\frac{10e^{-5x}}{3} + \frac{10e^{-2x}}{3} = \frac{10}{3}(e^{-2x} - e^{-5x})$

b) $P(Y|x=1) = \frac{f(1,y)}{f(1)} = \frac{10e^{-2-3y}}{\frac{10}{3}(e^{-2} - e^{-5})} = 2.3329e^{-2-3y}$

c) $E(Y|x = 1) = 10(2.3329)\int_0^1 ye^{-2-3y}dy = .28094$

d) $P(X|y=2) = \frac{f(x,2)}{f(2)} = \frac{10e^{-2x-6}}{5e^{-6}} = 2e^{-2x}$, where $f(y) = 5e^{-3y}$

6-46. $c\int_0^{\infty}\int_x^{\infty} e^{-2x-3y}dydx = \frac{1}{15}c$, $c = 15$

6-47. a) $P(X<1, Y<2) = 15\int_0^{1}\int_x^{2} e^{-2x-3y}dydx = 0.9879$

b) $P(1<X<2) = 15\int_1^{2}\int_x^{\infty} e^{-2x-3y}dydx = 0.0067$

c) $P(Y>2) = 15\int_0^{2}\int_2^{\infty} e^{-2x-3y}dydx + 152\int_2^{\infty}\int_x^{\infty} e^{-2x-3y}dydx = 0.00613$

d) $P(X<2, Y<2) = 15\int_0^{2}\int_x^{2} e^{-2x-3y}dydx = 0.99387$

e) $E(X) = 15\int_0^{\infty}\int_x^{\infty} xe^{-2x-3y}dydx = \frac{1}{5}$

f) $E(Y) = 15\int_0^{\infty}\int_x^{\infty} ye^{-2x-3y}dydx = \frac{8}{15}$

6-48. a) $f(x) = 15\int_x^{\infty} 3^{-2x-3y}dy = 5e^{-5x}$

b) $f(x=1) = 5e^{-5}$, $f(1,y) = 15e^{-2-3y}$

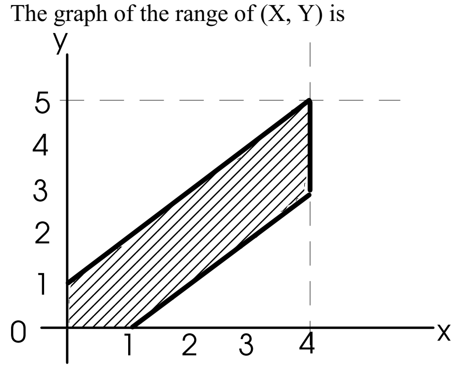
$P(Y|x=1) = \frac{15e^{-2-3y}}{5e^{-5}} = 3e^{3-3y}$

c) $E(Y|x=1) = 15\int_1^{\infty} 3ye^{3-3y}dy = 45\int_1^{\infty} ye^{3-3y}dy = 20$

d) $f(y) = 15\int_0^{\infty} e^{-2x-3y}dx = \frac{15e^{-3y}}{2}$, $f(2) = \frac{15e^{-6}}{2}$

$P(X|y=2) = \frac{15e^{-2x-6}}{\frac{15e^{-6}}{2}} = 2e^{-2x}$

6-49. The graph of the range of (X, Y) is



$$\int_0^1\int_0^{x+1} c\,dydx + \int_1^4\int_{x-1}^{x+1} c\,dydx = 1$$

$$= c\int_0^1 (x+1)dx + 2c\int_1^4 dx$$

$$= \frac{3}{2}c + 6c = 7.5c = 1$$

Therefore, c = 1/7.5

6-50. a) $P(X < 0.5, Y < 0.5) = \int_0^{0.5}\int_0^{0.5} \frac{1}{7.5}dydx = \frac{1}{30}$

b) $P(X < 0.5) = \int_0^{0.5}\int_0^{x+1} \frac{1}{7.5}dydx = \frac{1}{7.5}\int_0^{0.5}(x+1)dx = \frac{2}{15}(\frac{5}{8}) = \frac{1}{12}$

c)

$$E(X) = \int_0^1\int_0^{x+1} \frac{x}{7.5}dydx + \int_1^4\int_{x-1}^{x+1} \frac{x}{7.5}dydx$$

$$= \frac{1}{7.5}\int_0^1 (x^2 + x)dx + \frac{2}{7.5}\int_1^4 (x)dx = \frac{12}{15}(\frac{5}{6}) + \frac{2}{7.5}(7.5) = \frac{19}{9}$$

d)

$$E(Y) = \frac{1}{7.5}\int_0^1\int_0^{x+1} ydydx + \frac{1}{7.5}\int_1^4\int_{x-1}^{x+1} ydydx$$

$$= \frac{1}{7.5}\int_0^1 \frac{(x+1)^2}{2}dx + \frac{1}{7.5}\int_1^4 \frac{(x+1)^2 - (x-1)^2}{2}dx$$

$$= \frac{1}{15}\int_0^1 (x^2 + 2x + 1)dx + \frac{1}{15}\int_1^4 4xdx$$

$$= \frac{1}{15}\left(\frac{7}{3}\right) + \frac{1}{15}(30) = \frac{97}{45}$$

6-51. $f(x) = \int_0^{x+1} \frac{1}{7.5}dy + \int_{x-1}^{x+1} \frac{1}{7.5}dy = \frac{x}{7.5} + \frac{2}{5}$

6-52. a)

$$f_X(x) = \int_{-\infty}^{\infty} \frac{1}{1.2\pi} e^{-\frac{1}{0.72}\left[(x-1)^2 - 1.6(x-1)(y-2) + (y-2)^2\right]}dy$$

$$= \frac{1}{1.2\pi}e^{-\frac{1}{0.72}(x-1)^2}\int_{-\infty}^{\infty} e^{-\frac{1}{0.72}\left\{\left[(y-2)-0.8(x-1)\right]^2 - 0.64(x-1)^2\right\}}dy$$

$$= \frac{1}{1.2\pi}e^{-0.5(x-1)^2}\int_{-\infty}^{\infty} e^{-\frac{1}{0.72}[y-2-0.8(x-1)^2]}dy$$

$$= \frac{1}{1.2\pi}e^{-0.5(x-1)^2}(0.6\sqrt{2\pi})\frac{\int_{-\infty}^{\infty} e^{-0.5\frac{[y-2-0.8(x-1)^2]}{0.6^2}}dy}{(0.6\sqrt{2\pi})}$$

$$= \frac{e^{-0.5(x-1)^2}}{\sqrt{2\pi}} \quad \text{for } -\infty < x < \infty$$

because the integral of a normal probability density is one.

b)The marginal distribution of X is recognized to be normal with $\mu_X = 1$ and $\sigma_X = 1$. Therefore, E(X) = 1.

6-53.

$$f_{Y|X}(y) = \frac{f_{XY}(1,y)}{f_X(1)} = \frac{\frac{1}{1.2\pi}e^{-\frac{1}{0.72}[y-2]^2}}{\frac{1}{\sqrt{2\pi}}} = \frac{e^{-0.5\frac{(y-2)^2}{0.6^2}}}{0.6\sqrt{2\pi}}$$

Therefore, the conditional distribution of Y given that X = 1 is normal with $\mu_{Y|X} = 2$ and $\sigma_{Y|X} = 0.6$.

6-54. Let X, Y, and Z denote the time until a problem on line 1, 2, and 3, respectively.

a)

$$P(X > 40, Y > 40, Z > 40) = P(X > 40)P(Y > 40)P(Z > 40) = [P(X > 40)]^3$$

because the random variables are independent with the same distribution. Now,

$P(X > 40) = \int_{40}^{\infty} \frac{1}{40} e^{-x/40} dx = -e^{-x/40}\Big|_{40}^{\infty} = e^{-1}$ and the answer is $\left(e^{-1}\right)^3 = e^{-3} = 0.0498$.

b) $P(20 < X < 40, 20 < Y < 40, 20 < Z < 40) = [P(20 < X < 40)]^3$ and

$P(20 < X < 40) = -e^{-x/40}\Big|_{20}^{40} = e^{-0.5} - e^{-1} = 0.2387$.

The answer is $0.2387^3 = 0.0136$.

Section 6-4

6-55. a) $P(X < 0.5) = \int_0^{0.5}\int_0^1\int_0^1 (8xyz)dzdydx = \int_0^{0.5}\int_0^1 (4xy)dydx = \int_0^{0.5}(2x)dx = x^2\Big|_0^{0.5} = 0.25$

b)

$$P(X < 0.5, Y < 0.5) = \int_0^{0.5}\int_0^{0.5}\int_0^1 (8xyz)dzdydx$$
$$= \int_0^{0.5}\int_0^{0.5} (4xy)dydx = \int_0^{0.5}(0.5x)dx = \frac{x^2}{4}\Big|_0^{0.5} = 0.0625$$

c) P(Z < 1.5) = 1, because the range of Z is from 0 to 1.

d) P(X < 0.5 or Z < 2) = P(X < 0.5) + P(Z < 2) - P(X < 0.5, Z < 2). Now, P(Z < 2) =1 and P(X < 0.5, Z < 2) = P(X < 0.5). Therefore, the answer is 1.

e) $E(X) = \int_0^1\int_0^1\int_0^1 (8x^2yz)dzdydx = \int_0^1 (2x^2)dx = \frac{2x^3}{3} = 2/3$

6-56. a) $P(X < 0.5 | Y = 0.5)$ is the integral of the conditional density $f_{X|Y}(x)$. Now, $f_{X|0.5}(x) = \frac{f_{XY}(x,0.5)}{f_Y(0.5)}$ and

$f_{XY}(x,y) = \int_0^1 (8xyz)dz = 4xy$ for 0 < x < 1 and 0 < y < 1. Also, $f_Y(y) = \int_0^1\int_0^1 (8xyz)dzdx = 2y$ for 0 < y < 1. Therefore, $f_{X|0.5}(x) = \frac{2x}{1} = 2x$ for 0 < x < 1.

Then, $P(X < 0.5 | Y = 0.5) = \int_0^{0.5} 2xdx = 0.25$.

b) $P(X < 0.5, Y < 0.5 | Z = 0.8)$ is the integral of the conditional density of X and Y. Now, $f_Z(z) = 2z$ for 0 < z < 1 as in part a. and $f_{XY|Z}(x,y) = \frac{f_{XYZ}(x,y,z)}{f_Z(z)} = \frac{8xy(0.8)}{2(0.8)} = 4xy$ for 0 < x < 1 and 0 < y < 1. Then, $P(X < 0.5, Y < 0.5 | Z = 0.8) = \int_0^{0.5}\int_0^{0.5} (4xy)dydx = \int_0^{0.5} (x/2)dx = \frac{1}{16} = 0.0625$

6-57. a) $f_{YZ}(y,z) = \int_0^1 (8xyz)dx = 4yz$ for 0 < y < 1 and 0 < z < 1.

Then, $f_{X|YZ}(x) = \frac{f_{XYZ}(x,y,z)}{f_{YZ}(y,z)} = \frac{8x(0.5)(0.8)}{4(0.5)(0.8)} = 2x$ for 0 < x < 1.

b) Therefore, $P(X < 0.5 | Y = 0.5, Z = 0.8) = \int_0^{0.5} 2xdx = 0.25$

6-58. a) $\iint_{x^2+y^2\le 4} \int_0^4 cdzdydx$ = the volume of a cylinder with a base of radius 2 and a height of 4 = $(\pi 2^2)4 = 16\pi$. Therefore, $c = \frac{1}{16\pi}$

b) $P(X^2 + Y^2 \le 1)$ equals the volume of a cylinder of radius 1 and a height of 4 ($= 4\pi$) times c. Therefore, the answer is $\frac{4\pi}{16\pi} = 1/4$.

c) P(Z < 2) equals half the volume of the region where $f_{XYZ}(x,y,z)$ is positive times 1/c. Therefore, the answer is 0.5.

d) $E(X) = \int_{-2}^{2} \int_{-\sqrt{4-x^2}}^{\sqrt{4-x^2}} \int_0^4 \frac{x}{c} dzdydx = \frac{1}{c}\int_{-2}^{2}\left[4xy\Big|_{-\sqrt{4-x^2}}^{\sqrt{4-x^2}}\right]dx = \frac{1}{c}\int_{-2}^{2}(8x\sqrt{4-x^2})dx$. Using substitution,

$u = 4 - x^2$, du = -2x dx, and $E(X) = \frac{1}{c}\int 4\sqrt{u}du = \frac{-4}{c}\frac{2}{3}(4-x^2)^{\frac{3}{2}}\Big|_{-2}^{2} = 0$.

6-59. a) $f_{X|1}(x) = \frac{f_{XY}(x,1)}{f_Y(1)}$ and $f_{XY}(x,y) = \frac{1}{c}\int_0^4 dz = \frac{4}{c} = \frac{1}{4\pi}$ for $x^2 + y^2 < 4$.

Also, $f_Y(y) = \frac{1}{c}\int_{-\sqrt{4-y^2}}^{\sqrt{4-y^2}}\int_0^4 dzdx = \frac{8}{c}\sqrt{4-y^2}$ for $-2 < y < 2$.

Then, $f_{X|y}(x) = \frac{4/c}{\frac{8}{c}\sqrt{4-y^2}}$ evaluated at y = 1. That is, $f_{X|1}(x) = \frac{1}{2\sqrt{3}}$ for $-\sqrt{3} < x < \sqrt{3}$.

Therefore, $P(X < 1 | Y < 1) = \int_{-\sqrt{3}}^{1} \frac{1}{2\sqrt{3}} dx = \frac{1+\sqrt{3}}{2\sqrt{3}} = 0.7887$

b) $f_{XY|1}(x,y) = \frac{f_{XYZ}(x,y,1)}{f_Z(1)}$ and $f_Z(z) = \int_{-2}^{2}\int_{-\sqrt{4-x^2}}^{\sqrt{4-x^2}} \frac{1}{c} dydx = \int_{-2}^{2} \frac{2}{c}\sqrt{4-x^2}dx$

Because $f_Z(z)$ is a density over the range 0 < z < 4 that does not depend on Z, $f_Z(z)$ =1/4 for 0 < z < 4.

Then, $f_{XY|1}(x,y) = \frac{1/c}{1/4} = \frac{1}{4\pi}$

for $x^2 + y^2 < 4$. Then, $P(x^2 + y^2 < 1 | Z = 1) = \frac{\text{area in } x^2 + y^2 < 1}{4\pi} = 1/4$.

6-60. $f_{Z|xy}(z) = \frac{f_{XYZ}(x,y,z)}{f_{XY}(x,y)}$ and from part a., $f_{XY}(x,y) = \frac{1}{4\pi}$ for $x^2 + y^2 < 4$. Therefore,

$f_{Z|xy}(z) = \frac{\frac{1}{16\pi}}{\frac{1}{4\pi}} = 1/4$ for 0 < z < 4.

6-61. Let X denote the production yield on a day. Then,

$P(X > 1400) = P(Z > \frac{1400-1500}{\sqrt{10000}}) = P(Z > -1) = 0.84134$.

a) Let Y denote the number of days out of five such that the yield exceeds 1400. Then, by independence, Y has a binomial distribution with n = 5 and p = 0.8413. Therefore, the answer is

$$P(Y = 5) = \binom{5}{5} 0.8413^5 (1-0.8413)^0 = 0.4215.$$

b) As in part a., the answer is

$$P(Y \ge 4) = P(Y = 4) + P(Y = 5)$$
$$= \binom{5}{4} 0.8413^4 (1-0.8413)^1 + 0.4215 = 0.8190$$

6-62. a) Let X denote the weight of a brick. Then, $P(X > 2.75) = P(Z > \frac{2.75-3}{0.25}) = P(Z > -1) = 0.84134$.

Let Y denote the number of bricks in the sample of 20 that exceed 2.75 pounds. Then, by independence, Y has a binomial distribution with n = 20 and p = 0.84134. Therefore, the answer is

$$P(Y = 20) = \binom{20}{20} 0.84134^{20} = 0.032.$$

b) Let A denote the event that the heaviest brick in the sample exceeds 3.75 pounds. Then, P(A) = 1 - P(A') and A' is the event that all bricks weigh less than 3.75 pounds. As in part a., P(X < 3.75) = P(Z < 3) and

$$P(A) = 1 - [P(Z < 3)]^{20} = 0.0267.$$

Section 6-5

6-63. E(X) = 7/4 = 1.75
E(Y) = 37/8 = 4.625

$$E(XY) = [1\times 3\times(1/8)] + [1\times 4\times(1/4)] + [2\times 5\times(1/2)] + [3\times 6\times(1/8)]$$
$$= 69/8 = 8.625$$

V(X) = 0.4375 V(Y) = 0.7344

$$\sigma_{XY} = 8.625 - (1.75)(4.625) = 0.5313$$

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X\sigma_Y} = 0.937$$

6-64. E(X) = 0/4 = 0
E(Y) =0/8 = 0

$$E(XY) = [-1\times -2\times(1/8)] + [-0.5\times -1\times(1/4)] + [0.5\times 1\times(1/2)] + [1\times 2\times(1/8)]$$ V(X) = 0.4375
$$= 0.875$$

V(Y) = 1.75

$$\sigma_{XY} = 0.875 - (0)(0) = 0.875$$

$$\rho_{XY} = \frac{\sigma_{xy}}{\sigma_x\sigma_y} = 1$$

6-65.

$$\sum_{x=1}^{3}\sum_{y=1}^{3} c(x+y) = 36c, \qquad c = 1/36$$

$$E(X) = \frac{13}{6} \qquad E(Y) = \frac{13}{6} \qquad E(XY) = \frac{14}{3} \qquad \sigma_{xy} = \frac{14}{3} - \left(\frac{13}{6}\right)^2 = \frac{-1}{36}$$

$$E(X^2) = \frac{16}{3} \qquad E(Y^2) = \frac{16}{3} \qquad V(X) = V(Y) = \frac{23}{36}$$

$$\rho = \frac{\frac{-1}{36}}{\sqrt{\frac{23}{36}}\sqrt{\frac{23}{36}}} = -0.0435$$

6-66. E(X) = 1(0.9702 + 0.0198) = 0.99
E(Y) = 1(0.0098 + 0.9702) = 0.98
V(X) = 0.0099 V(Y) = 0.0196
E(XY) = 1 × 1 × 0.9702 = 0.9702

$$\sigma_{XY} = 0.9702 - E(X)E(Y) = 0 \text{ and } \rho_{XY} = 0$$

6-68. From Exercise 6-40, c=8/81.
From Exercise 6-41, E(X) = 12/5, and E(Y) = 8/5
Now E(XY) = 4

$$\sigma_{xy} = 4 - \left(\frac{12}{5}\right)\left(\frac{8}{5}\right) = 0.16$$

$$E(X^2) = 6 \qquad E(Y^2) = 3$$

$$v(x) = 0.24, \qquad v(y) = 0.44$$

$$\rho = \frac{0.16}{\sqrt{0.24}\sqrt{0.44}} = 0.4923$$

6-69. From probem 6-50c) and 6-50d), E(X) = 2.11, E(Y) = 2.155, respectively.

Now, E(XY) = $E(XY) = \int_0^1 \int_0^{x+1} \frac{xy}{7.5} dydx + \int_1^4 \int_{x-1}^{x+1} \frac{xy}{7.5} dydx = 5.6944$.

$$\sigma_{xy} = 5.6944 - (2.11)(2.155) = 1.1437$$

$$E(X^2) = 5.6778 \qquad E(Y^2) = 6.033$$

$$v(x) = 1.211, \qquad v(y) = 1.3867$$

$$\rho = \frac{1.1437}{\sqrt{1.211}\sqrt{2.155}} = 0.87891$$

6-70.

$$E(X) = 6 \times 10^{-6} \int_0^{\infty} \int_x^{\infty} xe^{-.001x-.002y} dydx = 333.33$$

$$E(X^2) = 6 \times 10^{-6} \int_0^{\infty} \int_x^{\infty} x^2 e^{-.001x-.002y} dydx = 2222200$$

$$V(X) = 2222200 - (333.33)^2 = 2111091.11$$

$$E(Y) = 6 \times 10^{-6} \int_0^{\infty} \int_x^{\infty} ye^{-.001x-.002y} dydx = 833.33$$

$$E(Y^2) = 6 \times 10^{-6} \int_0^{\infty} \int_x^{\infty} y^2 e^{-.001x-.002y} dydx = 10556000$$

$$V(Y) = 10556000 - (833.33)^2 = 9861561.11$$

$$E(XY) = 6 \times 10^{-6} \int_0^{\infty} \int_x^{\infty} xye^{-.001x-.002y} dydx = 3888900$$

$$\sigma_{xy} = 3888900 - (333.33)(833.33) = 3611126.11$$

$$\rho = \frac{3611126.11}{\sqrt{2111091.11}\sqrt{9861561.11}} = 0.7914$$

6-71. a) E(X) = 1 E(Y) = 1

$$E(XY) = \int_0^{\infty} \int_0^{\infty} xye^{-x-y} dxdy$$

$$= \int_0^{\infty} xe^{-x} dx \int_0^{\infty} ye^{-y} dy$$

$$= E(X)E(Y)$$

Therefore, $\sigma_{XY} = \rho_{XY} = 0$.

b) E(X) = -1(1/4)+0(1/2)+1(1/4) = 0, E(Y)=0, E(XY)=0

Therefore, $\sigma_{XY} = \rho_{XY} = 0$. However, P(Y=0) = 1/2 and P(Y=0|X=-1) = 1. Therefore, X and Y are <u>not</u> independent.

6-72. Suppose the correlation between X and Y is ρ. for constants a, b, c, and d, what is the correlation between the random variables U = aX+b and V = cY+d?

Now, E(U) = a E(X) + b and E(V) = c E(Y) + d.

Also, U - E(U) = a[X - E(X)] and V - E(V) = c[Y - E(Y)]. Then,

$$\sigma_{UV} = E\{[U - E(U)][V - E(V)]\} = acE\{[X - E(X)][Y - E(Y)]\} = ac\sigma_{XY}$$

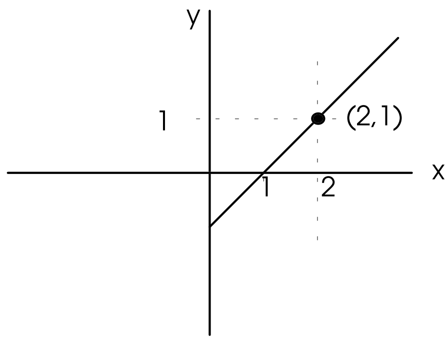
Also, $\sigma_U^2 = E[U - E(U)]^2 = a^2 E[X - E(X)]^2 = a^2 \sigma_X^2$ and $\sigma_V^2 = c^2 \sigma_Y^2$. Then,

$$\rho_{UV} = \frac{ac\sigma_{XY}}{\sqrt{a^2\sigma_X^2}\sqrt{c^2\sigma_Y^2}} = \begin{cases} \rho_{XY} & \text{if a and c are of the same sign} \\ -\rho_{XY} & \text{if a and c differ in sign} \end{cases}$$

6-73. If X and Y are independent, then $f_{XY}(x, y) = f_X(x)f_Y(y)$ and the range of (X, Y) is rectangular. Therefore, $E(XY) = \iint xyf_X(x)f_Y(y)dxdy = \int xf_X(x)dx\int yf_Y(y)dy = E(X)E(Y)$

Section 6-6

6-74.



Because $\rho > 0$, the major axis of the ellipse has positive slope and it is centered at x = 2 and y = 1.

6-75. Because $\rho = 0$ and X and Y are normally distributed, X and Y are independent. Therefore,
P(2.95 < X < 3.05, 7.60 < Y < 7.80) = P(2.95 < X < 3.05) P(7.60 < Y < 7.80) =
$P(\frac{2.95-3}{0.04} < Z < \frac{3.05-3}{0.04})P(\frac{7.60-7.70}{0.08} < Z < \frac{7.80-7.70}{0.08}) =$
$0.7887^2 = 0.6220$

6-76.

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x, y)dxdy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left[\frac{1}{2\pi\sigma_X\sigma_Y}e^{-\frac{1}{2}\left[\frac{(x-\mu_X)^2}{\sigma_X^2}+\frac{(y-\mu_Y)^2}{\sigma_Y^2}\right]}\right]dxdy =$$

$$\int_{-\infty}^{\infty}\left[\frac{1}{\sqrt{2\pi}\sigma_X}e^{-\frac{1}{2}\left[\frac{(x-\mu_X)^2}{\sigma_X^2}\right]}\right]dx\int_{-\infty}^{\infty}\left[\frac{1}{\sqrt{2\pi}\sigma_Y}e^{-\frac{1}{2}\left[\frac{(y-\mu_Y)^2}{\sigma_Y^2}\right]}\right]dy$$

and each of the last two integrals is recognized as the integral of a normal probability density function from $-\infty$ to ∞. That is, each integral equals one.

6-77.

$$f_X(x) = \int_{-\infty}^{\infty} \left[\frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} e^{\frac{-0.5}{1-\rho^2}\left[\frac{(x-\mu_X)^2}{\sigma_X^2} - \frac{2\rho(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y} + \frac{(y-\mu_Y)^2}{\sigma_Y^2}\right]} \right] dy$$

$$= \frac{1}{\sqrt{2\pi}\sigma_X} e^{\frac{-0.5}{1-\rho^2}\frac{(x-\mu_X)^2}{\sigma_X^2}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_Y\sqrt{1-\rho^2}} e^{\frac{-0.5}{1-\rho^2}\left\{\left[\frac{(y-\mu_Y)}{\sigma_Y} - \frac{\rho(x-\mu_X)}{\sigma_X}\right]^2 - \left[\frac{\rho(x-\mu_X)}{\sigma_X}\right]^2\right\}} dy$$

$$= \frac{1}{\sqrt{2\pi}\sigma_X} e^{-0.5\frac{(x-\mu_X)^2}{\sigma_X^2}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_Y\sqrt{1-\rho^2}} e^{\frac{-0.5}{1-\rho^2}\left[\frac{(y-\mu_Y)}{\sigma_Y} - \frac{\rho(x-\mu_X)}{\sigma_X}\right]^2} dy$$

The last integral is recognized as the integral of a normal probability density with mean $\mu_Y + \frac{\sigma_Y\rho(x-\mu_X)}{\sigma_X}$ and variance $\sigma_Y{}^2(1-\rho^2)$. Therefore, the last integral equals one and the requested result is obtained.

6-78. $E(X) = \mu_X, E(Y) = \mu_Y, V(X) = \sigma_X^2,$ and $V(Y) = \sigma_Y^2$. Also,

$$E(X-\mu_X)(Y-\mu_Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{(x-\mu_X)(y-\mu_Y)e^{\frac{-0.5}{1-\rho^2}\left[\frac{(x-\mu_X)^2}{\sigma_X^2} - \frac{2\rho(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y} + \frac{(y-\mu_Y)^2}{\sigma_Y^2}\right]}}{2\pi\sigma_X\sigma_Y(1-\rho^2)^{1/2}} dxdy$$

Let $u = \frac{x-\mu_X}{\sigma_X}$ and $v = \frac{y-\mu_Y}{\sigma_Y}$. Then,

$$E(X-\mu_X)(Y-\mu_Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{uve^{\frac{-0.5}{1-\rho^2}\left[u^2 - 2\rho uv + v^2\right]}}{2\pi(1-\rho^2)^{1/2}} \sigma_X\sigma_Y dudv$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{uve^{\frac{-0.5}{1-\rho^2}\left\{[u-\rho v]^2 + (1-\rho^2)v^2\right\}}}{2\pi(1-\rho^2)^{1/2}} \sigma_X\sigma_Y dudv$$

The integral with respect to u is recognized as a constant times the mean of a normal random variable with mean ρv and variance $1-\rho^2$. Therefore,

$$E(X-\mu_X)(Y-\mu_Y) = \int_{-\infty}^{\infty} \frac{v}{\sqrt{2\pi}} e^{-0.5v^2} \rho v\sigma_X\sigma_Y dudv = \rho\sigma_X\sigma_Y \int_{-\infty}^{\infty} \frac{v^2}{2\pi} e^{-0.5v^2} dv.$$

The last integral is recognized as the variance of a normal random variable with mean 0 and variance 1. Therefore, $E(X-\mu_X)(Y-\mu_Y) = \rho\sigma_X\sigma_Y$ and the correlation between X and Y is ρ.

Section 6-7

6-79. a) E(2X + 3Y) = 2(0) + 3(10) = 30
b) V(2X + 3Y) = 4V(X) + 9V(Y) = 97
c) 2X + 3Y is normally distributed with mean 30 and variance 97. Therefore,

$$P(2X+3Y<30)=P(Z<\frac{30-30}{\sqrt{97}})=P(Z<0)=0.5$$

d) $P(2X+3Y<40)=P(Z<\frac{40-30}{\sqrt{97}})=P(Z<1.02)=0.846$

6-80. Y = 25.4 X and E(Y) = 25.4 E(X) = 127 mm.

6-81. a) Let T denote the total thickness. Then, T = X + Y and E(T) = 3 mm,
V(T) = $0.1^2+0.1^2=0.02mm^2$, and $\sigma_T=0.1414$ mm.
b) P(T > 3.3) = P(Z > $\frac{3.3-3}{0.1414}$) = P(Z > 2.12) = 1 - 0.983 = 0.017

6-82. Let D denote the width of the casing minus the width of the door. Then, D is normally distributed.
a) E(D) = 1/8 V(D) = $(\frac{1}{8})^2+(\frac{1}{16})^2=\frac{5}{256}$

b) $P(D>\frac{1}{4})=P(Z>\frac{\frac{1}{4}-\frac{1}{8}}{\sqrt{5/256}})=P(Z>0.89)=0.187$

c) $P(D<0)=P(Z<\frac{0-\frac{1}{8}}{\sqrt{5/256}})=P(Z<-0.89)=0.187$

6-83. D = A - B - C
a) E(D) = 10 - 2 - 2 = 6 mm

$$V(D)=0.1^2+0.05^2+0.05^2=0.015mm^2$$

$$\sigma_D=0.1225mm$$

b) P(D < 5.9) = P(Z < $\frac{5.9-6}{0.1225}$) = P(Z < -0.82) = 0.206.

6-84. a) Let $\overline{X}$ denote the average fill-volume of 100 cans. $\sigma_{\overline{X}}=\sqrt{0.5^2/100}=0.05$.
b) E($\overline{X}$) = 12.1 and P($\overline{X}$ < 12) = P(Z < $\frac{12-12.1}{0.05}$) = P(Z < -2) = 0.023.
c) P($\overline{X}$ < 12) = 0.005 implies that P(Z < $\frac{12-\mu}{0.05}$) = 0.005. Then $\frac{12-\mu}{0.05}$ = -2.58 and $\mu=12.129$.

6-85. Let $\overline{X}$ denote the average thickness of 10 wafers. Then, E($\overline{X}$) = 10 and V($\overline{X}$) = 0.1.
a) $P(9<\overline{X}<11)=P(\frac{9-10}{\sqrt{0.1}}<Z<\frac{11-10}{\sqrt{0.1}})=P(-3.16<Z<3.16)=0.998$. The answer is 1 – 0.998 = 0.002

b) $P(\overline{X}>11)=0.01$ and $\sigma_{\overline{X}}=\frac{1}{\sqrt{n}}$.

Therefore, $P(\overline{X}>11)=P(Z>\frac{11-10}{1/\sqrt{n}})=0.01$, $\frac{11-10}{1/\sqrt{n}}$ = 2.33 and n = 5.43 which is rounded up to 6.

6-86. X~N(160, 900)
a) Let Y = 25X, E(Y)=25E(X)=4000, V(Y) = 25^2(900)=562500

P(Y>4500) = $P\left(\frac{Y-\mu_y}{\sigma_y}>\frac{4500-4000}{\sqrt{562500}}\right)=P(Z>0.67)=1-P(Z<0.67)=1-0.74857=0.25143$

Section 6-8

6-87. Use Chebychev's inequality with c = 4. Then, $P(|X-10|>4) \le \frac{1}{16}$.

6-88. E(X) = 5 and $\sigma_X = 2.887$. Then, $P(|X-5|>2\sigma_X) \le \frac{1}{4}$.

The actual probability is $P(|X-5|>2\sigma_X) = P(|X-5|>5.77) = 0$.

6-89. E(X) = 20 and V(X) = 400. Then, $P(|X-20|>2\sigma) \le \frac{1}{4}$ and $P(|X-20|>3\sigma) \le \frac{1}{9}$ The actual probabilities are

$$P(|X-20|>2\sigma) = 1 - P(|X-20|<40)$$
$$= 1 - \int_0^{60} 0.05e^{-0.05x}dx = 1 - \left[-e^{-0.05x}\Big|_0^{60}\right] = 0.0498$$
$$P(|X-20|>3\sigma) = 1 - P(|X-20|<60)$$
$$= 1 - \int_0^{80} 0.05e^{-0.05x}dx = 1 - \left[-e^{-0.05x}\Big|_0^{80}\right] = 0.0183$$

6-90. E(X) = 4 and $\sigma_X = 2$

$P(|X-4| \ge 4) \le \frac{1}{4}$ and $P(|X-4| \ge 6) \le \frac{1}{9}$. The actual probabilities are

$$P(|X-4| \ge 4) = 1 - P(|X-4|<4) = 1 - \sum_{x=1}^{7} \frac{e^{-2}2^x}{x!} = 1 - 0.8636 = 0.1364$$
$$P(|X-4| \ge 6) = 1 - P(|X-4|<6) = 1 - \sum_{x=1}^{9} \frac{e^{-2}2^x}{x!} = 0.000046$$

6-91. Let $\overline{X}$ denote the average of 500 diameters. Then, $\sigma_{\overline{X}} = \frac{0.01}{500} = 2x10^{-5}$.

a) $P(|\overline{X}-\mu| \ge 4\sigma_{\overline{X}}) \le \frac{1}{16}$ and $P(|\overline{X}-\mu| < 8x10^{-5}) \ge \frac{15}{16}$. Therefore, the bound is $8x10^{-5}$.

b) If $P(|\overline{X}-\mu| < x) = \frac{15}{16}$, then $P(\frac{-x}{\sigma_{\overline{X}}} < \frac{\overline{X}-\mu}{\sigma_{\overline{X}}} < \frac{x}{\sigma_{\overline{X}}}) = 0.9375.$ Then,

$P(\frac{-x}{2x10^{-5}} < Z < \frac{x}{2x10^{-5}}) = 0.9375.$ and $\frac{x}{2x10^{-5}} = 1.86$. Therefore, x = $3.75x10^{-5}$.

Supplemental Exercises

6-92. If R denotes the range of (X, Y), then $\sum\sum_R f_{XY}(x,y) = 1$ and $f_{XY}(x,y) \ge 0$.

6-93. a) $P(X<0.5, Y<1.5) = f_{XY}(0,1) = 1/8$.

b) $P(X \le 1) = f_{XY}(0,1) + f_{XY}(1,0) + f_{XY}(1,1) = 1/2$

c) $P(Y < 1.5) = f_{XY}(0,1) + f_{XY}(1,0) + f_{XY}(1,1) = 1/2$

d) $P(X > 0.5, Y < 1.5) = f_{XY}(1,0) + f_{XY}(1,1) = 3/8$

6-94. E(X) = 0(1/8) + 1(3/8) + 2(1/2) = 11/8.

$$V(X) = (0 - \frac{11}{8})^2 \frac{1}{8} + (1 - \frac{11}{8})^2 \frac{3}{8} + (2 - \frac{11}{8})^2 \frac{1}{2} = \frac{31}{64} = 0.4844 .$$

E(Y) = 1(3/8) + 0(1/8) + 2(1/2) = 11/8.

$$V(Y) = (1 - \frac{11}{8})^2 \frac{3}{8} + (0 - \frac{11}{8})^2 \frac{1}{8} + (2 - \frac{11}{8})^2 \frac{1}{2} = \frac{31}{64} = 0.4844 .$$

6-95. a) $f_X(x) = \sum_y f_{XY}(x,y)$ and $f_X(0) = 1/8$, $f_X(1) = 3/8$, $f_X(2) = 1/2$.

b) $f_{Y|1}(y) = \frac{f_{XY}(1,y)}{f_X(1)}$ and $f_{Y|1}(0) = \frac{1/8}{3/8} = 1/3$, $f_{Y|1}(1) = \frac{1/4}{3/8} = 2/3$.

6-96. Because the range of (X, Y) is not rectangular, X and Y are not independent.

6-97. E(XY) = 0x1x(1/8) + 1x0x(1/8) + 1x1x(1/4) + 2x2x(1/2) = 5/4.
From Exercise 5-79, E(X) = E(Y) = 11/8.
Then, $\rho_{XY} = \frac{5}{4} - \frac{11}{8}(\frac{11}{8}) = \frac{-41}{64} = -0.641$.

6-98. Let X, Y, and Z denote the number of bolts rated high, moderate, and low.
Then, X, Y, and Z have a multinomial distribution.

a) $P(X = 12, Y = 6, Z = 2) = \frac{20!}{12!6!2!} 0.6^{12} 0.3^6 0.1^2 = 0.0560$.

b) Because X, Y, and Z are multinomial, the marginal distribution of Z is binomial with n = 20 and p = 0.1.

c) E(Z) = np = 20(0.1) = 2.

d)

$$P(Z > 2) = 1 - P(Z \le 2) = 1 - \left[\binom{20}{0} 0.1^0 0.9^{20} + \binom{20}{1} 0.1^1 0.9^{19} + \binom{20}{2} 0.1^2 0.9^{18}\right]$$

$$= 0.3231.$$

6-99. a) $f_{Z|16}(z) = \frac{f_{XZ}(16,z)}{f_X(16)}$ and $f_{XZ}(x,z) = \frac{20!}{x!z!(20-x-z)!} 0.6^x 0.3^{(20-x-z)} 0.1^z$ for

$x + z \le 20$ and $0 \le x, 0 \le z$. Then,

$$f_{Z|16}(z) = \frac{\frac{20!}{16!z!(4-z)!} 0.6^{16} 0.3^{(4-z)} 0.1^z}{\frac{20!}{16!4!} 0.6^{16} 0.4^4} = \frac{4!}{z!(4-z)!} \left(\frac{0.3}{0.4}\right)^{4-z} \left(\frac{0.1}{0.4}\right)^z$$

for $0 \le z \le 4$. That is the distribution of Z given X = 16 is binomial with n = 4 and p = 0.25.

b) From part a., E(Z) = 4 (0.25) = 1.

c) Because the conditional distribution of Z given X = 16 does not equal the marginal distribution of Z, X and Z are not independent.

6-100. Let X, Y, and Z denote the number of calls answered in two rings or less, three or four rings, and five rings or more, respectively.

a) $P(X = 8, Y = 1, Z = 1) = \frac{10!}{8!1!1!} 0.7^8 0.25^1 0.05^1 = 0.0649$

b) Let W denote the number of calls answered in four rings or less. Then, W is a binomial random variable with n = 10 and p = 0.95.

Therefore, P(W = 10) = $\binom{10}{10} 0.95^{10} 0.05^0 = 0.5987$.

c) E(W) = 10(0.95) = 9.5.

6-101. a) $f_{Z|8}(z) = \frac{f_{XZ}(8,z)}{f_X(8)}$ and $f_{XZ}(x,z) = \frac{10!}{x!z!(10-x-z)!} 0.70^x 0.25^{(10-x-z)} 0.05^z$ for

$x + z \le 10$ and $0 \le x, 0 \le z$. Then,

$$f_{Z|8}(z) = \frac{\frac{10!}{8!z!(2-z)!} 0.70^8 0.25^{(2-z)} 0.05^z}{\frac{10!}{8!2!} 0.70^8 0.30^2} = \frac{2!}{z!(2-z)!} \left(\frac{0.25}{0.30}\right)^{2-z} \left(\frac{0.05}{0.30}\right)^z$$

for $0 \le z \le 2$. That is Z is binomial with n =2 and p = 0.05/0.30 = 1/6.

b) E(Z) given X = 8 is 2(1/6) = 1/3.

c) Because the conditional distribution of Z given X = 8 does not equal the marginal distribution of Z, X and Z are not independent.

6-102. $\int_0^3 \int_0^2 cx^2 y\,dy\,dx = \int_0^3 cx^2 \left.\frac{y^2}{2}\right|_0^2 dx = 2c \left.\frac{x^3}{3}\right|_0^3 = 18c$. Therefore, c = 1/18.

6-103. a) $P(X < 1, Y < 1) = \int_0^1 \int_0^1 \frac{1}{18} x^2 y\,dy\,dx = \int_0^1 \frac{1}{18} x^2 \left.\frac{y^2}{2}\right|_0^1 dx = \frac{1}{36} \left.\frac{x^3}{3}\right|_0^1 = \frac{1}{108}$

b) $P(X < 2.5) = \int_0^{2.5} \int_0^2 \frac{1}{18} x^2 y\,dy\,dx = \int_0^{2.5} \frac{1}{18} x^2 \left.\frac{y^2}{2}\right|_0^2 dx = \frac{1}{9} \left.\frac{x^3}{3}\right|_0^{2.5} = 0.5787$

c) $P(1 < Y < 2.5) = \int_0^3 \int_1^2 \frac{1}{18} x^2 y\,dy\,dx = \int_0^3 \frac{1}{18} x^2 \left.\frac{y^2}{2}\right|_1^2 dx = \frac{1}{12} \left.\frac{x^3}{3}\right|_0^3 = \frac{3}{4}$

d)

$$P(X > 2, 1 < Y < 1.5) = \int_2^3 \int_1^{1.5} \frac{1}{18} x^2 y\,dy\,dx = \int_2^3 \frac{1}{18} x^2 \left.\frac{y^2}{2}\right|_1^{1.5} dx = \frac{5}{144} \left.\frac{x^3}{3}\right|_2^3$$

$$= \frac{95}{432} = 0.2199$$

e) $E(X) = \int_0^3 \int_0^2 \frac{1}{18} x^3 y\,dy\,dx = \int_0^3 \frac{1}{18} x^3\, 2\,dx = \frac{1}{9} \left.\frac{x^4}{4}\right|_0^3 = \frac{9}{4}$

f) $E(Y) = \int_0^3 \int_0^2 \frac{1}{18} x^2 y^2\,dy\,dx = \int_0^3 \frac{1}{18} x^2 \frac{8}{3} dx = \frac{4}{27} \left.\frac{x^3}{3}\right|_0^3 = \frac{4}{3}$

6-104. a) $f_X(x) = \int_0^2 \frac{1}{18} x^2 y\,dy = \frac{1}{9} x^2$ for $0 < x < 3$

b) $f_{Y|X}(y) = \frac{f_{XY}(1,y)}{f_X(1)} = \frac{\frac{1}{18} y}{\frac{1}{9}} = \frac{y}{2}$ for $0 < y < 2$.

c) $f_{X|1}(x) = \frac{f_{XY}(x,1)}{f_Y(1)} = \frac{\frac{1}{18} x^2}{f_Y(1)}$ and $f_Y(y) = \int_0^3 \frac{1}{18} x^2 y\,dx = \frac{y}{2}$ for $0 < y < 2$.

Therefore, $f_{X|1}(x) = \frac{\frac{1}{18} x^2}{1/2} = \frac{1}{9} x^2$ for $0 < x < 3$.

6-105. The region $x^2 + y^2 \le 1$ and 0 < z < 4 is a cylinder of radius 1 (and base area π) and height 4. Therefore, the volume of the cylinder is 4π and $f_{XYZ}(x,y,z) = \frac{1}{4\pi}$ for $x^2 + y^2 \le 1$ and 0 < z < 4.

a) The region $X^2 + Y^2 \le 0.5$ is a cylinder of radius $\sqrt{0.5}$ and height 4. Therefore,
$P(X^2 + Y^2 \le 0.5) = \frac{4(0.5\pi)}{4\pi} = 1/2$.

b) The region $X^2 + Y^2 \le 0.5$ and 0 < z < 2 is a cylinder of radius $\sqrt{0.5}$ and height 2. Therefore,
$P(X^2 + Y^2 \le 0.5, Z < 2) = \frac{2(0.5\pi)}{4\pi} = 1/4$

c) $f_{XY|1}(x,y) = \frac{f_{XYZ}(x,y,1)}{f_Z(1)}$ and $f_Z(z) = \iint_{x^2+y^2 \le 1} \frac{1}{4\pi} dydx = 1/4$

for 0 < z < 4. Then, $f_{XY|1}(x,y) = \frac{1/4\pi}{1/4} = \frac{1}{\pi}$ for $x^2 + y^2 \le 1$.

d) $f_X(x) = \int_0^4 \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \frac{1}{4\pi} dydz = \int_0^4 \frac{1}{2\pi}\sqrt{1-x^2}\,dz = \frac{2}{\pi}\sqrt{1-x^2}$ for -1 < x < 1

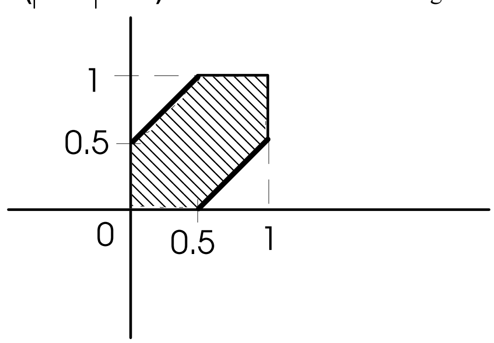
6-106. a) $f_{Z|0,0}(z) = \frac{f_{XYZ}(0,0,z)}{f_{XY}(0,0)}$ and $f_{XY}(x,y) = \int_0^4 \frac{1}{4\pi} dz = 1/\pi$ for $x^2 + y^2 \le 1$. Then,

$f_{Z|0,0}(z) = \frac{1/4\pi}{1/\pi} = 1/4$ for 0 < z < 4 and $\mu_{Z|0,0} = 2$.

b) $f_{Z|xy}(z) = \frac{f_{XYZ}(x,y,z)}{f_{XY}(x,y)} = \frac{1/4\pi}{1/\pi} = 1/4$ for 0 < z < 4. Then, E(Z) given X = x and Y = y is $\int_0^4 \frac{z}{4} dz = 2$.

6-107. $f_{XY}(x,y) = c$ for 0 < x < 1 and 0 < y < 1. Then, $\int_0^1 \int_0^1 c\,dxdy = 1$ and c = 1. Because $f_{XY}(x,y)$ is constant,

$P(|X - Y| < 0.5)$ is the area of the shaded region below



That is, $P(|X - Y| < 0.5)$ = 3/4.

6-108. a) Let $X_1, X_2, \ldots, X_6$ denote the lifetimes of the six components, respectively. Because of independence,
$P(X_1 > 5000, X_2 > 5000, \ldots, X_6 > 5000) = P(X_1 > 5000)P(X_2 > 5000)\ldots P(X_6 > 5000)$

If X is exponentially distributed with mean θ , then $\lambda = \frac{1}{\theta}$ and

$P(X > x) = \int_x^\infty \frac{1}{\theta} e^{-t/\theta} dt = -e^{-t/\theta}\Big|_x^\infty = e^{-x/\theta}$. Therefore, the answer is

$e^{-5/8}e^{-0.5}e^{-0.5}e^{-0.25}e^{-0.25}e^{-0.2} = e^{-2.325} = 0.0978$.

b) Because the component lifetimes are independent, the covariance is zero.

6-109. Let X, Y, and Z denote the number of problems that result in functional, minor, and no defects, respectively.

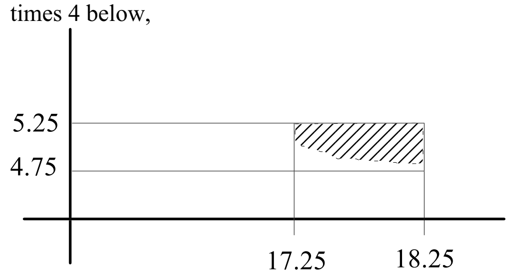
a) $P(X = 2, Y = 5) = P(X = 2, Y = 5, Z = 3) = \frac{10!}{2!5!3!} 0.2^2 0.5^5 0.3^3 = 0.085$

b) Z is binomial with n = 10 and p = 0.3.

c) E(Z) = 10(0.3) = 3.

6-110. Let $\overline{X}$ denote the mean weight of the 25 bricks in the sample. Then, E($\overline{X}$) = 3 and $\sigma_{\overline{X}} = \frac{0.25}{\sqrt{25}} = 0.05$. Then, P($\overline{X}$ < 2.95) = P(Z < $\frac{2.95-3}{0.05}$) = P (Z < -1) = 0.159.

6-111. Because $\int_{17.75}^{18.25} \int_{4.75}^{5.25} c\,dy\,dx = 0.25c$, c = 4. The area of a panel is XY and P(XY > 90) is the shaded area times 4 below,



That is, $\int_{17.75}^{18.25} \int_{90/x}^{5.25} 4\,dy\,dx = 4 \int_{17.75}^{18.25} 5.25 - \frac{90}{x} dx = 4(5.25x - 90 \ln x \Big|_{17.75}^{18.25}) = 0.499$

6-112. a) Let Y denote the total weight of the 16 candies. Then, E(Y) = 16(0.1) = 1.6 and $V(Y) = 16(0.01^2) = 0.0016$.

b) P(Y < 1.6) = P(Z < $\frac{1.6-1.6}{\sqrt{0.0016}}$) = P(Z < 0) = 0.5

c) Let X denote the weight of 17 candies. Then, E(X) = 17 (0.1) = 1.7 and $V(X) = 17(0.01^2) = 0.0017$. Also, P(Y < 1.6) = P(Z < $\frac{1.6-1.7}{\sqrt{0.0017}}$) = P(Z < -2.43) = 0.008.

6-113. Let $\overline{X}$ denote the average time to locate 10 parts. Then, E($\overline{X}$) =45 and $\sigma_{\overline{X}} = \frac{30}{\sqrt{10}}$

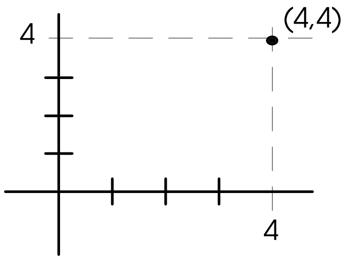
a) $P(\overline{X} > 60) = P(Z > \frac{60-45}{30/\sqrt{10}}) = P(Z > 1.58) = 0.057$

b) Let Y denote the total time to locate 10 parts. Then, Y > 600 if and only if $\overline{X}$ > 60. Therefore, the answer is the same as part a.

6-114. a) Let Y denote the weight of an assembly. Then, E(Y) = 4 + 5.5 + 10 + 8 = 27.5 and V(Y) = $0.4^2 + 0.5^2 + 0.2^2 + 0.5^2 = 0.7$. $P(Y > 29.5) = P(Z > \frac{29.5-27.5}{\sqrt{0.7}}) = P(Z > 2.39) = 0.0084$

b) Let $\overline{X}$ denote the mean weight of 8 independent assemblies. Then, E($\overline{X}$) = 27.5 and V($\overline{X}$) = 0.7/8 = 0.0875. Also, $P(\overline{X} > 29) = P(Z > \frac{29-27.5}{\sqrt{0.0875}}) = P(Z > 5.07) = 0$.

6-115. $\rho = -0.2$ implies negative slope.



6-116. Using the same method as in exercise 6-52, E(X) = 1, E(Y) = 2, E(XY) = 2.8, $E(X^2) = 2$, $E(Y^2) = 5$,
$V(X) = 2 - (1)^2 = 1$, $V(Y) = 5 - (2)^2 = 1$,
$\sigma_{xy} = E(XY) - E(X)E(Y) = 2.8 - (1)(2) = 0.8$, $\rho = 0.8/(1)(1) = 0.80$

Mind-Expanding Exercises

6-117. a) $E(Y) = P(|X - \mu| \geq c\sigma)$

b) Because $Y \leq 1$, $(X-\mu)^2 \geq (X-\mu)^2 Y$

If $|X-\mu| \geq c\sigma$, then Y = 1 and $(X-\mu)^2 Y \geq c^2\sigma^2 Y$

If $|X-\mu| < c\sigma$, then Y = 0 and $(X-\mu)^2 Y = c^2\sigma^2 Y$.

c) Because $(X-\mu)^2 \geq c^2\sigma^2 Y$, $E[(X-\mu)^2] \geq c^2\sigma^2 E(Y)$.

d) From part a., $E(Y) = P(|X-\mu| \geq c\sigma)$. From part c., $\sigma^2 \geq c^2\sigma^2 P(|X-\mu| \geq c\sigma)$. Therefore,

$$\frac{1}{c^2} \geq P(|X-\mu| \geq c\sigma).$$

6-118. By the independence,

$$P(X_1 \in A_1, X_2 \in A_2, \ldots, X_p \in A_p) = \int_{A_1}\int_{A_2}\ldots\int_{A_p} f_{X_1}(x_1) f_{X_2}(x_2)\ldots f_{X_p}(x_p)\,dx_1 dx_2 \ldots dx_p$$

$$= \left[\int_{A_1} f_{X_1}(x_1)dx_1\right]\left[\int_{A_2} f_{X_2}(x_2)dx_2\right]\ldots\left[\int_{A_p} f_{X_p}(x_p)dx_p\right]$$

$$= P(X_1 \in A_1)P(X_2 \in A_2)\ldots P(X_p \in A_p)$$

6-119. $E(Y) = c_1\mu_1 + c_2\mu_2 + \ldots + c_p\mu_p$. Also,

$$V(Y) = \int \left[c_1x_1 + c_2x_2 + \ldots + c_px_p - (c_1\mu_1 + c_2\mu_2 + \ldots + c_p\mu_p)\right]^2 f_{X_1}(x_1)f_{X_2}(x_2)\ldots f_{X_p}(x_p)\,dx_1dx_2\ldots dx_p$$

$$= \int \left[c_1(x_1 - \mu_1) + \ldots + c_p(x_p - \mu_p)\right]^2 f_{X_1}(x_1)f_{X_2}(x_2)\ldots f_{X_p}(x_p)\,dx_1dx_2\ldots dx_p$$

Now, the crossterm

$$\int c_1c_2(x_1-\mu_1)(x_2-\mu_2) f_{X_1}(x_1)f_{X_2}(x_2)\ldots f_{X_p}(x_p)\,dx_1dx_2\ldots dx_p$$

$$= c_1c_2\left[\int (x_1-\mu_1)f_{X_1}(x_1)dx_1\right]\left[\int (x_2-\mu_2)f_{X_2}(x_2)dx_2\right] = 0$$

from the definition of the mean. Therefore, each crossterm in the last integral for V(Y) is zero and

$$V(Y) = \left[\int c_1^2(x_1-\mu_1)^2 f_{X_1}(x_1)dx_1\right]\ldots\left[\int c_p^2(x_p-\mu_p)^2 f_{X_p}(x_p)dx_p\right]$$

$$= c_1^2 V(X_1) + \ldots + c_p^2 V(X_p).$$

6-120. $\int_0^a \int_0^b f_{XY}(x,y)dydx = \int_0^a \int_0^b cdydx = cab$. Therefore, c = 1/ab. Then, $f_X(x) = \int_0^b cdy = \frac{1}{a}$ for 0 < x < a, and

$f_Y(y) = \int_0^a cdx = \frac{1}{b}$ for 0 < y < b. Therefore, $f_{XY}(x,y) = f_X(x)f_Y(y)$ for all x and y and X and Y are independent.

6-121. $f_X(x) = \int_0^b g(x)h(y)dy = g(x)\int_0^b h(y)dy = ag(x)$ where $a = \int_0^b h(y)dy$. Also,

$f_Y(y) = bh(y)$ where $b = \int_0^a g(x)dx$. Because $f_{XY}(x,y)$ is a probability density function,

$\int_0^a \int_0^b g(x)h(y)dydx = \left[\int_0^a g(x)dx\right]\left[\int_0^b h(y)dy\right] = 1$. Therefore, ab = 1 and $f_{XY}(x,y) = f_X(x)f_Y(y)$ for all x and y.

CHAPTER 7

Section 7-3

7-1. $$E\left(\overline{X}_1\right) = E\left(\frac{\sum_{i=1}^{2n} X_i}{2n}\right) = \frac{1}{2n}E\left(\sum_{i=1}^{2n} X_i\right) = \frac{1}{2n}\left(2n\mu\right) = \mu$$

$$E\left(\overline{X}_2\right) = E\left(\frac{\sum_{i=1}^{n} X_i}{n}\right) = \frac{1}{n}E\left(\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\left(n\mu\right) = \mu ,$$ $\overline{X}_1$ and $\overline{X}_2$ are unbiased estimators of μ.

The variances are $V\left(\overline{X}_1\right) = \frac{\sigma^2}{2n}$ and $V\left(\overline{X}_2\right) = \frac{\sigma^2}{n}$; compare the MSE (variance in this case),

$$\frac{MSE(\hat{\theta}_1)}{MSE(\hat{\theta}_2)} = \frac{\sigma^2 / 2n}{\sigma^2 / n} = \frac{n}{2n} = \frac{1}{2}$$

Since both estimators are unbiased, examination of the variances would conclude that $\overline{X}_1$ is the "better" estimator with the smaller variance.

7-2. $$E\left(\hat{\theta}_1\right) = \frac{1}{7}\left[E(X_1) + E(X_2) + \cdots + E(X_7)\right] = \frac{1}{7}(7E(X)) = \frac{1}{7}(7\mu) = \mu$$

$$E\left(\hat{\theta}_2\right) = \frac{1}{2}\left[E(2X_1) + E(X_6) + E(X_7)\right] = \frac{1}{2}[2\mu - \mu + \mu] = \mu$$

a) Both $\hat{\theta}_1$ and $\hat{\theta}_2$ are unbiased estimates of μ since the expected values of these statistics are equivalent to the true mean, μ.

b) $$V\left(\hat{\theta}_1\right) = V\left[\frac{X_1 + X_2 + ... + X_7}{7}\right] = \frac{1}{7^2}\left(V(X_1) + V(X_2) + \cdots + V(X_7)\right) = \frac{1}{49}(7\sigma^2) = \frac{1}{7}\sigma^2$$

$$V(\hat{\theta}_1) = \frac{\sigma^2}{7}$$

$$V\left(\hat{\theta}_2\right) = V\left[\frac{2X_1 - X_6 + X_4}{2}\right] = \frac{1}{2^2}\left(V(2X_1) + V(X_6) + V(X_4)\right) = \frac{1}{4}(4V(X_1) + V(X_6) + V(X_4))$$

$$= \frac{1}{4}\left(4\sigma^2 + \sigma^2 + \sigma^2\right)$$

$$= \frac{1}{4}(6\sigma^2)$$

$$V(\hat{\theta}_2) = \frac{3\sigma^2}{2}$$

Since both estimators are unbiased, the variances can be compared to decide which is the better estimator. The variance of $\hat{\theta}_1$ is smaller than that of $\hat{\theta}_2$, $\hat{\theta}_1$ is the better estimator.

7-3. Since both $\hat{\theta}_1$ and $\hat{\theta}_2$ are unbiased, the variances of the estimators can be examined to determine which is the "better" estimator. The variance of $\hat{\theta}_2$ is smaller than that of $\hat{\theta}_1$ thus $\hat{\theta}_2$ may be the better estimator.

$$\text{Relative Efficiency} = \frac{MSE(\hat{\theta}_1)}{MSE(\hat{\theta}_2)} = \frac{V(\hat{\theta}_1)}{V(\hat{\theta}_2)} = \frac{10}{4} = 2.5$$

7-4. Since both estimators are unbiased:

$$\text{Relative Efficiency} = \frac{MSE(\hat{\theta}_1)}{MSE(\hat{\theta}_2)} = \frac{V(\hat{\theta}_1)}{V(\hat{\theta}_2)} = \frac{\sigma^2/7}{3\sigma^2/7} = \frac{2}{21}$$

7-5. $$\frac{MSE(\hat{\theta}_1)}{MSE(\hat{\theta}_2)} = \frac{V(\hat{\theta}_1)}{V(\hat{\theta}_2)} = \frac{10}{4} = 2.5$$

7-6. $E(\hat{\theta}_1) = \theta \qquad E(\hat{\theta}_2) = \theta/2$

$$\text{Bias} = E(\hat{\theta}_2) - \theta = \frac{\theta}{2} - \theta = -\frac{\theta}{2}$$

$V(\hat{\theta}_1) = 10 \qquad V(\hat{\theta}_2) = 4$

For unbiasedness, use $\hat{\theta}_1$ since it is the only unbiased estimator.
As for minimum variance and efficiency we have:

$$\text{Relative Efficiency} = \frac{(V(\hat{\theta}_1) + \text{Bias}^2)_1}{(V(\hat{\theta}_2) + \text{Bias}^2)_2}$$ where, Bias for θ_1 is 0.

Thus,

$$\text{Relative Efficiency} = \frac{(10+0)}{\left(4 + \left(\frac{-\theta}{2}\right)^2\right)} = \frac{40}{\left(16+\theta^2\right)}$$

If the relative efficiency is less than or equal to 1, $\hat{\theta}_1$ is the better estimator.

Use $\hat{\theta}_1$, when $\frac{40}{(16+\theta^2)} \le 1$

$$40 \le (16+\theta^2)$$
$$24 \le \theta^2$$
$$\theta \le -4.899 \text{ or } \theta \ge 4.899$$

If $-4.899 < \theta < 4.899$ then use $\hat{\theta}_2$.

For unbiasedness, use $\hat{\theta}_1$. For efficiency, use $\hat{\theta}_1$ when $\theta \le -4.899$ or $\theta \ge 4.899$ and use $\hat{\theta}_2$ when $-4.899 < \theta < 4.899$.

7-7.

$E(\hat{\theta}_1) = \theta$	No bias	$V(\hat{\theta}_1) = 12 = MSE(\hat{\theta}_1)$	
$E(\hat{\theta}_2) = \theta$	No bias	$V(\hat{\theta}_2) = 10 = MSE(\hat{\theta}_2)$	
$E(\hat{\theta}_3) \ne \theta$	Bias	$V(\hat{\theta}_3) = 6$	includes (bias^2)

To compare the three estimators, calculate the relative efficiencies:

$\frac{MSE(\hat{\theta}_1)}{MSE(\hat{\theta}_2)} = \frac{12}{10} = 1.2$, since rel. eff. > 1 use $\hat{\theta}_2$ as the estimator for θ

$\frac{MSE(\hat{\theta}_1)}{MSE(\hat{\theta}_3)} = \frac{12}{6} = 2$, since rel. eff. > 1 use $\hat{\theta}_3$ as the estimator for θ

$\frac{MSE(\hat{\theta}_2)}{MSE(\hat{\theta}_3)} = \frac{10}{6} = 1.8$, since rel. eff. > 1 use $\hat{\theta}_3$ as the estimator for θ

Conclusion:

$\hat{\theta}_3$ is the most efficient estimator with bias. $\hat{\theta}_2$ is would be the best “unbiased” estimator.

7-8. $n_1 = 20, n_2 = 10, n_3 = 8$

Show that S^2 is unbiased:

$$E\left(S^2\right) = E\left(\frac{20S_1^2 + 10S_2^2 + 8S_3^2}{38}\right)$$

$$= \frac{1}{38}\left(E\left(20S_1^2\right) + E\left(10S_2^2\right) + E\left(8S_3^2\right)\right)$$

$$= \frac{1}{38}\left(20\sigma_1^2 + 10\sigma_2^2 + 8\sigma_3^2\right)$$

$$= \frac{1}{38}\left(38\sigma^2\right)$$

$$= \sigma^2$$

$\therefore$ S^2 is an unbiased estimator of σ^2 .

7-9. Show that $\dfrac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}{n}$ is a biased estimator of σ^2 :

a)

$$E\left(\frac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}{n}\right)$$

$$= \frac{1}{n}E\left(\sum_{i=1}^{n}X_i - n\overline{X}\right)^2$$

$$= \frac{1}{n}\left(\sum_{i=1}^{n}E\left(X_i^2\right) - nE\left(\overline{X}^2\right)\right)$$

$$= \frac{1}{n}\left(\sum_{i=1}^{n}\left(\mu^2 + \sigma^2\right) - n\left(\mu^2 + \frac{\sigma^2}{n}\right)\right)$$

$$= \frac{1}{n}\left(n\mu^2 + n\sigma^2 - n\mu^2 - \sigma^2\right)$$

$$= \frac{1}{n}((n-1)\sigma^2)$$

$$= \sigma^2 - \frac{\sigma^2}{n}$$

$\therefore$ $\dfrac{\sum\left(X_i - \overline{X}\right)^2}{n}$ is a biased estimator of σ^2 .

b) $\text{Bias} = E\left[\dfrac{\sum\left(X_i^2 - n\overline{X}\right)^2}{n}\right] - \sigma^2 = \sigma^2 - \dfrac{\sigma^2}{n} - \sigma^2 = -\dfrac{\sigma^2}{n}$

c) Bias decreases as n increases.

7-10. Show that $\overline{X}^2$ is a biased estimator of μ. Using $E\left(X^2\right) = V(X) + [E(X)]^2$

$$E\left(\overline{X}^2\right) = \frac{1}{n^2} E\left(\sum_{i=1}^{n} X_i\right)^2$$

$$= \frac{1}{n^2}\left(V\left(\sum_{i=1}^{n} X_i\right) + \left[E\left(\sum_{i=1}^{n} X_i\right)\right]^2\right)$$

$$= \frac{1}{n^2}\left(n\sigma^2 + \left(\sum_{i=1}^{n} \mu\right)^2\right)$$

$$= \frac{1}{n^2}\left(n\sigma^2 + (n\mu)^2\right)$$

$$= \frac{1}{n^2}\left(n\sigma^2 + n^2\mu^2\right)$$

$$E\left(\overline{X}^2\right) = \frac{\sigma^2}{n} + \mu^2$$

$\therefore$ $\overline{X}^2$ is a biased estimator of μ.

b) Bias = $E\left(\overline{X}^2\right) - \mu^2 = \frac{\sigma^2}{n} + \mu^2 - \mu^2 = \frac{\sigma^2}{n}$

c) Bias decreases as n increases.

7-11. Using B=100, and n=8, an estimate of $\hat{\lambda} = 6.2$ was found with standard error of 0.90 (Using Equation 7-3). Note: the estimates will vary.

7-12. An estimate of the 10% trimmed mean was found to be 51.3. The standard error of this estimate was found to be 2.46 (Using Equation 7-3). Note: The original data points were treated as the population. Randomly select 100 samples of size n.

7-13. An estimate of the population standard deviation was found to be 6.14 with a standard error of 1.37 (Using Eq. 7-3).

7-14. An estimate of the population range was found to be 16.9 with a standard error of 4.21 (Using Eq. 7-3).

Section 7-4

7-15. $f(x) = \frac{e^{-\lambda}\lambda^x}{x!}$ $\quad L(\lambda) = \prod_{i=1}^{n} \frac{e^{-\lambda}\lambda^{x_i}}{x_i!} = \frac{e^{-n\lambda}\lambda^{\sum_{i=1}^{n} x_i}}{\prod_{i=1}^{n} x_i!}$

$$\ln L(\lambda) = -n\lambda \ln e + \sum_{i=1}^{n} x_i \ln\lambda - \sum_{i=1}^{n} \ln x_i!$$

$$\frac{d\ln L(\lambda)}{d\lambda} = -n + \frac{1}{\lambda}\sum_{i=1}^{n} x_i \equiv 0$$

$$= -n + \frac{\sum_{i=1}^{n} x_i}{\lambda} = 0$$

$$\sum_{i=1}^{n} x_i = n\lambda$$

$$\hat{\lambda} = \frac{\sum_{i=1}^{n} x_i}{n}$$

7-16. a) The joint pdf is given by

$$f(x_1,\dots,x_n;\lambda,\theta) = \lambda^n e^{-\lambda\sum(x_i-\theta)} \quad \text{for} \quad x_1 \ge \theta, \dots x_n \ge \theta$$
$$= 0 \qquad \text{otherwise}$$

where $x_1 \ge \theta, \dots x_n \ge \theta$ only if $\min(x_i) \ge \theta$. The exponent can be written as $-\lambda\sum(x_i-\theta) = -\lambda\sum x_i + n\lambda\theta$.

The likelihood is found to be: $L(\lambda,\theta) = \lambda^n e^{(-\lambda\sum x_i)(\exp(n\lambda\theta))}$ for $\min(x_i) \ge \theta$. The likelihood is 0 otherwise.

As long as $\min(x_i) \ge \theta$, maximizing the likelihood with respect to θ would increase the likelihood since the exponent is positive. Thus, the mle of θ is $\hat{\theta} = \min(x_i)$.

The log likelihood is $\ln L(\lambda,\theta) = n\ln(\lambda) - \lambda\sum(x_i-\hat{\theta}) = n\ln(\lambda) - \lambda\sum(x_i - \min(x_i))$. Finally,

$$\frac{\partial L(\lambda,\theta)}{\partial\lambda} = \frac{n}{\lambda} - \sum(x_i-\hat{\theta}) = 0$$

and

$$\hat{\lambda} = \frac{n}{\sum(x_i-\hat{\theta})}$$

b) Example: Consider traffic flow and let the time that has elapsed between one car passing a fixed point and the instant that the next car begins to pass that point be considered time headway. This headway can be modeled by the shifted exponential distribution.

Example in Reliability: Consider a process where failures are of interest. Say that a sample or population is put into operation at x = 0, but no failures will occur until θ period of operation. Failures will occur after the time θ.

7-17.

$$f(x) = p(1-p)^{x-1}$$

$$L(p) = \prod_{i=1}^{n} p(1-p)^{x_i-1}$$

$$= p^n(1-p)^{\sum_{i=1}^{n} x_i - n}$$

$$\ln L(p) = n\ln p + \left(\sum_{i=1}^{n} x_i - n\right)\ln(1-p)$$

$$\frac{\partial \ln L(p)}{\partial p} = \frac{n}{p} - \frac{\sum_{i=1}^{n} x_i - n}{1-p} \equiv 0$$

$$0 = \frac{(1-p)n - p\left(\sum_{i=1}^{n} x_i - n\right)}{p(1-p)}$$

$$0 = \frac{n - np - p\sum_{i=1}^{n} x_i + pn}{p(1-p)}$$

$$0 = n - \sum_{i=1}^{n} x_i$$

$$\hat{P} = \frac{n}{\sum_{i=1}^{n} x_i}$$

7-18. $f(x) = (\alpha+1)x^{\alpha}$

$$L(\alpha) = \prod_{i=1}^{n}(\alpha+1)x_i^{\alpha} = (\alpha+1)x_1^{\alpha} \times (\alpha+1)x_2^{\alpha} \times ...$$

$$= (\alpha+1)^n \prod_{i=1}^{n} x_i^{\alpha}$$

$$\ln L(\alpha) = n\ln(\alpha+1) + \alpha \ln x_1 + \alpha \ln x_2 + ...$$

$$= n\ln(\alpha+1) + \alpha \sum_{i=1}^{n} \ln x_i$$

$$\frac{\partial \ln L(\alpha)}{\partial \alpha} = \frac{n}{\alpha+1} + \sum_{i=1}^{n} \ln x_i = 0$$

$$\frac{n}{\alpha+1} = -\sum_{i=1}^{n} \ln x_i$$

$$\alpha = \frac{n}{-\sum_{i=1}^{n} \ln x_i} - 1$$

7-19. a)

$$\ln L(\beta,\delta) = \sum_{i=1}^{n} \ln\left[\frac{\beta}{\delta}\left(\frac{x_i}{\delta}\right)^{\beta-1} e^{-(x_i/\beta)^{\beta}}\right]$$

$$= n\ln\left(\frac{\beta}{\delta}\right) + (\beta-1)\sum \ln\left(\frac{x_i}{\delta}\right) - \sum\left(\frac{x_i}{\delta}\right)^{\beta}$$

b)

$$\frac{\partial \ln L(\beta,\delta)}{\partial \beta} = \frac{n}{\beta} + \sum \ln\left(\frac{x_i}{\delta}\right) - \sum \ln\left(\frac{x_i}{\delta}\right)\left(\frac{x_i}{\delta}\right)^{\beta}$$

$$\frac{\partial \ln L(\beta,\delta)}{\partial \delta} = -\frac{n}{\delta} - (\beta-1)\frac{n}{\delta} + \beta\frac{\sum x_i^{\beta}}{\delta^{\beta+1}}$$

Upon setting $\frac{\partial \ln L(\beta,\delta)}{\partial \delta}$ equal to zero, we obtain

$$\delta^{\beta} n = \sum x_i^{\beta} \quad \text{and} \quad \delta = \left[\frac{\sum x_i^{\beta}}{n}\right]^{1/\beta}$$

Upon setting $\frac{\partial \ln L(\beta,\delta)}{\partial \beta}$ equal to zero and substituting for δ, we obtain

$$\frac{n}{\beta} + \sum \ln x_i - n\ln\delta = \frac{1}{\delta^{\beta}}\sum x_i^{\beta}(\ln x_i - \ln\delta)$$

$$\frac{n}{\beta} + \sum \ln x_i - \frac{n}{\beta}\ln\left(\frac{\sum x_i^{\beta}}{n}\right) = \frac{n}{\sum x_i^{\beta}}\sum x_i^{\beta}\ln x_i - \frac{n}{\sum x_i^{\beta}}\sum x_i^{\beta}\frac{1}{\beta}\ln\left(\frac{\sum x_i^{\beta}}{n}\right)$$

and $\frac{1}{\beta} = \left[\frac{\sum x_i^{\beta}\ln x_i}{\sum x_i^{\beta}} + \frac{\sum \ln x_i}{n}\right]$

c) Numerical iteration is required.

7-20. a) $L(\lambda,r)=\prod_{i=1}^{n}\frac{\lambda^r x_i^{r-1}e^{-\lambda x_i}}{\Gamma(r)}$ $\ln L(\lambda,r)=nr\ln\lambda+(r-1)\sum x_i-\lambda\sum x_i-n\ln\Gamma(r)$

b) $\frac{\partial \ln L(\lambda,r)}{\partial\lambda}=\frac{nr}{\lambda}-\sum x_i$

$\frac{\partial \ln L(\lambda,r)}{\partial r}=n\ln\lambda+\sum x_i-\frac{n\Gamma'(r)}{\Gamma(r)}$

where $\Gamma(r)=\int_0^\infty x^{r-1}e^{-x}dx$. Upon setting each derivative equal to zero, the maximum likelihood equations are $\frac{r}{\lambda}=\bar{x}$ $\ln\lambda+\bar{x}=\frac{\Gamma'(r)}{\Gamma(r)}$

c) From the first equation in part b., the solutions for r and λ satisfy $\frac{r}{\lambda}=\bar{x}$. Therefore, $\hat{\mu}=\bar{x}$.

7-21. a) Because $\hat{a}$ is less than or equal to "a" for every sample, E($\hat{a}$) cannot equal "a".

b) Yes, E($\hat{a}$) is less than "a" by a factor of $\frac{n}{n+1}$. As $n\to\infty$, $\frac{n}{n+1}\to 1$, and E($\hat{a}$) → a.

c) $\hat{a}\frac{(n+1)}{n}$, because $E\left[\hat{a}\frac{(n+1)}{n}\right]=\frac{n+1}{a}E(\hat{a})=a$.

d) $F_Y(y)=P(Y\le y)=P(X_1\le y, X_2\le y,\ldots, X_n\le y)$
$=P(X_1\le y)P(X_2\le y)\ldots P(X_n\le y)$
$=(y/a)^n$ for $0\le y\le a$.

$$f(y)=\frac{\partial F_Y(y)}{\partial y}=\frac{ny^{n-1}}{a^n}\quad \text{for}\quad 0\le y\le a$$
$$=0\quad\text{otherwise}$$

The maximum likelihood estimator for a is Y. To show that the mle for a is biased, need to show that E(Y) ≠ a:

$$E(Y)=\int_0^a y\frac{ny^{n-1}}{a^n}=\left.\frac{ny^n}{a^n(n+1)}\right|_0^a=\frac{n}{n+1}a.$$

Thus, E(Y) ≠ a, and the mle of a is biased.

Section 7-6

7-22. $P(1.009\le\bar{X}\le 1.012)=P\left(\frac{1.009-1.01}{0.003/\sqrt{9}}\le\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\le\frac{1.012-1.01}{0.003/\sqrt{9}}\right)$

$=P(-1\le Z\le 2)=P(Z\le 2)-P(Z\le -1)$

$=0.9772-0.1587=0.8385$

7-23. $X_i\sim N(100,10^2)$ $n=25$

$\mu_{\bar{X}}=100$ $\sigma_{\bar{X}}=\frac{\sigma}{\sqrt{n}}=\frac{10}{\sqrt{25}}=2$

$P[(100-1.8(2))\le\bar{X}\le(100+2)]=P(96.4\le\bar{X}\le 102)$

$=P\left(\frac{96.4-100}{2}\le\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\le\frac{102-100}{2}\right)$

$=P(-1.8\le Z\le 1)=P(Z\le 1)-P(Z\le -1.8)$

$=0.8413-0.0359=0.8054$

7-24. $\mu_{\overline{X}} = 75.5\text{psi}$ $\sigma_{\overline{X}} = \dfrac{\sigma}{\sqrt{n}} = \dfrac{3.5}{\sqrt{6}} = 1.429$

$$P(\overline{X} \geq 75.75) = P\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \geq \frac{75.75-75.5}{1.429}\right)$$

$$= P(Z \geq 0.175) = 1 - P(Z \leq 1.75)$$

$$= 1 - 0.56945 = 0.43055$$

7-25.

n = 6	n = 49
$\sigma_{\overline{X}} = \dfrac{\sigma}{\sqrt{n}} = \dfrac{3.5}{\sqrt{6}}$	$\sigma_{\overline{X}} = \dfrac{\sigma}{\sqrt{n}} = \dfrac{3.5}{\sqrt{49}}$
$= 1.429$	$= 0.5$
$\sigma_{\overline{X}}$ is reduced by 0.929 psi	

7-26. Assuming a normal distribution,

$\mu_{\overline{X}} = 2500$ $\sigma_{\overline{X}} = \dfrac{\sigma}{\sqrt{n}} = \dfrac{50}{\sqrt{5}} = 22.361$

$$P(2490 \leq \overline{X} \leq 2510) = P\left(\frac{2490-2500}{22.361} \leq \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \leq \frac{2510-2500}{22.361}\right)$$

$$= P(-0.45 \leq Z \leq 0.45) = P(Z \leq 0.45) - P(Z \leq -0.45)$$

$$= 0.6736 - 0.3264 = 0.3472$$

7-27. $\sigma_{\overline{X}} = \dfrac{\sigma}{\sqrt{n}} = \dfrac{50}{\sqrt{5}} = 22.361\text{psi}$ = standard error of $\overline{X}$

7-28. $\sigma^2 = 25$

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}$$

$$n = \left(\frac{\sigma}{\sigma_{\overline{X}}}\right)^2$$

$$= \left(\frac{5}{1.5}\right)^2$$

$$= 11.11 \sim 12$$

7-29. Let $Y = \overline{X} - 6$

$$\mu_X = \frac{a+b}{2} = \frac{(0+1)}{2} = \frac{1}{2}$$

$$\mu_{\overline{X}} = \mu_X$$

$$\sigma_X^2 = \frac{(b-a)^2}{12} = \frac{1}{12}$$

$$\sigma_{\overline{X}}^2 = \frac{\sigma^2}{n} = \frac{\frac{1}{12}}{12} = \frac{1}{144}$$

$$\sigma_{\overline{X}} = \frac{1}{12}$$

$$\mu_Y = \frac{1}{2} - 6 = -5\frac{1}{2}$$

$$\sigma_Y^2 = \frac{1}{144}$$

$Y = \overline{X} - 6 \sim N(-5\frac{1}{2}, \frac{1}{144})$, approximately, using the central limit theorem.

7-30. n = 36

$$\mu_X = \frac{a+b}{2} = \frac{(3+1)}{2} = 2$$

$$\sigma_X = \sqrt{\frac{(b-a+1)^2 - 1}{12}} = \sqrt{\frac{(3-1+1)^2 - 1}{12}} = \sqrt{\frac{8}{12}} = \sqrt{\frac{2}{3}}$$

$$\mu_{\overline{X}} = 2, \sigma_{\overline{X}} = \frac{\sqrt{2/3}}{\sqrt{36}} = \frac{\sqrt{2/3}}{6}$$

$$z = \frac{\overline{X} - \mu}{\sigma / \sqrt{n}}$$

Using the central limit theorem:

$$P(2.1 < \overline{X} < 2.5) = P\left(\frac{2.1-2}{\frac{\sqrt{2/3}}{6}} < Z < \frac{2.5-2}{\frac{\sqrt{2/3}}{6}}\right)$$

$$= P(0.7348 < Z < 3.6742)$$

$$= P(Z < 3.6742) - P(Z < 0.7348)$$

$$= 1 - 0.7688 = 0.2312$$

7-31.

$\mu_X = 8.2$ minutes	n = 49
$\sigma_X = 1.5$ minutes	$\sigma_{\overline{X}} = \frac{\sigma_X}{\sqrt{n}} = \frac{1.5}{\sqrt{49}} = 0.2143$ $\mu_{\overline{X}} = \mu_X = 8.2$ mins

Using the central limit theorem, $\overline{X}$ is approximately normally distributed.

a) $P(\overline{X} < 10) = P(Z < \frac{10 - 8.2}{0.2143}) = P(Z < 8.4) = 1$

b) $P(5 < \overline{X} < 10) = P(\frac{5-8.2}{0.2143} < Z < \frac{10-8.2}{0.2143})$

$$= P(Z < 8.4) - P(Z < -14.932) = 1 - 0 = 1$$

c) $P(\overline{X} < 6) = P(Z < \frac{6 - 8.2}{0.2143}) = P(Z < -10.27) = 0$

7-32.

$n_1 = 16$	$n_2 = 9$	$\overline{X}_1 - \overline{X}_2 \sim N(\mu_{\overline{X}_1} - \mu_{\overline{X}_2}, \sigma^2_{\overline{X}_1} + \sigma^2_{\overline{X}_2})$
$\mu_1 = 75$	$\mu_2 = 70$	$\sim N(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})$
$\sigma_1 = 8$	$\sigma_2 = 12$	$\sim N(75 - 70, \frac{8^2}{16} + \frac{12^2}{9})$
		$\sim N(5, 20)$

a) $P(\overline{X}_1 - \overline{X}_2 > 4)$

$$P(Z > \frac{4-5}{\sqrt{20}}) = P(Z > -0.2236) = 1 - P(Z \le -0.2236)$$
$$= 1 - 0.4115 = 0.5885$$

b) $P(3.5 \le \overline{X}_1 - \overline{X}_2 \le 5.5)$

$$P(\frac{3.5-5}{\sqrt{20}} \le Z \le \frac{5.5-5}{\sqrt{20}}) = P(Z \le 0.1118) - P(Z \le -0.3354)$$
$$= 0.5445 - 0.3686 = 0.1759$$

7-33. If $\mu_B = \mu_A$, then $\overline{X}_B - \overline{X}_A$ is approximately normal with mean 0 and variance $\frac{\sigma_B^2}{25} + \frac{\sigma_A^2}{25} = 20.48$.

Then, $P(\overline{X}_B - \overline{X}_A > 3.5) = P(Z > \frac{3.5-0}{\sqrt{20.48}}) = P(Z > 0.773) = 0.221$

The probability that $\overline{X}_B$ exceeds $\overline{X}_A$ by 3.5 or more is not that unusual when μ_B and μ_A are equal. Therefore, there is not strong evidence that μ_B is greater than μ_A.

7-34. Assume approximate normal distributions.

$$(\overline{X}_{high} - \overline{X}_{low}) \sim N(60 - 55, \frac{4^2}{16} + \frac{4^2}{16})$$
$$\sim N(5, 2)$$
$$P(\overline{X}_{high} - \overline{X}_{low} \ge 2) = P(Z \ge \frac{2-5}{\sqrt{2}}) = 1 - P(Z \le -2.12) = 1 - 0.0170 = 0.983$$

Supplemental Exercises

7-35. $f(x_1, x_2, x_3, x_4, x_5) = \frac{1}{\sqrt{2\pi}\sigma} \sum_{i=1}^{5} e^{-\frac{(x_i - \mu)^2}{2\sigma^2}}$

7-36. $f(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} \lambda e^{-\lambda x_i}$ for $x_1 > 0, x_2 > 0, \ldots, x_n > 0$

7-37. $f(x_1, x_2, x_3, x_4) = 1$ for $0 \le x_1 \le 1, 0 \le x_2 \le 1, 0 \le x_3 \le 1, 0 \le x_4 \le 1$

7-38.

$$\overline{X}_1 - \overline{X}_2 \sim N(100 - 105, \frac{1.5^2}{25} + \frac{2^2}{25})$$
$$\sim N(-5, 0.2233)$$

7-39. $\sigma_{\overline{X}_1-\overline{X}_2} = \sqrt{0.2233} = 0.4726$

7-40. $X \sim N(50,144)$

$$P(47 \le \overline{X} \le 53) = P\left(\frac{47-50}{12/\sqrt{36}} \le Z \le \frac{53-50}{12/\sqrt{36}}\right)$$
$$= P(-1.5 \le Z \le 1.5)$$
$$= P(Z \le 1.5) - P(Z \le -1.5)$$
$$= 0.9332 - 0.0668 = 0.8664$$

7-41. No, because Central Limit Theorem states that with large samples ($n \ge 30$), $\overline{X}$ is approximately normally distributed.

7-42. Assume $\overline{X}$ is approximately normally distributed.

$$P(\overline{X} > 4985) = 1 - P(\overline{X} \le 4985) = 1 - P(Z \le \frac{4985-5500}{100/\sqrt{9}})$$
$$= 1 - P(Z \le -15.45) = 1 - 0 = 1$$

7-43. $t = \dfrac{\overline{X} - \mu}{s/\sqrt{n}} = \dfrac{52-50}{1.5/\sqrt{16}} = 5.333$

$t_{.05,15} = 1.753$. Since 5.33 >> $t_{.05,15}$, the results are very unusual.

7-44. $P(\overline{X} \le 37) = P(Z \le -5.36) = 0$

7-45. Binomial with p equal to the proportion of defective chips and n = 100.

7-46.

$$V(\overline{X}) = V[a\overline{X}_1 + (1-a)\overline{X}_2]$$
$$= a^2 V(\overline{X}_1) + (1-a)^2 V(\overline{X}_2)$$
$$= a^2(\frac{\sigma^2}{n_1}) + (1 - 2a + a^2)(\frac{\sigma^2}{n_2})$$
$$= \frac{a^2\sigma^2}{n_1} + \frac{\sigma^2}{n_2} - \frac{2a\sigma^2}{n_2} + \frac{a^2\sigma^2}{n_2}$$
$$= (n_2 a^2 + n_1 - 2n_1 a + n_1 a^2)(\frac{\sigma^2}{n_1 n_2})$$
$$\frac{\partial V(\overline{X})}{\partial a} = (\frac{\sigma^2}{n_1 n_2})(2n_2 a - 2n_1 + 2n_1 a) \equiv 0$$
$$0 = 2n_2 a - 2n_1 + 2n_1 a$$
$$2a(n_2 + n_1) = 2n_1$$
$$a(n_2 + n_1) = n_1$$
$$a = \frac{n_1}{(n_2 + n_1)}$$

Mind-Expanding Exercises

7-47. $P(X_1 = 0, X_2 = 0) = \frac{M(M-1)}{N(N-1)}$

$P(X_1 = 0, X_2 = 1) = \frac{M(N-M)}{N(N-1)}$

$P(X_1 = 1, X_2 = 0) = \frac{(N-M)M}{N(N-1)}$

$P(X_1 = 1, X_2 = 1) = \frac{(N-M)(N-M-1)}{N(N-1)}$

$P(X_1 = 0) = M/N$

$P(X_1 = 1) = \frac{N-M}{N}$

$P(X_2 = 0) = P(X_2 = 0|X_1 = 0)P(X_1 = 0) + P(X_2 = 0|X_1 = 1)P(X_1 = 1)$

$= \frac{M-1}{N-1} \times \frac{M}{N} + \frac{M}{N-1} \times \frac{N-M}{N} = \frac{M}{N}$

$P(X_2 = 1) = P(X_2 = 1|X_1 = 0)P(X_1 = 0) + P(X_2 = 1|X_1 = 1)P(X_1 = 1)$

$= \frac{N-M}{N-1} \times \frac{M}{N} + \frac{N-M-1}{N-1} \times \frac{N-M}{N} = \frac{N-M}{N}$

Because $P(X_2 = 0|X_1 = 0) = \frac{M-1}{N-1}$ is not equal to $P(X_2 = 0) = \frac{M}{N}$, X_1 and X_2 are not independent.

7-48. Because $\frac{(n-1)S^2}{\sigma^2}$ has a chi-square distribution with (n-1) degrees of freedom,

V[$\frac{(n-1)S^2}{\sigma^2}$] = 2 (n-1). Therefore, $V(S^2) = \frac{2\sigma^4}{n-1}$ and this decreases as n increases.

7-49. If X has an $F_{\nu,u}$ distribution, then $\alpha = P(X > f_{\alpha,\nu,u}) = P(\frac{Y/\nu}{W/u} > f_{\alpha,\nu,u}) = P(\frac{W/u}{Y/\nu} < \frac{1}{f_{\alpha,\nu,u}})$

That is, the probability a $F_{u,\nu}$ random variable is less than $\frac{1}{f_{\alpha,\nu,u}}$ is α. Therefore, $f_{1-\alpha,u,\nu} = \frac{1}{f_{\alpha,\nu,u}}$

7-50. a) $c_n = \frac{\Gamma[\frac{n-1}{2}]}{\Gamma\left(\frac{n}{2}\right)\sqrt{\frac{2}{n-1}}}$

b)

n	c_n
10	$\frac{\Gamma(4.5)}{\Gamma(5)\sqrt{2/9}} = \frac{11.6317}{4\sqrt{2/9}} = 1.028$
25	$\frac{\Gamma(12)}{\Gamma(12.5)\sqrt{2/24}} = 1.010$

The bias is quite small for n as large as 25.

7-51. $P\left(|\overline{X} - \mu| \ge \frac{c\sigma}{\sqrt{n}}\right) \le \frac{1}{c^2}$ from Chebyshev's inequality.

Then, $P\left(|\overline{X} - \mu| < \frac{c\sigma}{\sqrt{n}}\right) \ge 1 - \frac{1}{c^2}$. Given an ε, n and c can be chosen sufficiently large that the last probability is near 1 and $\frac{c\sigma}{\sqrt{n}}$ is equal to ε.

7-52. $P\left(X_{(n)} \le t\right) = P\left(X_i \le t \text{ for } i = 1,\ldots,n\right) = [F(t)]^n$

$P\left(X_{(1)} > t\right) = P\left(X_i > t \text{ for } i = 1,\ldots,n\right) = [1 - F(t)]^n$

Then, $P\left(X_{(1)} \le t\right) = 1 - [1 - F(t)]^n$

$f_{X_{(1)}}(t) = \frac{\partial}{\partial t} F_{X_{(1)}}(t) = n[1 - F(t)]^{n-1} f(t)$

$f_{X_{(n)}}(t) = \frac{\partial}{\partial t} F_{X_{(n)}}(t) = n[F(t)]^{n-1} f(t)$

7-53. $P\left(X_{(1)} = 0\right) = F_{X_{(1)}}(0) = 1 - [1 - F(0)]^n = 1 - p^n$ because F(0) = 1 - p.

$P\left(X_{(n)} = 1\right) = 1 - F_{X_{(n)}}(0) = 1 - [F(0)]^n = 1 - (1-p)^n$

7-54. $P(X \le t) = F(t) = \Phi\left[\frac{t-\mu}{\sigma}\right]$. From Exercise 7-51,

$$f_{X_{(1)}}(t) = n\left\{1 - \Phi\left[\frac{t-\mu}{\sigma}\right]\right\}^{n-1} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(t-\mu)^2}{2\sigma^2}}$$

$$f_{X_{(n)}}(t) = n\left\{\Phi\left[\frac{t-\mu}{\sigma}\right]\right\}^{n-1} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(t-\mu)^2}{2\sigma^2}}$$

7-55. $P(X \le t) = 1 - e^{-\lambda t}$. From Exercise 7-51,

$F_{X_{(1)}}(t) = 1 - e^{-n\lambda t}$	$f_{X_{(1)}}(t) = ne^{-(n-1)\lambda t}\lambda e^{-\lambda t}$ $= n\lambda e^{-n\lambda t}$
$F_{X_{(n)}}(t) = [1 - e^{-\lambda t}]^n$	$f_{X_{(n)}}(t) = n[1 - e^{-\lambda t}]^{n-1}\lambda e^{-\lambda t}$

7-56. $P\left(F\left(X_{(n)}\right) \le t\right) = P\left(X_{(n)} \le F^{-1}(t)\right) = t^n$ from Exercise 7-51 for $0 \le t \le 1$.

If $Y = F(X_{(n)})$, then $f_Y(y) = ny^{n-1}, 0 \le y \le 1$. Then, $E(Y) = \int_0^1 ny^n dy = \frac{n}{n+1}$

$P\left(F\left(X_{(1)}\right) > t\right) = P\left(X_{(1)} > F^{-1}(t)\right) = 1 - (1-t)^n$ from Exercise 7-51 for $0 \le t \le 1$.

If $Y = F(X_{(1)})$, then $f_Y(y) = n(1-t)^{n-1}, 0 \le y \le 1$.

Then, $E(Y) = \int_0^1 yn(1-y)^{n-1} dy = \frac{1}{n+1}$ where integration by parts is used. Therefore,

$E[F(X_{(n)})] = \frac{n}{n+1}$ and $E[F(X_{(1)})] = \frac{1}{n+1}$

7-57. $E(V) = k\sum_{i=1}^{n-1}[E(X_{i+1}^2) + E(X_i^2) - 2E(X_iX_{i+1})]$

$= k\sum_{i=1}^{n-1}(\sigma^2 + \mu^2 + \sigma^2 + \mu^2 - 2\mu^2)$

$= k(n-1)2\sigma^2$

Therefore, $k = \frac{1}{2(n-1)}$

CHAPTER 8

Section 8-1

8-1. a) α = P(reject H_0 when H_0 is true)

$= P(\overline{X} \le 11.5 \text{ when } \mu = 12) = P\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \le \frac{11.5-12}{0.5/\sqrt{4}}\right) = P(Z \le -2) = 1 - P(Z \le 2)$

$= 1 - 0.97725$
$= 0.02275.$

The probability of rejecting the null hypothesis when it is true is 0.02275.

b) $\beta = P(\text{accept } H_0 \text{ when } \mu = 11.25) = P(\overline{X} > 11.5 \text{ when } \mu = 11.25) = P\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} > \frac{11.5-11.25}{0.5/\sqrt{4}}\right)$

$P(Z > 1.0) = 1 - P(Z \le 1.0) = 1 - 0.84134 = 0.15866$

The probability of accepting the null hypothesis when it is false is 0.15866.

8-2. a) $\alpha = P(\overline{X} \le 11.5 \text{ when } \mu = 12) = P\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \le \frac{11.5-12}{0.5/\sqrt{16}}\right) = P(Z \le -4) = 1 - P(Z \le 4)$

$= 1 - 1 = 0.$

The probability of rejecting the null, when the null is true, is 0 with a sample size of 16.

b) $\beta = P(\overline{X} > 11.5 \text{ when } \mu = 11.25) = P\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} > \frac{11.5-11.25}{0.5/\sqrt{16}}\right) = P(Z > 2) = 1 - P(Z \le 2)$

$= 1 - 0.97725 = 0.02275.$

The probability of accepting the null hypothesis when it is false is 0.02275.

8-3. Find the boundary of the critical region if $\alpha = 0.01$:

$0.01 = P\left(Z \le \frac{c-12}{0.5/\sqrt{4}}\right)$

What Z value will give a probability of 0.01? Using Table 2 in the appendix, Z value is -2.33.

Thus, $\frac{c-12}{0.5/\sqrt{4}} = -2.33$, $c = 11.4175$

8-4. $0.05 = P\left(Z \le \frac{c-12}{0.5/\sqrt{4}}\right)$

What Z value will give a probability of 0.05? Using Table 2 in the appendix, Z value is -1.65.

Thus, $\frac{c-12}{0.5/\sqrt{4}} = -1.65$, $c = 11.5875$

8-5. a) $\alpha = P(\overline{X} \le 98.5) + P(\overline{X} > 101.5)$

$$= P\left(\frac{\overline{X}-100}{2/\sqrt{9}} \le \frac{98.5-100}{2/\sqrt{9}}\right) + P\left(\frac{\overline{X}-100}{2/\sqrt{9}} > \frac{101.5-100}{2/\sqrt{9}}\right)$$

$= P(Z \le -2.25) + P(Z > 2.25)$
$= (1 - P(Z \le 2.25)) + (1 - P(Z \le 2.25))$
$= 1 - 0.98778 + 1 - 0.98778$
$= 0.01222 + 0.01222 = 0.02444$

b) $\beta = P(98.5 \le \overline{X} \le 101.5 \text{ when } \mu = 103)$

$$= P\left(\frac{98.5-103}{2/\sqrt{9}} \le \frac{\overline{X}-103}{2/\sqrt{9}} \le \frac{101.5-103}{2/\sqrt{9}}\right)$$

$= P(-6.75 \le Z \le -2.25)$
$= P(Z \le -2.25) - P(Z \le -6.75)$
$= 1 - P(Z \le 2.25) - (1 - P(Z \le 6.75))$
$= 0.01222 - 0 = 0.01222$

c) $\beta = P(98.5 \le \overline{X} \le 101.5 \text{ when } \mu = 105)$

$$= P\left(\frac{98.5-105}{2/\sqrt{9}} \le \frac{\overline{X}-105}{2/\sqrt{9}} \le \frac{101.5-105}{2/\sqrt{9}}\right)$$

$= P(-9.75 \le Z \le -5.25)$
$= P(Z \le -5.25) - P(Z \le -9.75)$
$= 1 - P(Z \le 5.25) - (1 - P(Z \le 9.75))$
$= 0 - 0$
$= 0.$

The probability of accepting the null hypothesis when it is actually false is smaller in part c since the true mean, $\mu = 105$, is further from the acceptance region. A larger difference exists.

8-6. Use n = 5, everything else held constant:

a) $P(\overline{X} \le 98.5) + P(\overline{X} > 101.5)$

$$= P\left(\frac{\overline{X}-100}{2/\sqrt{5}} \le \frac{98.5-100}{2/\sqrt{5}}\right) + P\left(\frac{\overline{X}-100}{2/\sqrt{5}} > \frac{101.5-100}{2/\sqrt{5}}\right)$$

$= P(Z \le -1.68) + P(Z > 1.68)$
$= (1 - P(Z \le 1.68)) + (1 - P(Z \le 1.68))$
$= (1 - 0.95352) + (1 - 0.95352)$
$= 0.09296$

b) $\beta = P(98.5 \le \overline{X} \le 101.5 \text{ when } \mu = 103)$

$$= P\left(\frac{98.5-103}{2/\sqrt{5}} \le \frac{\overline{X}-103}{2/\sqrt{5}} \le \frac{101.5-103}{2/\sqrt{5}}\right)$$

$= P(-5.03 \le Z \le -1.68)$
$= P(Z \le -1.68) - P(Z \le -5.03)$
$= 1 - P(Z \le 1.68) - (1 - P(Z \le 5.03))$
$= 0.04648 - 0$
$= 0.04648$

c) $\beta = P(98.5 \le \overline{x} \le 101.5 \text{ when } \mu = 105)$

$$= P\left(\frac{98.5-105}{2/\sqrt{5}} \le \frac{\overline{X}-105}{2/\sqrt{5}} \le \frac{101.5-105}{2/\sqrt{5}}\right)$$

$= P(-7.27 \le Z \le -3.91)$
$= P(Z \le -3.91) - P(Z \le -7.27)$
$= 1 - P(Z \le 3.91) - (1 - P(Z \le 7.27))$
$= (1 - 0.99995) - 0$
$= 0.00005 - 0$
$= 0.00005$

8-7. a) $\alpha = P(\overline{X} > 185 \text{ when } \mu = 175)$

$$= P\left(\frac{\overline{X} - 175}{20/\sqrt{10}} > \frac{185 - 175}{20/\sqrt{10}}\right)$$

$= P(Z > 1.58)$
$= 1 - P(Z \le 1.58)$
$= 1 - 0.94295$
$= 0.05705$

b) $\beta = P(\overline{X} \le 185 \text{ when } \mu = 195)$

$$= P\left(\frac{\overline{X} - 195}{20/\sqrt{10}} \le \frac{185 - 195}{20/\sqrt{10}}\right)$$

$= P(Z \le -1.58)$
$= 1 - P(Z \le 1.58)$
$= 0.05705.$

8-8. a) Reject the null hypothesis and conclude that the mean foam height is greater than 175 mm.

b) $P(\overline{X} > 190 \text{ when } \mu = 175)$

$$= P\left(\frac{\overline{X} - 175}{20/\sqrt{10}} > \frac{190 - 175}{20/\sqrt{10}}\right)$$

$= P(Z > 2.37) = 1 - P(Z \le 2.37)$
$= 1 - 0.99111$
$= 0.00889.$

The probability that a value of at least 190 mm would be observed (if the true mean height is 175 mm) is only 0.00889. Thus, the sample value of $\overline{x} = 190$ mm would be an unusual result.

8-9. Using n = 16:

a) $\alpha = P(\overline{X} > 185 \text{ when } \mu = 175)$

$$= P\left(\frac{\overline{X} - 175}{20/\sqrt{16}} > \frac{185 - 175}{20/\sqrt{16}}\right)$$

$= P(Z > 2)$
$= 1 - P(Z \le 2)$
$= 1 - 0.97725$
$= 0.02275$

b) $\beta = P(\overline{X} \le 185 \text{ when } \mu = 195)$

$$= P\left(\frac{\overline{X} - 195}{20/\sqrt{16}} \le \frac{185 - 195}{20/\sqrt{16}}\right)$$

$= P(Z \le -2)$
$= 1 - P(Z \le 2)$
$= 1 - 0.97725$
$= 0.02275.$

8-10. a) $0.0571 = P\left(\overline{X} > c \text{ when } \mu = 175\right) = P\left(Z > \frac{c - 175}{20/\sqrt{16}}\right) = P(Z \ge 1.58)$

Thus, $1.58 = \frac{c - 175}{20/\sqrt{16}}$, and $c = 182.9$

b) If the true mean foam height is 195 mm, then

$\beta = P(\overline{X} \le 182.9 \text{ when } \mu = 195)$

$$= P\left(Z \le \frac{182.9 - 195}{20/\sqrt{16}}\right)$$

$= P(Z \le -2.42)$
$= 1 - P(Z \le 2.42)$
$= 0.00776.$

c) For the same level of α, with the increased sample size, β is reduced.

8-11. a) $\alpha = P(\overline{X} \le 4.85 \text{ when } \mu = 5) + P(\overline{X} > 5.15 \text{ when } \mu = 5)$

$$= P\left(\frac{\overline{X}-5}{0.25/\sqrt{8}} \le \frac{4.85-5}{0.25/\sqrt{8}}\right) + P\left(\frac{\overline{X}-5}{0.25/\sqrt{8}} > \frac{5.15-5}{0.25/\sqrt{8}}\right)$$

$= P(Z \le -1.7) + P(Z > 1.7)$
$= (1 - P(Z \le 1.7)) + (1 - P(Z \le 1.7)$
$= (1 - 0.95543) + (1 - 0.95543)$
$= 0.08914.$

b) Power $= 1 - \beta$

$\beta = P(4.85 \le \overline{X} \le 5.15 \text{ when } \mu = 5.1)$

$$= P\left(\frac{4.85-5.1}{0.25/\sqrt{8}} \le \frac{\overline{X}-5.1}{0.25/\sqrt{8}} \le \frac{5.15-5.1}{0.25/\sqrt{8}}\right)$$

$= P(-2.83 \le Z \le 0.565)$
$= P(Z \le 0.565) - P(Z \le -2.83)$
$= P(Z \le 0.565) - (1 - P(Z \le 2.83))$
$= 0.71566 - (1 - 0.99767)$
$= 0.71566 - 0.00233$
$= 0.71333$
$1 - \beta = 0.2867.$

8-12. Using n = 16:

a) $\alpha = P(\overline{X} \le 4.85 \text{ when } \mu = 5) + P(\overline{X} > 5.15 \text{ when } \mu = 5)$

$$= P\left(\frac{\overline{X}-5}{0.25/\sqrt{16}} \le \frac{4.85-5}{0.25/\sqrt{16}}\right) + P\left(\frac{\overline{X}-5}{0.25/\sqrt{16}} > \frac{5.15-5}{0.25/\sqrt{16}}\right)$$

$= P(Z \le -2.4) + P(Z > 2.4)$
$= (1 - P(Z \le 2.4)) + (1 - P(Z \le 2.4))$
$= 2(1 - P(Z \le 2.4))$
$= 2(1 - 0.99180)$
$= 2(0.0082)$
$= 0.0164.$

b) $\beta = P(4.85 \le \overline{X} \le 5.15 \text{ when } \mu = 5.1)$

$$= P\left(\frac{4.85-5.1}{0.25/\sqrt{16}} \le \frac{\overline{X}-5.1}{0.25/\sqrt{16}} \le \frac{5.15-5.1}{0.25/\sqrt{16}}\right)$$

$= P(-4 \le Z \le 0.8) = P(Z \le 0.8) - P(Z \le -4)$
$= P(Z \le 0.8) - (1 - P(Z \le 4)$
$= 0.78814 - (1 - 0)$
$= 0.78814 - 0$
$= 0.78814$
$1 - \beta = 0.21186$

8-13. $\alpha = P\left(\overline{X} \le c_L\right) + P\left(\overline{X} > c_H\right)$

$0.025 = P\left(\overline{X} \le c_L\right),\ 0.025 = P\left(\overline{X} > c_H\right)$

$P\left(\frac{c_L - 5}{0.25/\sqrt{8}}\right) = 0.025,$ thus, $-1.96 = \frac{c_L - 5}{0.25/\sqrt{8}}$; $c_L = 4.83$

$P\left(\frac{c_H - 5}{0.25/\sqrt{8}}\right) = 0.025,$ thus, $1.96 = \frac{c_H - 5}{0.25/\sqrt{8}}$; $c_H = 5.17$

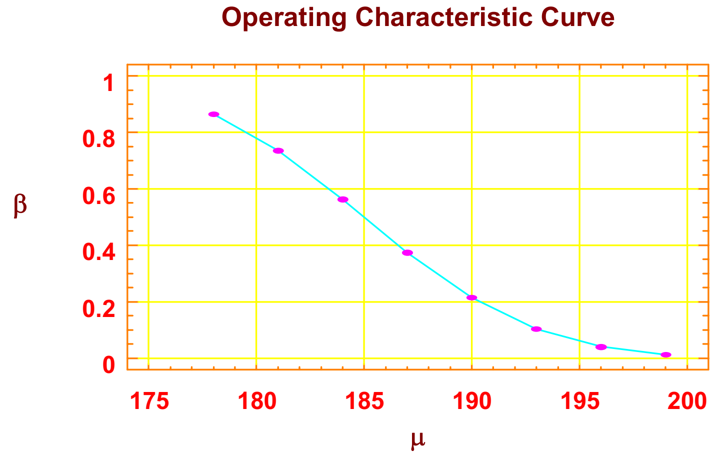
$P\left(4.83 \le \overline{X} \le 5.17\right)$

8-14. Operating characteristic curve:

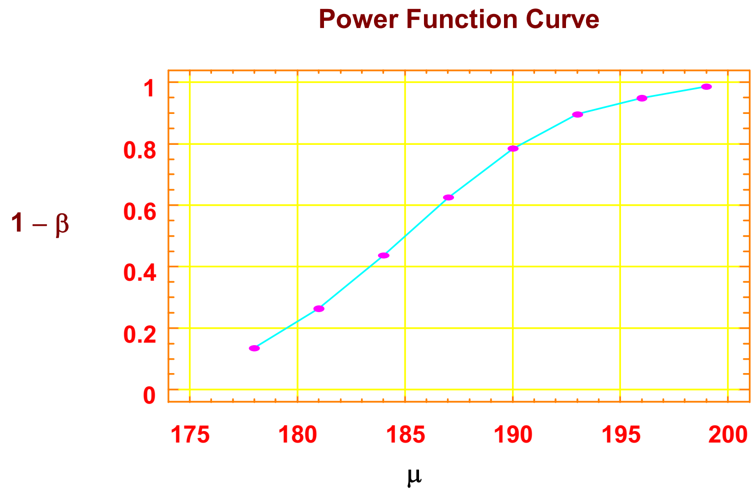
$\bar{x} = 185$

$$\beta = P\left(Z \le \frac{\bar{x} - \mu}{20/\sqrt{10}}\right) = P\left(Z \le \frac{185 - \mu}{20/\sqrt{10}}\right)$$

μ	$P\left(Z \le \frac{185-\mu}{20/\sqrt{10}}\right) =$	β	$1 - \beta$
178	$P(Z \le 1.11) =$	0.8643	0.1357
181	$P(Z \le 0.63) =$	0.7357	0.2643
184	$P(Z \le 0.16) =$	0.5636	0.4364
187	$P(Z \le -0.32) =$	0.3745	0.6255
190	$P(Z \le -0.79) =$	0.2148	0.7852
193	$P(Z \le -1.26) =$	0.1038	0.8962
196	$P(Z \le -1.74) =$	0.0409	0.9491
199	$P(Z \le -2.21) =$	0.0136	0.9864



8-15.



8-16. The problem statement implies H_0: p = 0.6, H_1: p > 0.6 and defines an acceptance region as $\hat{p} \le \frac{315}{500} = 0.63$ and rejection region as $\hat{p} > 0.63$

a) $\alpha = P(\hat{P} \ge 0.63 \text{ when } p = 0.6) = P\left(Z \ge \frac{0.63 - 0.6}{\sqrt{\frac{0.6(0.4)}{500}}}\right) = P(Z \ge 1.37) = 1 - P(Z < 1.37) = 0.08535$.

b) $\beta = P(\hat{P} \le 0.63 \text{ when } p = 0.75) = P(Z \le -6.196) = 0.$

8-17. X ~ bin(10, 0.3) Implicitly, H_0: p = 0.3 and H_1: p < 0.3
n = 20
Accept region: $\hat{p} > 0.1$
Reject region: $\hat{p} \le 0.1$
Use the normal approximation for parts a), b) and c):

a) When p =0.3 $\alpha = P(\hat{p} < 0.1) = P\left(Z \le \frac{0.1 - 0.3}{\sqrt{\frac{0.3(0.7)}{20}}}\right)$

$= P(Z \le -1.95)$
$= 0.02559$

b) When p = 0.2 $\beta = P(\hat{p} > 0.1) = P\left(Z > \frac{0.1 - 0.2}{\sqrt{\frac{0.2(0.8)}{20}}}\right)$

$= P(Z > -1.12)$
$= 1 - P(Z < -1.12)$
$= 0.86864$

c) Power = 1 – β = 1 – 0.86864 = 0.13136.

8-18. X ~ bin(15, 0.4) H_0: p = 0.4 and H_1: p ≠ 0.4

4/15 = 0.267 8/15 = 0.533

Accept Region: $0.267 \le \hat{p} \le 0.533$
Reject Region: $\hat{p} \le 0.267$ or $\hat{p} > 0.533$

Use the normal approximation for parts a) and b)

a) When p = 0.4, $\alpha = P(\hat{p} \le 0.267) + P(\hat{p} > 0.533)$

$$= P\left(Z \le \frac{0.267 - 0.4}{\sqrt{\frac{0.4(0.6)}{15}}}\right) + P\left(Z > \frac{0.533 - 0.4}{\sqrt{\frac{0.4(0.6)}{15}}}\right)$$
$= P(Z \le -1.05) + P(Z > 1.05)$
$= P(Z \le -1.05) + (1 - P(Z < 1.05))$
$= 0.14686 + 0.14686$
$= 0.29372$

b) When p = 0.2, $\beta = P(0.267 < \hat{p} \le 0.533) = P\left(\frac{0.267-0.2}{\sqrt{\frac{0.2(0.8)}{15}}} < Z \le \frac{0.533-0.2}{\sqrt{\frac{0.2(0.8)}{15}}}\right)$

$$= P(0.65 < Z \le 3.22)$$
$$= P(Z \le 3.22) - P(Z < 0.65)$$
$$= 0.99936 - 0.74215$$
$$= 0.25721$$

Section 8-2

8-19. a) 1) The parameter of interest is the true mean breaking strength, μ.

2) $H_0 : \mu = 100$

3) $H_1 : \mu > 100$

4) $\alpha = 0.05$

5) $z_0 = \frac{\overline{x} - \mu}{\sigma / \sqrt{n}}$

6) Reject H_0 if $z_0 > z_\alpha$ where $z_{0.05} = 1.65$

7) $\overline{x} = 98$, $\sigma = 2$

$$z_0 = \frac{98 - 100}{2/\sqrt{9}} = -3$$

8) Since $-3 < 1.65$ do not reject H_0 and conclude that the fiber would not be judged acceptable at α=0.05.

b) P-value = $P(Z \ge -3) = 1 - \Phi(-3) = 0.99865$

c) For $\alpha = 0.05$, accept H_0 if $\overline{X} < 100 + 1.65\left(\frac{2}{\sqrt{9}}\right) = 101.1$

$$P\left(\overline{X} \le 101.1 \text{ when } \mu = 104\right) = P\left(Z \le \frac{101.1 - 104}{2/\sqrt{9}}\right) = P\left(Z \le -4.35\right) = \Phi(-4.35) \cong 0$$

The probability is 0 of accepting the null hypothesis if the true mean breaking strength is 104 psi, with a level of significance of $\alpha = 0.05$.

d) $z_{\alpha/2} = z_{0.025} = 1.96$

$$\overline{x} - z_{0.025}\left(\frac{\sigma}{\sqrt{n}}\right) \le \mu \le \overline{x} + z_{0.025}\left(\frac{\sigma}{\sqrt{n}}\right)$$

$$98 - 1.96\left(\frac{2}{\sqrt{9}}\right) \le \mu \le 98 + 1.96\left(\frac{2}{\sqrt{9}}\right)$$

$96.7 \le \mu \le 99.31$

With 95% confidence, we believe the true mean breaking strength is between 96.7 psi and 99.3 psi.

8-20. a) 1) The parameter of interest is the true mean yield, μ.

2) $H_0 : \mu = 90$

3) $H_1 : \mu \ne 90$

4) $\alpha = 0.05$

5) $z_0 = \frac{\overline{x} - \mu}{\sigma / \sqrt{n}}$

6) Reject H_0 if $z_0 < -z_{\alpha/2}$ where $-z_{0.025} = -1.96$ or $z_0 > z_{\alpha/2}$ where $z_{0.025} = 1.96$

7) $\overline{x} = 90.48$, $\sigma = 3$

$$z_0 = \frac{90.48 - 90}{3/\sqrt{5}} = 0.36$$

8) Since $-1.96 < 0.36 < 1.96$ do not reject H_0 and conclude the yield is not significantly different from 90% at $\alpha = 0.05$.

b) P-value = $2[1 - \Phi(0.36)] = 2[1 - 0.64058] = 0.71884$

c) $n = \dfrac{(z_{\alpha/2} + z_\beta)^2 \sigma^2}{\delta^2} = \dfrac{(z_{0.025} + z_{0.05})^2 3^2}{(85-90)^2} = \dfrac{(1.96+1.65)^2 9}{(-5)^2} = 4.67$

$n \cong 5.$

d) $\beta = \Phi\left(z_{0.025} + \dfrac{90-92}{3/\sqrt{5}}\right) - \Phi\left(-z_{0.025} + \dfrac{90-92}{3/\sqrt{5}}\right)$

$= \Phi(1.96 + -1.491) - \Phi(-1.96 + -1.491)$
$= \Phi(0.47) - \Phi(-3.45)$
$= \Phi(0.47) - (1 - \Phi(3.45))$
$= 0.68082 - (1 - 0.99972)$
$= 0.68054.$

e) For $\alpha = 0.05$, $z_{\alpha/2} = z_{0.025} = 1.96$

$$\bar{x} - z_{0.025}\left(\frac{\sigma}{\sqrt{n}}\right) \le \mu \le \bar{x} + z_{0.025}\left(\frac{\sigma}{\sqrt{n}}\right)$$

$$90.48 - 1.96\left(\frac{3}{\sqrt{5}}\right) \le \mu \le 90.48 + 1.96\left(\frac{3}{\sqrt{5}}\right)$$

$87.85 \le \mu \le 93.11$

With 95% confidence, we believe the true mean yield of the chemical process is between 87.85% and 93.11%.

8-21. a) 1) The parameter of interest is the true mean hole diameter, μ.

2) $H_0 : \mu = 1.50$
3) $H_1 : \mu \neq 1.50$
4) $\alpha = 0.01$
5) $z_0 = \dfrac{\bar{x} - \mu}{\sigma/\sqrt{n}}$
6) Reject H_0 if $z_0 < -z_{\alpha/2}$ where $-z_{0.005} = -2.58$ or $z_0 > z_{\alpha/2}$ where $z_{0.005} = 2.58$
7) $\bar{x} = 1.5045$, $\sigma = 0.01$

$$z_0 = \frac{1.5045 - 1.50}{0.01/\sqrt{10}} = 1.423$$

8) Since $-2.58 < 1.423 < 2.58$, do not reject the null hypothesis and conclude the true mean hole diameter is not significantly different from 1.5 in. at $\alpha = 0.01$.

b) P-value $= 2[1 - \Phi(1.423)] \cong 2[1 - 0.92219] = 0.15562$

c) Set $\beta = 1 - 0.90 = 0.10$

$n = \dfrac{(z_{\alpha/2} + z_\beta)^2 \sigma^2}{\delta^2} = \dfrac{(z_{0.005} + z_{0.10})^2 \sigma^2}{(1.505 - 1.50)^2} \cong \dfrac{(2.58 + 1.29)^2 (0.01)^2}{(0.005)^2} = 59.91,$

$n \cong 60.$

d) $\beta = \Phi\left(z_{0.005} + \dfrac{1.50 - 1.505}{0.01/\sqrt{10}}\right) - \Phi\left(-z_{0.005} + \dfrac{1.50 - 1.505}{0.01/\sqrt{10}}\right)$

$\cong \Phi(2.58 + -1.581) - \Phi(-2.58 + -1.581)$
$= \Phi(0.999) - \Phi(-4.161)$
$= 0.84134 - 0$

$\beta \cong 0.84134.$

e) For $\alpha = 0.01$, $z_{\alpha/2} = z_{0.005} = 2.58$

$$\bar{x} - z_{0.005}\left(\frac{\sigma}{\sqrt{n}}\right) \le \mu \le \bar{x} + z_{0.005}\left(\frac{\sigma}{\sqrt{n}}\right)$$

$$1.5045 - 2.58\left(\frac{0.01}{\sqrt{10}}\right) \le \mu \le 1.5045 + 2.58\left(\frac{0.01}{\sqrt{10}}\right)$$

$1.496 \le \mu \le 1.513$

The confidence interval constructed contains the value 1.5, thus the true mean hole diameter could possibly be 1.5 in. using a 99% level of confidence. Since a two-sided 99% confidence interval is equivalent to a two-sided hypothesis test at $\alpha = 0.01$, the conclusions necessarily must be consistent.

8-22. a) 1) The parameter of interest is the true mean piston ring diameter, μ.

2) $H_0 : \mu = 74.035$

3) $H_1 : \mu \neq 74.035$

4) $\alpha = 0.01$

5) $z_0 = \dfrac{\overline{x} - \mu}{\sigma / \sqrt{n}}$

6) Reject H_0 if $z_0 < -z_{\alpha/2}$ where $-z_{0.005} = -2.58$ or $z_0 > z_{\alpha/2}$ where $z_{0.005} = 2.58$

7) $\overline{x} = 74.036$, $\sigma = 0.001$

$$z_0 = \frac{74.036 - 74.035}{0.001/\sqrt{15}} = 3.87$$

8) Since 3.87 > 2.58, reject the null hypothesis and conclude the true mean piston ring diameter is not 74.035 mm.

b) P-value = $2[1 - \Phi(3.87)] = 2[1 - 0.99995] = 0.00010$

Again since the P-value is less than the level of significance, α, we would reject the null hypothesis.

c) For $\alpha = 0.01$, $z_{\alpha/2} = z_{0.005} = 2.58$

$$\overline{x} - z_{0.005}\left(\frac{\sigma}{\sqrt{n}}\right) \le \mu \le \overline{x} + z_{0.005}\left(\frac{\sigma}{\sqrt{n}}\right)$$

$$74.036 - 2.58\left(\frac{0.001}{\sqrt{15}}\right) \le \mu \le 74.036 + 2.58\left(\frac{0.001}{\sqrt{15}}\right)$$

$$74.0353 \le \mu \le 74.0367$$

With 99% confidence, we believe the true mean piston ring diameter is between 74.0353 and 74.0367 mm.

d) For $\alpha = 0.05$, $z_\alpha = z_{0.05} = 1.65$

$$\overline{x} - z_{0.05}\frac{\sigma}{\sqrt{n}} \le \mu$$

$$74.036 - 1.65\frac{0.001}{\sqrt{15}} \le \mu$$

$$74.0356 \le \mu$$

With 95% confidence, the true mean piston ring diameter is at least 74.0356 mm.

8-23. a) 1) The parameter of interest is the true mean bulb life, μ.

2) $H_0 : \mu = 1000$

3) $H_1 : \mu > 1000$

4) $\alpha = 0.05$

5) $z_0 = \dfrac{\overline{x} - \mu}{\sigma / \sqrt{n}}$

6) Reject H_0 if $z_0 > z_\alpha$ where $z_{0.05} = 1.65$

7) $\overline{x} = 1014$, $\sigma = 25$

$$z_0 = \frac{1014 - 1000}{25 / \sqrt{20}} = 2.504$$

8) Since 2.504 > 1.65, reject the null hypothesis and conclude there is sufficient evidence to support the claim the bulb life exceeds 1000 hrs at $\alpha = 0.05$.

b) P-value = $P(Z > 2.504) = 1 - P(Z \le 2.504) = 1 - \Phi(2.504) = 0.99461 = 0.00539$.

c) $\beta = \Phi\left(z_{0.05} - \dfrac{(1000 - 1050)\sqrt{20}}{25} \right) = \Phi(1.65 + -8.9445) = \Phi(-7.30) = 0$

d) $n = \dfrac{\left(z_\alpha + z_\beta\right)^2 \sigma^2}{\delta^2} = \dfrac{\left(z_{0.05} + z_{0.10}\right)^2 \sigma^2}{(1025 - 1000)^2} = \dfrac{(1.65 + 1.29)^2 (25)^2}{(25)^2} = 8.6436,$

$n \cong 9$.

e) For $\alpha = 0.05$, $z_{\alpha/2} = z_{0.025} = 1.96$

$$\overline{x} - z_{0.025}\left(\frac{\sigma}{\sqrt{n}}\right) \le \mu \le \overline{x} + z_{0.025}\left(\frac{\sigma}{\sqrt{n}}\right)$$

$$1014 - 1.96\left(\frac{25}{\sqrt{20}}\right) \le \mu \le 1014 + 1.96\left(\frac{25}{\sqrt{20}}\right)$$

$$1003 \le \mu \le 1025$$

With 95% confidence, we believe the true mean bulb life is between 1003 hrs and 1025 hrs.

f) $\overline{x} - z_{0.050}\dfrac{\sigma}{\sqrt{n}} \le \mu$

$$1014 - 1.645\frac{25}{\sqrt{20}} \le \mu$$

$$1005 \le \mu$$

With 95% confidence, the true mean bulb life is at least 1005 hrs.

8-24. a) 1) The parameter of interest is the true mean compressive strength, μ.

2) $H_0 : \mu = 3500$

3) $H_1 : \mu \ne 3500$

4) $\alpha = 0.01$

5) $z_0 = \dfrac{\overline{x} - \mu}{\sigma / \sqrt{n}}$

6) Reject H_0 if $z_0 < -z_{\alpha/2}$ where $-z_{0.005} = -2.58$ or $z_0 > z_{\alpha/2}$ where $z_{0.005} = 2.58$

7) $\overline{x} = 3250$, $\sigma = 31.62$

$$z_0 = \frac{3250 - 3500}{31.62 / \sqrt{12}} = -27.39$$

8) Since $-27.39 < -2.58$, reject the null hypothesis and conclude the true mean compressive strength is significantly different from 3500 at $\alpha = 0.01$.

b) Smallest level of significance = P-value = $2[1 - \Phi(27.39)] = 2[1 - 1] = 0$

The smallest level of significance at which we are willing to reject the null hypothesis is 0.

c) $z_{\alpha/2} = z_{0.025} = 1.96$

$$\bar{x} - z_{0.025}\left(\frac{\sigma}{\sqrt{n}}\right) \le \mu \le \bar{x} + z_{0.025}\left(\frac{\sigma}{\sqrt{n}}\right)$$

$$3250 - 1.96\left(\frac{31.62}{\sqrt{12}}\right) \le \mu \le 3250 + 1.96\left(\frac{31.62}{\sqrt{12}}\right)$$

$$3232.11 \le \mu \le 3267.89$$

With 95% confidence, we believe the true mean compressive strength is between 3232.11 psi and 3267.89 psi.

d) $z_{\alpha/2} = z_{0.005} = 2.58$

$$\bar{x} - z_{0.005}\left(\frac{\sigma}{\sqrt{n}}\right) \le \mu \le \bar{x} + z_{0.005}\left(\frac{\sigma}{\sqrt{n}}\right)$$

$$3250 - 2.58\left(\frac{31.62}{\sqrt{12}}\right) \le \mu \le 3250 + 2.58\left(\frac{31.62}{\sqrt{12}}\right)$$

$$3226.5 \le \mu \le 3273.5$$

With 99% confidence, we believe the true mean compressive strength is between 3226.5 psi and 3273.5 psi.

The 99% confidence interval is wider than the 95% confidence interval. The confidence interval with the larger level of confidence will always result in a wider confidence interval when $\bar{x}$, σ^2, and n are held constant.

8-25. $z_{\alpha/2} = z_{0.025} = 1.96$, $E = 5$

$$n \cong \left[\frac{25(1.96)}{5}\right]^2 = 96.04,\ n \cong 97.$$

8-26. $z_{\alpha/2} = z_{0.025} = 1.96$, $E = 3$ since the desired width is 6, set $E = \frac{6}{2} = 3$

$$n \cong \left[\frac{25(1.96)}{3}\right]^2 = 266.7,\ n \cong 267.$$

8-27. $z_{\alpha/2} = z_{0.005} = 2.58$, $E = 15$

$$n = \left[\frac{31.62(2.58)}{15}\right]^2 = 29.58,\ n \cong 30.$$

Section 8-3

8-28. a) 1) The parameter of interest is the true mean life, μ.

2) $H_0 : \mu = 60000$

3) $H_1 : \mu > 60000$

4) $\alpha = 0.05$

5) $t_0 = \dfrac{\bar{x} - \mu}{s / \sqrt{n}}$

6) Reject H_0 if $t_0 > t_{\alpha,n-1}$ where $t_{0.05,15} = 1.753$

7) $\bar{x} = 60139.7$ $s = 3645.94$ $n = 16$

$$t_0 = \frac{60139.7 - 60000}{3645.94 / \sqrt{16}} = 0.153$$

8) Since 0.153 < 1.753, do not reject the null hypothesis and conclude there is insufficient evidence to indicate that the mean life of this new tire is in excess of 60,000 km at $\alpha = 0.05$.

b) $d = \frac{\delta}{\sigma} = \frac{\mu - \mu_0}{\sigma} = \frac{61000 - 60000}{3645.94} = 0.274$

Using the OC curve, Chart VIc) for $\alpha = 0.05$, d = 0.274, and n = 16, we get $\beta \cong 0.72$ and power of 1−0.72 = 0.28. With the power being smaller than the acceptable level, 16 is not an adequate sample size for detecting a difference with probability of at least 0.90.

c) $t_{\alpha,n-1} = t_{0.025,15} = 2.131$

$$\bar{x} - t_{0.025,15}\left(\frac{s}{\sqrt{n}}\right) \le \mu \le \bar{x} + t_{0.025,15}\left(\frac{s}{\sqrt{n}}\right)$$

$$60139.7 - 2.131\left(\frac{3645.94}{\sqrt{16}}\right) \le \mu \le 60139.7 + 2.131\left(\frac{3645.94}{\sqrt{16}}\right)$$

$$58,197.3 \le \mu \le 62,082.07$$

With 95% confidence, we believe the true mean tire life is between 58,197.3 km and 62,082.07 km.

8-29. In order to use t statistics in hypothesis testing, we need to assume that the underlying distribution is normal.

1) The parameter of interest is the true Izod impact strength, μ.
2) $H_0 : \mu = 1.0$
3) $H_1 : \mu > 1.0$
4) $\alpha = 0.01$
5) $t_0 = \frac{\bar{x} - \mu}{s/\sqrt{n}}$
6) Reject H_0 if $t_0 > t_{\alpha,n-1}$ where $t_{0.01,19} = 2.539$
7) $\bar{x} = 1.25$ s = 0.25 n = 20

$$t_0 = \frac{1.25 - 1.0}{0.25/\sqrt{20}} = 4.47$$

8) Since 4.47 > 2.539, reject the null hypothesis and conclude there is sufficient evidence to indicate that the true Izod impact strength is greater than 1.0 ft-lb/in at $\alpha = 0.01$.

8-30. In order to use t statistics in hypothesis testing, we need to assume that the underlying distribution is normal.

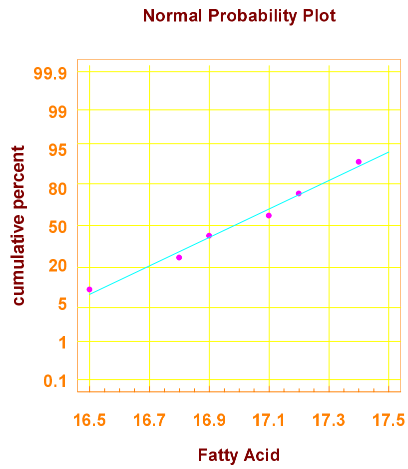
a) 1) The parameter of interest is the true mean amount of current, μ.
2) $H_0 : \mu = 300$
3) $H_1 : \mu \neq 300$
4) $\alpha = 0.05$
5) $t_0 = \frac{\bar{x} - \mu}{s/\sqrt{n}}$
6) Reject H_0 if $t_0 < -t_{\alpha/2,n-1}$ where $-t_{0.025,9} = -2.262$ or $t > t_{\alpha/2,n-1}$ where $t_{0.025,9} = 2.262$
7) $\bar{x} = 317.2$ s = 15.75 n = 10

$$t_0 = \frac{317.2 - 300}{15.75/\sqrt{10}} = 3.46$$

8) Since 3.46 > 2.262, reject the null hypothesis and conclude there is sufficient evidence to indicate that the amount of current necessary is not 300 microamps at $\alpha = 0.05$.

P-value = 2P(t < 3.46): for degrees of freedom of 9 we obtain
2(0.025) < P-value < 2(0.005); 0.005 < P-value < 0.010

8-31.



The normality assumption appears to be satisfied. This is evident by the fact that the data fall along a straight line.

a) 1) The parameter of interest is the true mean level of polyunsaturated fatty acid, μ.
2) $H_0 : \mu = 17$
3) $H_1 : \mu \neq 17$
4) $\alpha = 0.01$
5) $t_0 = \dfrac{\bar{x} - \mu}{s / \sqrt{n}}$
6) Reject H_0 if $t_0 < -t_{\alpha/2,n-1}$ where $-t_{0.005,5} = -4.032$ or $t > t_{\alpha/2,n-1}$ where $t_{0.005,15} = 4.032$
7) $\bar{x} = 16.98 \quad s = 0.319 \quad n = 6$

$$t_0 = \frac{16.98 - 17}{0.319 / \sqrt{6}} = -0.154$$

8) Since $-4.032 < -0.154 < 4.032$, accept the null hypothesis and conclude the true mean level is not significantly different from 17% at $\alpha = 0.01$.

b) P-value = $2P(t > 0.154)$: for degrees of freedom of 5 we obtain

$$2(0.40) < \text{P-value}$$
$$0.80 < \text{P-value}$$

c) Using the OC curves on Chart VIb, with $d = \dfrac{0.5}{0.319} = 1.567$, $n = 10$, when $\beta \cong 0.1$. Therefore, the current sample size of 6 is inadequate.

d) For $\alpha = 0.01$, $t_{\alpha/2,n-1} = t_{0.005,5} = 4.032$

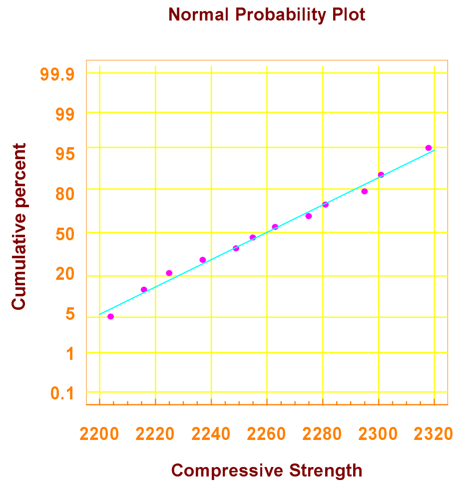
$$\bar{x} - t_{0.005,5}\left(\frac{s}{\sqrt{n}}\right) \le \mu \le \bar{x} + t_{0.005,5}\left(\frac{s}{\sqrt{n}}\right)$$

$$16.98 - 4.032\left(\frac{0.319}{\sqrt{6}}\right) \le \mu \le 16.98 + 4.032\left(\frac{0.319}{\sqrt{6}}\right)$$

$$16.455 \le \mu \le 17.505$$

With 99% confidence, we believe the true mean level of polyunsaturated fatty acid is between 16.455% and 17.505%.

8-32.. a) According to the normal probability plot, the data appear to follow a normal distribution. This is evident by the fact that the data fall along a straight line.



b) 1) The parameter of interest is the true mean compressive strength, μ.

2) $H_0 : \mu = 2250$

3) $H_1 : \mu \neq 2250$

4) $\alpha = 0.05$

5) $t_0 = \dfrac{\bar{x} - \mu}{s / \sqrt{n}}$

6) Reject H_0 if $t_0 < -t_{\alpha/2,n-1}$ where $-t_{0.025,11} = -2.201$ or $t > t_{\alpha/2,n-1}$ where $t_{0.025,11} = 2.201$

7) $\bar{x} = 2260$ $s = 35.57$ $n = 12$

$$t_0 = \frac{2260 - 2250}{35.57 / \sqrt{12}} = 0.974$$

8) Since $-2.201 < 0.974 < 2.201$, do not reject the null hypothesis and conclude there is insufficient evidence to indicate that the true mean compressive strength is significantly different from 2250 psi at $\alpha = 0.05$.

c) For $\alpha = 0.05$ and $n = 12$, $t_{\alpha/2,n-1} = t_{0.025,11} = 2.201$

$$\bar{x} - t_{0.025,11}\left(\frac{s}{\sqrt{n}}\right) \le \mu \le \bar{x} + t_{0.025,11}\left(\frac{s}{\sqrt{n}}\right)$$

$$2260 - 2.201\left(\frac{35.57}{\sqrt{12}}\right) \le \mu \le 2260 + 2.201\left(\frac{35.57}{\sqrt{12}}\right)$$

$2237.4 \le \mu \le 2282.6$

With 95% confidence, we believe the true mean compressive strength is between 2237.4 psi and 2282.6 psi.

d) For $\alpha = 0.05$ and $n = 12$, $t_{\alpha,n-1} = t_{0.05,11} = 1.796$

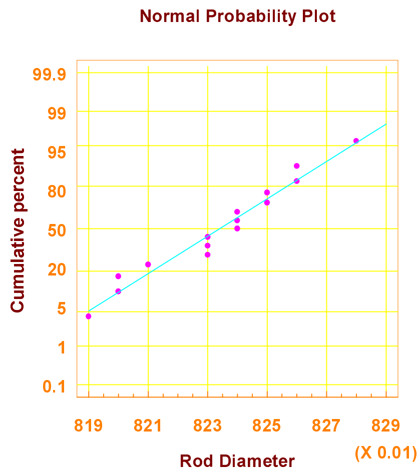
$$\bar{x} - t_{0.05,11}\left(\frac{s}{\sqrt{n}}\right) \le \mu$$

$$16.98 - 1.796\left(\frac{35.57}{\sqrt{12}}\right) \le \mu$$

$2241.56 \le \mu$

With 95% confidence, we believe the true mean compressive strength to be at least 2241.56 psi.

8-33. a) According to the normal probability plot there does not seem to be a severe deviation from normality for this data. This is evident by the fact that the data appears to fall along a straight line.



b) 1) The parameter of interest is the true mean rod diameter, μ.

2) $H_0 : \mu = 8.2$

3) $H_1 : \mu > 8.2$

4) $\alpha = 0.05$

5) $t_0 = \dfrac{\overline{x} - \mu}{s/\sqrt{n}}$

6) Reject H_0 if $t_0 > t_{\alpha,n-1}$ where $t_{0.05,14} = 1.761$

7) $\overline{x} = 8.23$ $s = 0.025$ $n = 15$

$$t_0 = \frac{8.23 - 8.2}{0.025/\sqrt{15}} = 4.65$$

8) Since 4.65 > 1.761, reject the null hypothesis and conclude the true mean rod diameter exceeds 8.2 mm at the 0.05 level of significance.

c) P-value = $P(t > 4.65)$: for degrees of freedom of 14: P-value < 0.0005

d) For $\alpha = 0.05$ and $n = 25$, $t_{\alpha/2,n-1} = t_{0.025,14} = 2.145$

$$\overline{x} - t_{0.025,14}\left(\frac{s}{\sqrt{n}}\right) \le \mu \le \overline{x} + t_{0.025,14}\left(\frac{s}{\sqrt{n}}\right)$$

$$8.23 - 2.145\left(\frac{0.025}{\sqrt{15}}\right) \le \mu \le 8.23 + 2.145\left(\frac{0.025}{\sqrt{15}}\right)$$

$$8.216 \le \mu \le 8.244$$

With 95% confidence, we believe the true mean rod diameter is between 8.216 mm and 8.244 mm.

8-34.. a) 1) The parameter of interest is the true mean wall thickness, μ.

2) $H_0 : \mu = 4.0$

3) $H_1 : \mu > 4.0$

4) $\alpha = 0.05$

5) $t_0 = \dfrac{\overline{x} - \mu}{s/\sqrt{n}}$

6) Reject H_0 if $t_0 > t_{\alpha,n-1}$ where $t_{0.05,24} = 1.711$

7) $\overline{x} = 4.05$ $s = 0.08$ $n = 25$

$$t_0 = \frac{4.05 - 4.0}{0.08/\sqrt{25}} = 3.125$$

8) Since 3.125 > 1.711, reject the null hypothesis and conclude there is sufficient evidence to indicate that the true mean wall thickness is greater than 8.23 mm at $\alpha = 0.05$.

$$\text{P-value} = P(t > 3.125);\ \ 0.001 < \text{P-value} < 0.0025$$

b) $t_{\alpha/2,n-1} = t_{0.025,24} = 2.064$

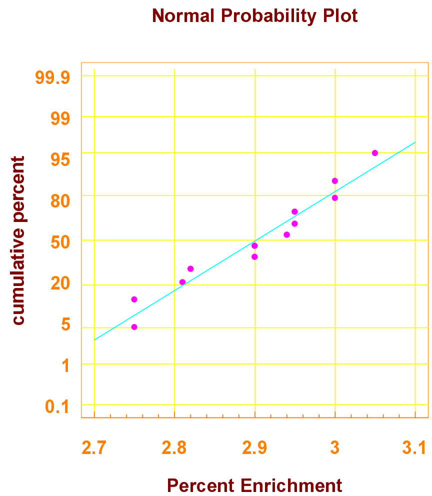
$$\bar{x} - t_{0.025,24}\left(\frac{s}{\sqrt{n}}\right) \le \mu \le \bar{x} + t_{0.025,24}\left(\frac{s}{\sqrt{n}}\right)$$

$$4.05 - 2.064\left(\frac{0.08}{\sqrt{25}}\right) \le \mu \le 4.05 + 2.064\left(\frac{0.08}{\sqrt{25}}\right)$$

$4.017 \le \mu \le 4.083$

With 95% confidence, we believe the true mean wall thickness is between 4.017 mm and 4.083 mm.

8-35. a)



The normality assumption appears to be reasonable. This is evident by the fact that the data appear to fall along a straight line.

b) 1) The parameter of interest is the true mean percent enrichment, μ.

2) $H_0 : \mu = 2.95$

3) $H_1 : \ \mu \ne 2.95$

4) $\alpha = 0.05$

5) $t_0 = \dfrac{\bar{x} - \mu}{s/\sqrt{n}}$

6) Reject H_0 if $t_0 < -t_{\alpha/2,n-1}$ where $-t_{0.025,24} = -2.201$ or $t > t_{\alpha/2,n-1}$ where $t_{0.025,24} = 2.201$

7) $\bar{x} = 2.9 \quad s = 0.099 \quad n = 12$

$$t_0 = \frac{2.9 - 2.95}{0.099/\sqrt{12}} = -1.748$$

8) Since $-2.201 < 1.748 < 2.201$, accept the null hypothesis and conclude there is no strong evidence to indicate that the true mean percent enrichment is significantly different from 2.95 at $\alpha = 0.05$.

c) For $\alpha = 0.01$ and $n = 12$, $t_{\alpha/2,n-1} = t_{0.005,11} = 3.106$

$$\bar{x} - t_{0.005,11}\left(\frac{s}{\sqrt{n}}\right) \le \mu \le \bar{x} + t_{0.005,11}\left(\frac{s}{\sqrt{n}}\right)$$

$$2.9 - 3.106\left(\frac{0.099}{\sqrt{12}}\right) \le \mu \le 2.9 + 3.106\left(\frac{0.099}{\sqrt{12}}\right)$$

$2.811 \le \mu \le 2.99$

With 99% confidence, we believe the true mean percent enrichment is between 2.811% and 2.99%. Since the interval contains the value 2.95% with a large level of confidence, we conclude that the mean percent enrichment is not significantly different 2.95%.

8-36. a) 1) The parameter of interest is the true mean quantity of syrup, μ.

2) $H_0 : \mu = 1.0$

3) $H_1 : \mu \neq 1.0$

4) $\alpha = 0.05$

5) $t_0 = \dfrac{\overline{x} - \mu}{s / \sqrt{n}}$

6) Reject H_0 if $t_0 < -t_{\alpha/2,n-1}$ where $t_{0.025,24} = -2.064$ or $t > t_{\alpha/2,n-1}$ where $t_{0.025,24} = 2.064$

7) $\overline{x} = 1.10$ $s = 0.015$ $n = 25$

$$t_0 = \frac{1.10 - 1.0}{0.015 / \sqrt{25}} = 33.3$$

8) Since 33.3 > 2.201, reject the null hypothesis and conclude the true mean quantity of syrup is significantly different from 1.0 fl. oz. at $\alpha = 0.05$

b) 2) $H_0 : \mu = 1.0$

3) $H_1 : \mu > 1.0$

4) $\alpha = 0.05$

5) $t_0 = \dfrac{\overline{x} - \mu}{s / \sqrt{n}}$

6) Reject H_0 if $t > t_{\alpha/n-1}$ where $t_{0.05,24} = 1.711$

7) $t_0 = 33.3$

8) Since 33.3 > 1.711, reject the null hypothesis and conclude there is sufficient evidence to support the claim that the mean quantity of syrup dispensed exceeds 1.0 fl. oz. at $\alpha = 0.05$.

c) Using the OC curve in Chart VI e), with d = 0.05/0.015 = 3.33 and n = 25 we get $\beta \cong 0$. Thus the power is given by $1 - \beta = 1$. The sample size is adequate for the experiment.

d) For $\alpha = 0.05$ and n = 25, $t_{\alpha/2,n-1} = t_{0.025,24} = 2.064$

$$\overline{x} - t_{0.025,24}\left(\frac{s}{\sqrt{n}}\right) \leq \mu \leq \overline{x} + t_{0.025,24}\left(\frac{s}{\sqrt{n}}\right)$$

$$1.10 - 2.064\left(\frac{0.015}{\sqrt{25}}\right) \leq \mu \leq 1.10 + 2.064\left(\frac{0.015}{\sqrt{25}}\right)$$

$$1.09 \leq \mu \leq 1.106$$

With 99% confidence, we believe the true mean quantity is between 1.09 fl. oz. and 1.106 fl. oz.

8-37. The parameter of interest is the true mean natural frequency, μ.

For $\alpha = 0.10$ and n = 5, $t_{\alpha/2,n-1} = t_{0.05,4} = 2.132$

$$\overline{x} - t_{0.05,4}\left(\frac{s}{\sqrt{n}}\right) \leq \mu \leq \overline{x} + t_{0.05,4}\left(\frac{s}{\sqrt{n}}\right)$$

$$231.67 - 2.132\left(\frac{1.53}{\sqrt{5}}\right) \leq \mu \leq 231.67 + 2.132\left(\frac{1.53}{\sqrt{5}}\right)$$

$$230.21 \leq \mu \leq 233.13$$

With 90% confidence, we believe the true mean frequency is between 230.21 Hz and 233.13 Hz. The claim that the mean natural frequency is 253 Hz cannot be supported with a level of confidence of 90% since this interval did not contain the value 253.

Section 8-4

8-38. a) In order to use χ^2 statistic in hypothesis testing and confidence interval construction, we need to assume that the underlying distribution is normal.

1) The parameter of interest is the true standard deviation of the diameter, σ. However, the answer can be found by performing a hypothesis test on σ^2.

2) H_0 : $\sigma^2 = 0.0001$

3) H_1 : $\sigma^2 > 0.0001$

4) $\alpha = 0.01$

5) $\chi_0^2 = \dfrac{(n-1)s^2}{\sigma^2}$

6) Reject H_0 if $\chi_0^2 > \chi^2_{\alpha,n-1}$ where $\chi^2_{0.01,14} = 29.14$

7) n = 15, $s^2 = 0.008$

$$\chi_0^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{14(0.008)^2}{0.0001} = 8.96$$

8) Since 8.96 < 29.14 do not reject H_0 and conclude there is insufficient evidence to indicate the true standard deviation of the diameter exceeds 0.01 at $\alpha = 0.01$.

b) P-value = $P(\chi^2 > 8.96)$ for 14 degrees of freedom: $0.5 < \text{P-value} < 0.9$

c) 99% lower confidence interval on σ^2 :

For $\alpha = 0.01$ and n = 15, $\chi^2_{\alpha,n-1} = \chi^2_{0.01,14} = 29.14$

$$\frac{14(0.008)^2}{29.14} < \sigma^2$$

$0.00003075 < \sigma^2$

With 99% confidence, we believe the true variance of the hole diameter is more than 0.00003075 mm.

8-39. a) In order to use χ^2 statistic in hypothesis testing and confidence interval construction, we need to assume that the underlying distribution is normal.

1) The parameter of interest is the true variance of the sugar content, σ^2 .

2) H_0 : $\sigma^2 = 18$

3) H_1 : $\sigma^2 \neq 18$

4) $\alpha = 0.05$

5) $\chi_0^2 = \dfrac{(n-1)s^2}{\sigma^2}$

6) Reject H_0 if $\chi_0^2 < \chi^2_{1-\alpha/2,n-1}$ where $\chi_{0.975,9} = 2.70$ or $\chi_0^2 > \chi^2_{\alpha,2,n-1}$ where $\chi_{0.025,9} = 19.02$

7) n = 10, $s^2 = 23.04$

$$\chi_0^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{9(23.04)}{18} = 11.52$$

8) Since 2.70 < 11.52 < 19.02 accept H_0 and conclude the evidence indicates the true variance of the sugar content is not significantly different from 18 mg^2 at $\alpha = 0.05$.

b) P-value = $2P(\chi^2 > 11.52)$ for 9 degrees of freedom

$$0.1 < P(\chi^2 > 11.52) < 0.5$$
$$0.2 < \text{P-value} < 1.0$$

c) 95% confidence interval for σ:

First find a confidence interval for σ^2 :

For $\alpha = 0.05$ and n = 10, $\chi^2_{\alpha/2,n-1} = \chi^2_{0.025,9} = 19.02$ and $\chi^2_{1-\alpha/2,n-1} = \chi^2_{0.975,9} = 2.70$

$$\frac{9(23.04)}{19.02} \le \sigma^2 \le \frac{9(23.04)}{2.70}$$
$$10.9 \le \sigma^2 \le 76.8$$

Take the square root of the endpoints of this interval to find the confidence interval for σ:

$$3.3 \le \sigma \le 8.76$$

With 95% confidence, we believe the true standard deviation of the sugar content is between 3.3 mg and 8.76 mg.

8-40. a) In order to use χ^2 statistic in hypothesis testing and confidence interval construction, we need to assume that the underlying distribution is normal.

1) The parameter of interest is the standard deviation of tire life, σ. However, the answer can be found by performing a hypothesis test on σ^2.

2) $H_0 : \sigma^2 = 40{,}000$

3) $H_1 : \sigma^2 > 40{,}000$

4) $\alpha = 0.05$

5) $\chi_0^2 = \dfrac{(n-1)s^2}{\sigma^2}$

6) Reject H_0 if $\chi_0^2 > \chi^2_{\alpha,n-1}$ where $\chi^2_{0.05,15} = 25.00$

7) $n = 16$, $s^2 = (3645.94)^2$

$$\chi_0^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{15(3645.94)^2}{40000} = 4984.83$$

8) Since 4984.83 > 25.00 reject H_0 and conclude there is strong evidence to indicate the true standard deviation of tire life exceeds 200 km at $\alpha = 0.05$.

b) P-value = $P(\chi^2 > 4984.83)$ for 15 degrees of freedom P-value < 0.005

c) 95% lower confidence interval for σ^2:

For $\alpha = 0.05$ and $n = 16$, $\chi^2_{\alpha,n-1} = \chi^2_{0.05,15} = 25$

$$\frac{15(3645.94)^2}{25} < \sigma^2$$

$7{,}203{,}485.52 < \sigma^2$

With 95% confidence, we believe the true variance of tire life is more than 7,203,485.52 km^2

8-41. a) In order to use χ^2 statistic in hypothesis testing and confidence interval construction, we need to assume that the underlying distribution is normal.

1) The parameter of interest is the true standard deviation of Izod strength, σ. However, the answer can be found by performing a hypothesis test on σ^2.

2) $H_0 : \sigma^2 = 0.01$

3) $H_1 : \sigma^2 \neq 0.01$

4) $\alpha = 0.01$

5) $\chi_0^2 = \dfrac{(n-1)s^2}{\sigma^2}$

6) Reject H_0 if $\chi_0^2 < \chi^2_{1-\alpha/2,n-1}$ where $\chi^2_{0.995,19} = 6.84$ or $\chi_0^2 > \chi^2_{\alpha,2,n-1}$ where $\chi^2_{0.005,19} = 38.58$

7) $n = 20$, $s = 0.25$

$$\chi_0^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{19(0.25)^2}{0.01} = 118.75$$

8) Since 118.75 > 38.58 we would reject H_0 and conclude the true standard deviation of Izod strength is significantly different from 0.10 ft-lb/in at $\alpha = 0.01$.

b) P-value = $2P(\chi^2 > 118.75)$ for 19 degrees of freedom

$$P(\chi^2 > 118.75) < 0.005$$
$$\text{P-value} < 0.01$$

c) 99% confidence interval for σ^2 :

For $\alpha = 0.01$ and $n = 20$, $\chi^2_{\alpha/2,n-1} = \chi^2_{0.005,19} = 38.58$ and $\chi^2_{1-\alpha/2,n-1} = \chi^2_{0.995,19} = 6.84$

$$\frac{20(0.25)^2}{38.58} \leq \sigma^2 \leq \frac{20(0.25)^2}{6.84}$$

$0.031 \leq \sigma^2 \leq 0.176$

With 99% confidence, we believe the true variance of Izod strength is between 0.031 $(\text{ft-lb/in})^2$ and 0.176 (ft-lb/in).

8-42. a) In order to use χ^2 statistic in hypothesis testing and confidence interval construction, we need to assume that the underlying distribution is normal.

1) The parameter of interest is the true standard deviation of titanium percentage, σ. However, the answer can be found by performing a hypothesis test on σ^2.

2) $H_0: \sigma^2 = (0.25)^2$

3) $H_1: \sigma^2 \neq (0.25)^2$

4) $\alpha = 0.01$

5) $\chi_0^2 = \dfrac{(n-1)s^2}{\sigma^2}$

6) Reject H_0 if $\chi_0^2 < \chi^2_{1-\alpha/2,n-1}$ where $\chi^2_{0.995,50} = 27.99$ or $\chi_0^2 > \chi^2_{\alpha,2,n-1}$ where $\chi^2_{0.005,50} = 79.49$

7) n = 51, s = 0.37

$$\chi_0^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{50(0.37)^2}{(0.25)^2} = 109.52$$

8) Since 109.52 > 79.49 we would reject H_0 and conclude there is sufficient evidence to indicate the true standard deviation of titanium percentage is significantly different from 0.25 at $\alpha = 0.01$.

b) 95% confidence interval for σ:

First find the confidence interval for σ^2 :

For $\alpha = 0.05$ and n = 51, $\chi^2_{\alpha/2,n-1} = \chi^2_{0.025,50} = 71.42$ and $\chi^2_{1-\alpha/2,n-1} = \chi^2_{0.975,50} = 32.36$

$$\frac{50(0.37)^2}{(71.42)^2} \le \sigma^2 \le \frac{50(0.37)^2}{(32.36)^2}$$

$0.096 \le \sigma^2 \le 0.2115$

Taking the square root of the endpoints of this interval we obtain,

$$0.31 < \sigma < 0.46$$

With 95% confidence, we believe the true standard deviation of titanium percentage is between 0.31 and 0.46.

8-43. Using the appropriate chart in the Appendix, with n = 15 and λ =1.5, the probability of detecting this difference is 0.6.

8-44. Using the appropriate chart in the Appendix, with $\lambda = \sqrt{\dfrac{40}{18}} = 1.49$ and $\beta = 0.10$, we find n = 30.

Section 8-5

8-45. $\hat{p} = \dfrac{823}{1000} = 0.823$ n = 1000 For $\alpha = 0.05$, $z_{\alpha/2} = z_{0.025} = 1.96$

$$0.823 - 1.96\sqrt{\frac{0.823(0.177)}{1000}} \le p \le 0.823 + 1.96\sqrt{\frac{0.823(0.177)}{1000}}$$

$0.799 \le p \le 0.847$

With 95% confidence, we believe the true proportion of deaths resulting from cancer is between 0.799 and 0.847.

8-46. E = 0.03, $\alpha = 0.05$, $z_{\alpha/2} = z_{0.025} = 1.96$ and $\hat{p} = 0.823$ as the initial estimate of p,

$$n = \left(\frac{z_{\alpha/2}}{E}\right)^2 \hat{p}(1-\hat{p}) = \left(\frac{1.96}{0.03}\right)^2 0.823(1-0.823) = 621.79 ,$$

$n \cong 622$.

8-47. $n = 50$, $x = 18$, $\hat{p} = \frac{18}{50} = 0.36$, $\alpha = 0.05$, $z_{\alpha/2} = z_{0.025} = 1.96$

a) $0.36 - 1.96\sqrt{\frac{0.36(0.64)}{50}} \le p \le 0.36 + 1.96\sqrt{\frac{0.36(0.64)}{50}}$

$0.23 \le p \le 0.49$

With 95% confidence, we believe the true proportion of helmets damaged lies within 0.23 and 0.49.

b) $n = \left(\frac{z_{\alpha/2}}{E}\right)^2 \hat{p}(1-\hat{p}) = \left(\frac{1.96}{0.02}\right)^2 0.36(1-0.36) = 2212.76$, $n \cong 2213$

c) $n = \left(\frac{z_{\alpha/2}}{E}\right)^2 p(1-p) = \left(\frac{1.96}{0.02}\right)^2 0.5(1-0.5) = 2401$, $n \cong 2401$

8-48. The worst case would be for $p = 0.5$, thus with $E = 0.05$ and $\alpha = 0.01$, $z_{\alpha/2} = z_{0.005} = 2.58$ we obtain a sample size of:

$$n = \left(\frac{z_{\alpha/2}}{E}\right)^2 p(1-p) = \left(\frac{2.58}{0.05}\right)^2 0.5(1-0.5) = 665.64, \quad n \cong 666$$

8-49. $\hat{p} = \frac{10}{800} = 0.0125$, $n = 800$, $\alpha = 0.01$, $z_\alpha = z_{0.01} = 2.33$

The parameter of interest is the true fraction defective, p.

$p \le 0.0125 + 2.33\sqrt{\frac{0.0125(0.9875)}{800}}$

$p \le 0.0217$

With 99% confidence, the true fraction defective is at most 2.17%.

8-50. $E = 0.017$, $\alpha = 0.01$, $z_{\alpha/2} = z_{0.005} = 2.58$

$$n = \left(\frac{z_{\alpha/2}}{E}\right)^2 p(1-p) = \left(\frac{2.58}{0.017}\right)^2 0.5(1-0.5) = 5758.13, \quad n \cong 5759$$

8-51. a) 1) The parameter of interest is the true fraction of defective integrated circuits, p.

2) $H_0 : p = 0.05$

3) $H_1 : p \ne 0.05$

4) $\alpha = 0.05$

5) $z_0 = \frac{x - np_0}{\sqrt{np_0(1-p_0)}}$ or $z_0 = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$; Either approach will yield the same conclusion

6) Reject H_0 if $z_0 < -z_{\alpha/2}$ where $-z_{\alpha/2} = -z_{0.025} = -1.96$ or $z_0 > z_{\alpha/2}$ where $z_{\alpha/2} = z_{0.025} = 1.96$

7) $x = 13$ $n = 300$ $\hat{p} = \frac{13}{300} = 0.043$

$$z_0 = \frac{x - np_0}{\sqrt{np_0(1-p_0)}} = \frac{13 - 300(0.05)}{\sqrt{300(0.05)(0.95)}} = -0.53$$

8) Since $-1.96 < -0.53 < 1.96$, accept the null hypothesis and conclude the true fraction of defective integrated circuits is not significantly different from 0.05, at $\alpha = 0.05$.

b) P-value = $2(\Phi(-0.53)) = 2(1 - \Phi(0.53)) = 2(1 - 0.70194) = 0.59612$.

Again since the P-value is greater than the level of significance, we would not reject the null hypothesis.

8-52. a) Using the information from Exercise 8-51, test

2) $H_0 : p = 0.05$
3) $H_1 : p < 0.05$
4) $\alpha = 0.05$
5) $z_0 = \dfrac{x - np_0}{\sqrt{np_0(1-p_0)}}$ or $z_0 = \dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}}$; Either approach will yield the same conclusion
6) Reject H_0 if $z_0 < -z_\alpha$ where $-z_\alpha = -z_{0.05} = -1.65$
7) $x = 13$ $n = 300$ $\hat{p} = \dfrac{13}{300} = 0.043$

$$z_0 = \frac{x - np_0}{\sqrt{np_0(1-p_0)}} = \frac{13 - 300(0.05)}{\sqrt{300(0.05)(0.95)}} = -0.53$$

8) Since $-0.53 > -1.65$, accept the null hypothesis and conclude the true fraction of defective integrated circuits is not significantly less than 0.05, at $\alpha = 0.05$.

b) P-value = $1 - \Phi(0.53) = 0.29806$

8-53. a) 1) The parameter of interest is the true proportion of engineers who continue their education, p.

2) $H_0 : p = 0.50$
3) $H_1 : p \neq 0.50$
4) $\alpha = 0.05$
5) $z_0 = \dfrac{x - np_0}{\sqrt{np_0(1-p_0)}}$ or $z_0 = \dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}}$; Either approach will yield the same conclusion
6) Reject H_0 if $z_0 < -z_{\alpha/2}$ where $-z_{\alpha/2} = -z_{0.025} = -1.96$ or $z_0 > z_{\alpha/2}$ where $z_{\alpha/2} = z_{0.025} = 1.96$
7) $x = 117$ $n = 484$ $\hat{p} = \dfrac{117}{484} = 0.242$

$$z_0 = \frac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}} = \frac{0.242 - 0.50}{\sqrt{\dfrac{0.5(1-0.5)}{484}}} = -11.352$$

8) Since $-11.352 < -1.96$, reject the null hypothesis and conclude the data from "Engineering Horizons" yield results significantly different from the claim reported by "Fortune", at $\alpha = 0.05$.

b) P-value = $2(1 - \Phi(11.352)) = 2(1 - 1) = 0$

8-54. a) 1) The parameter of interest is the true percentage of polished lenses that contain surface defects, p.

2) $H_0 : p = 0.02$
3) $H_1 : p < 0.02$
4) $\alpha = 0.05$
5) $z_0 = \dfrac{x - np_0}{\sqrt{np_0(1-p_0)}}$ or $z_0 = \dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}}$; Either approach will yield the same conclusion
6) Reject H_0 if $z_0 < -z_\alpha$ where $-z_\alpha = -z_{0.05} = -1.65$
7) $x = 6$ $n = 250$ $\hat{p} = \dfrac{6}{250} = 0.024$

$$z_0 = \frac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}} = \frac{0.024 - 0.02}{\sqrt{\dfrac{0.02(1-0.02)}{250}}} = 0.452$$

8) Since $0.452 > -1.65$ accept the null hypothesis and conclude the machine cannot be qualified at the 0.05 level of significance.

b) P-value = $\Phi(0.452) = 0.67364$

8-55. a) 1) The parameter of interest is the true percentage of football helmets that have flaws, p.

2) $H_0 : p = 0.10$

3) $H_1 : p > 0.10$

4) $\alpha = 0.01$

5) $z_0 = \dfrac{x - np_0}{\sqrt{np_0(1-p_0)}}$ or $z_0 = \dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}}$; Either approach will yield the same conclusion

6) Reject H_0 if $z_0 > z_\alpha$ where $z_\alpha = z_{0.01} = 2.33$

7) $x = 16$ $n = 200$ $\hat{p} = \dfrac{16}{200} = 0.08$

$$z_0 = \frac{x - np_0}{\sqrt{np_0(1-p_0)}} = \frac{16 - 200(0.10)}{\sqrt{200(0.10)(0.90)}} = -0.943$$

8) Since $-0.943 < 2.33$ accept the null hypothesis and conclude there is insufficient evidence to support the claim that at least 10% of all football helmets have manufacturing flaws that could potentially cause injury to the wearer, at $\alpha = 0.01$.

b) P-value $= 1 - \Phi(-0.943) = \Phi(0.943) = 0.82639$.

8-56. The problem statement implies H_0: $p = 0.6$, H_1: $p > 0.6$ and defines an acceptance region as $\hat{p} \le \dfrac{315}{500} = 0.63$ and rejection region as $\hat{p} > 0.63$

a) $\alpha = P(\hat{P} \ge 0.63 \text{ when } p = 0.6) = P\left(Z \ge \dfrac{0.63 - 0.6}{\sqrt{\dfrac{0.6(0.4)}{500}}}\right) = P(Z \ge 1.37) = 1 - P(Z < 1.37) = 0.08535$.

b) $\beta = P(\hat{P} \le 0.63 \text{ when } p = 0.75) = P(Z \le -6.196) = 0$.

8-57. 1) The parameter of interest is the true percentage of batteries that will fail during the warranty period, p.

2) $H_0 : p = 0.002$

3) $H_1 : p < 0.002$

4) $\alpha = 0.01$

5) $z_0 = \dfrac{x - np_0}{\sqrt{np_0(1-p_0)}}$ or $z_0 = \dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}}$; Either approach will yield the same conclusion

6) Reject H_0 if $z_0 < -z_\alpha$ where $-z_\alpha = -z_{0.01} = -2.33$

7) $x = 15$ $n = 5000$ $\hat{p} = \dfrac{15}{5000} = 0.003$

$$z_0 = \frac{x - np_0}{\sqrt{np_0(1-p_0)}} = \frac{15 - 5000(0.002)}{\sqrt{5000(0.002)(0.998)}} = 1.583$$

8) Since $1.583 > -2.33$ do not reject the null hypothesis and conclude there is insufficient evidence to support the claim that less than 0.2 percent of the company's batteries will fail during the warranty period, with proper charging procedures at $\alpha = 0.01$.

Section 8-7

8-58.

Value	0	1	2	3	4
Observed Frequency	24	30	31	11	4
Expected Frequency	30.12	36.14	21.69	8.67	2.60

Since value 4 has an expected frequency less than 3, combine this category with the previous category:

Value	0	1	2	3-4
Observed Frequency	24	30	31	15
Expected Frequency	30.12	36.14	21.69	11.67

The degrees of freedom are $k - p - 1 = 4 - 0 - 1 = 3$

a) 1) The variable of interest is the form of the distribution for X.
2) H_0: The form of the distribution is Poisson
3) H_1: The form of the distribution is not Poisson
4) $\alpha = 0.05$
5) The test statistic is

$$\chi_0^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$$

6) Reject H_0 if $\chi_o^2 > \chi_{0.05,3}^2 = 7.81$
7)

$$\chi_0^2 = \frac{(24 - 30.12)^2}{30.12} + \cdots + \frac{(15 - 11.67)^2}{11.67} = 7.23$$

8) Since $7.23 < 7.81$ do not reject H_0. We are unable to reject the null hypothesis that the distribution of X is Poisson.

b) P-value = 0.0649 (found using Minitab)

8-59. Estimated mean = 4.907

Value	1	2	3	4	5	6	7	8
Observed Frequency	1	11	8	13	11	12	10	9
Expected Frequency	2.7214	6.6770	10.9213	13.3977	13.1485	10.7533	7.5381	4.6237

Since the first category has an expected frequency less than 3, combine it with the next category:

Value	1-2	3	4	5	6	7	8
Observed Frequency	12	8	13	11	12	10	9
Expected Frequency	9.3984	10.9213	13.3977	13.1485	10.7533	7.5381	4.6237

The degrees of freedom are $k - p - 1 = 7 - 1 - 1 = 5$

a) 1) The variable of interest is the form of the distribution for the number of flaws.
2) H_0: The form of the distribution is Poisson
3) H_1: The form of the distribution is not Poisson
4) $\alpha = 0.01$
5) The test statistic is

$$\chi_0^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$$

6) Reject H_0 if $\chi_o^2 > \chi_{0.01,5}^2 = 15.09$

7)

$$\chi_0^2 = \frac{(12-9.3984)^2}{9.3984} + \cdots + \frac{(9-4.6237)^2}{4.6237} = 6.955$$

8) Since 6.955 < 15.09 do not reject H_0. We are unable to reject the null hypothesis that the distribution of the number of flaws is Poisson.

b) P-value = 0.2237 (found using Minitab)

8-60. Estimated mean = 10.131

Value	5	6	8	9	10	11	12	13	14	15
Rel. Freq	0.067	0.067	0.100	0.133	0.200	0.133	0.133	0.067	0.033	0.067
Observed (Days)	2	2	3	4	6	4	4	2	1	2
Expected (Days)	1.0626	1.7942	3.2884	3.7016	3.7501	3.4538	2.9159	2.2724	1.6444	1.1106

Since there are several cells with expected frequencies less than 3, the revised table would be:

Value	5-8	9	10	11	12-15
Observed (Days)	7	4	6	4	9
Expected (Days)	6.1452	3.7016	3.7501	3.4538	7.9433

The degrees of freedom are $k - p - 1 = 5 - 1 - 1 = 3$

a) 1) The variable of interest is the form of the distribution for the number of calls arriving to a switchboard from noon to 1pm during business days.
2) H_0: The form of the distribution is Poisson
3) H_1: The form of the distribution is not Poisson
4) $\alpha = 0.05$
5) The test statistic is

$$\chi_0^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$$

6) Reject H_0 if $\chi_o^2 > \chi_{0.05,3}^2 = 7.81$
7)

$$\chi_0^2 = \frac{(7-6.1452)^2}{6.1452} + \cdots + \frac{(9-7.9433)^2}{7.9433} = 1.72$$

8) Since 1.72 < 7.81 do not reject H_0. We are unable to reject the null hypothesis that the distribution for the number of calls is Poisson.

b) P-value = 0.6325 (found using Minitab)

8-61. Mean = np = 6(0.25) = 1.5

Value	0	1	2	3	4
Observed	4	21	10	13	2
Expected	8.8989	17.7979	14.8315	6.5918	1.6479

The expected frequency for value 4 is less than 3. Combine this cell with value 3:

Value	0	1	2	3-4
Observed	4	21	10	15
Expected	8.8989	17.7979	14.8315	8.2397

The degrees of freedom are $k - p - 1 = 4 - 0 - 1 = 3$

a) 1) The variable of interest is the form of the distribution for the random variable X.
2) H_0: The form of the distribution is binomial with n = 6 and p = 0.25
3) H_1: The form of the distribution is not binomial with n = 6 and p = 0.25
4) $\alpha = 0.05$
5) The test statistic is

$$\chi_0^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$$

6) Reject H_0 if $\chi_o^2 > \chi_{0.05,3}^2 = 7.81$
7)

$$\chi_0^2 = \frac{(4 - 8.8989)^2}{8.8989} + \cdots + \frac{(15 - 8.2397)^2}{8.2397} = 10.39$$

8) Since 10.39 > 7.81 reject H_0. We can conclude that the distribution is not binomial with n = 6 and p = 0.25 at $\alpha = 0.05$.
b) P-value = 0.0155 (found using Minitab)

8-62. The value of p must be estimated. Let the estimate be denoted by $\hat{p}_{sample}$

$$\text{sample mean} = \frac{0(39) + 1(23) + 2(12) + 3(1)}{75} = 0.6667$$

$$\hat{p}_{sample} = \frac{\text{sample mean}}{n} = \frac{0.6667}{24} = 0.02778$$

Value	0	1	2	3
Observed	39	23	12	1
Expected	38.1426	26.1571	8.5952	1.8010

Since value 3 has an expected frequency less than 3, combine this category with that of value 2:

Value	0	1	2-3
Observed	39	23	13
Expected	38.1426	26.1571	10.3962

The degrees of freedom are k – p – 1 = 3 – 1 – 1 = 1
a) 1) The variable of interest is the form of the distribution for the number of underfilled cartons, X.
2) H_0: The form of the distribution is binomial
3) H_1: The form of the distribution is not binomial
4) $\alpha = 0.05$
5) The test statistic is

$$\chi_0^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$$

6) Reject H_0 if $\chi_o^2 > \chi_{0.05,1}^2 = 3.84$
7)

$$\chi_0^2 = \frac{(39 - 38.1426)^2}{38.1426} + \frac{(23 - 26.1571)^2}{26.1571} + \frac{(13 - 10.3962)^2}{10.39} = 1.053$$

8) Since 1.053 < 3.84 do not reject H_0. We are unable to reject the null hypothesis that the distribution of the number of underfilled cartons is binomial at $\alpha = 0.05$.

b) P-value = 0.3048 (found using Minitab)

8-63. Estimated mean = 49.6741
All expected frequencies are greater than 3.
The degrees of freedom are k – p – 1 = 26 – 1 – 1 = 24

a) 1) The variable of interest is the form of the distribution for the number of cars passing through the intersection.
2) H_0: The form of the distribution is Poisson
3) H_1: The form of the distribution is not Poisson
4) $\alpha = 0.05$
5) The test statistic is

$$\chi_0^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$$

6) Reject H_0 if $\chi_o^2 > \chi_{0.05,24}^2 = 36.42$
7) Estimated mean = 49.6741

$$\chi_0^2 = 769.57$$

8) Since 769.57 >>> 36.42, reject H_0. We can conclude that the distribution is not Poisson at $\alpha = 0.05$.
b) P-value = 0 (found using Minitab)

Section 8-8

8-64. 1. The variable of interest is breakdowns among shift.
2. H_0: Breakdowns are independent of shift.
3. H_1: Breakdowns are not independent of shift.
4. $\alpha = 0.05$
5. The test statistic is:

$$\chi_0^2 = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{(o_{ij} - E_{ij})^2}{E_{ij}}$$

6. The critical value is $\chi_{.05,6}^2 = 12.592$
7. The calculated test statistic is $\chi_0^2 = 11.65$
8. $\chi_0^2 \ngtr \chi_{0.05,6}^2$, do not reject H_0 and conclude that the data provide insufficient evidence to claim that machine breakdown and shift are dependent at $\alpha = 0.05$.
P-value = 0.070 (using Minitab)

8-65. 1. The variable of interest is calls by surgical-medical patients.
2. H_0:Calls by surgical-medical patients are independent of Medicare status.
3. H_1:Calls by surgical-medical patients are not independent of Medicare status.
4. $\alpha = 0.05$
5. The test statistic is:

$$\chi_0^2 = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{(o_{ij} - E_{ij})^2}{E_{ij}}$$

6. The critical value is $\chi_{.01,1}^2 = 6.637$
7. The calculated test statistic is $\chi_0^2 = 0.033$
8. $\chi_0^2 \ngtr \chi_{0.01,1}^2$, do not reject H_0 and conclude evidence is not sufficient to claim that surgical-medical patients and Medicare status are dependent.
P-value = 0.85

8-66. 1. The variable of interest is statistics grades and OR grades.
2. H_0: Statistics grades are independent of OR grades.
3. H_1: Statistics and OR grades are not independent.
4. $\alpha = 0.01$
5. The test statistic is:

$$\chi_0^2 = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{(o_{ij} - E_{ij})^2}{E_{ij}}$$

6. The critical value is $\chi^2_{.01,9} = 21.665$

7. The calculated critical value is $\chi^2_0 = 25.55$

8. $\chi^2_0 > \chi^2_{0.01,9}$ Therefore, reject H_0 and conclude that the grades are not independent at $\alpha = 0.01$.

P-value = 0.002

8-67. 1. The variable of interest is characteristic among deflections and ranges.
2. H_0: Deflection and range are independent.
3. H_1: Deflection and range are not independent.
4. $\alpha = 0.05$
5. The test statistic is:

$$\chi_0^2 = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{(o_{ij} - E_{ij})^2}{E_{ij}}$$

6. The critical value is $\chi^2_{.05,4} = 9.488$

7. The calculated test statistic is $\chi^2_0 = 2.46$

8. $\chi^2_0 \not> \chi^2_{0.05,4}$, do not reject H_0 and conclude evidence is not sufficient to claim that the data are not independent at $\alpha = 0.05$.
P-value = 0.652

8-68. 1. The variable of interest is failures of an electronic component.
2. H_0: Type of failure is independent of mounting position.
3. H_1: Type of failure is not independent of mounting position.
4. $\alpha = 0.01$
5. The test statistic is:

$$\chi_0^2 = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{(o_{ij} - E_{ij})^2}{E_{ij}}$$

6. The critical value is $\chi^2_{.01,3} = 11.344$

7. The calculated test statistic is $\chi^2_0 = 10.71$

8. $\chi^2_0 \not> \chi^2_{0.01,3}$, do not reject H_0 and conclude evidence is not sufficient to claim that the type of failure is not independent of the mounting position at $\alpha = 0.01$.
P-value = 0.013

8-69. 1. The variable of interest is opinion on core curriculum change.
2. H_0: Opinion of the change is independent of the class standing.
3. H_1: Opinion of the change is not independent of the class standing.
4. $\alpha = 0.05$
5. The test statistic is:

$$\chi_0^2 = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{(o_{ij}-E_{ij})^2}{E_{ij}}$$

6. The critical value is $\chi^2_{.05,3} = 7.815$

7. The calculated test statistic is $\chi^2_{.0} = 26.97$.

8. $\chi_0^2 >>> \chi^2_{0.05,3}$, reject H_0 and conclude that the opinions on the change are not independent of class standing.
P-value ≈ 0

Supplemental Exercises

8-70. Symmetric confidence interval: $\bar{x} - 1.96\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{x} + 1.96\frac{\sigma}{\sqrt{n}}$, since $z_{0.025} = 1.96$

The length of this interval is $2\left(1.96\frac{\sigma}{\sqrt{n}}\right) = 3.92\left(\frac{\sigma}{\sqrt{n}}\right)$

Asymmetric confidence interval: $\bar{x} - 2.33\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{x} + 1.76\frac{\sigma}{\sqrt{n}}$, since $-z_{0.01} = -2.325$ and $z_{0.04} = 1.75$

The length of this interval is $(2.33 + 1.76)\frac{\sigma}{\sqrt{n}} = 4.09\left(\frac{\sigma}{\sqrt{n}}\right)$

The symmetric confidence interval is the shorter of the two; the advantage to a symmetric confidence interval is the fact that in general they are shorter than an asymmetric confidence interval.

8-71. $\mu = 50$ σ unknown
a) $n = 16$ $\bar{x} = 52$ $s = 1.5$

$$t_o = \frac{52-50}{8/\sqrt{16}} = 1$$

The P-value for $t_0 = 1$, degrees of freedom = 15, is between 0.1 and 0.25. Thus we would conclude that the results are not very unusual.
b) $n = 30$

$$t_o = \frac{52-50}{8/\sqrt{30}} = 1.37$$

The P-value for $t_0 = 1.37$, degrees of freedom = 29, is between 0.05 and 0.1. Thus we would conclude that the results are somewhat unusual.
c) $n = 100$ (with $n > 30$, the standard normal table can be used for this problem)

$$z_o = \frac{52-50}{8/\sqrt{100}} = 2.5$$

The P-value for $z_0 = 2.5$, is 0.00539. Thus we would conclude that the results are very unusual.
d) For constant values of $\bar{x}$ and s, increasing only the sample size, we see that the standard error of $\bar{X}$ decreases and consequently a sample mean value of 52 when the true mean is 50 is more unusual for the larger sample sizes.

8-72. $\mu = 50 \quad \sigma^2 = 5$ Find $P(s^2 \geq 7.44)$ and $P(s^2 \leq 2.56)$

a) n = 16

$$P\left(\frac{(n-1)s^2}{\sigma^2} \geq \frac{15(7.44)}{5}\right) = P\left(\chi_{15}^2 \geq 22.32\right) \approx 0.10$$

$$P\left(\frac{(n-1)s^2}{\sigma^2} \leq \frac{15(2.56)}{5}\right) = P\left(\chi_{15}^2 \leq 7.68\right) = 1 - P\left(\chi_{15}^2 \geq 7.68\right); \quad 0.05 < P\left(\chi_{15}^2 \leq 7.68\right) < 0.10$$

b) n = 30

$$P\left(\frac{(n-1)s^2}{\sigma^2} \geq \frac{29(7.44)}{5}\right) = P\left(\chi_{29}^2 \geq 43.152\right); \quad 0.025 < P\left(\chi_{29}^2 \geq 43.152\right) < 0.05$$

$$P\left(\frac{(n-1)s^2}{\sigma^2} \leq \frac{29(2.56)}{5}\right) = P\left(\chi_{29}^2 \leq 14.85\right) = 1 - P\left(\chi_{29}^2 \geq 14.85\right); \quad 0.01 < P\left(\chi_{29}^2 \leq 14.85\right) < 0.025$$

c) n = 71

$$P\left(\frac{(n-1)s^2}{\sigma^2} \geq \frac{70(7.44)}{5}\right) = P\left(\chi_{70}^2 \geq 104.16\right); \quad 0.005 < P\left(\chi_{70}^2 \geq 104.16\right) \approx 0.005$$

$$P\left(\frac{(n-1)s^2}{\sigma^2} \leq \frac{70(2.56)}{5}\right) = P\left(\chi_{70}^2 \leq 35.84\right) = 1 - P\left(\chi_{70}^2 \geq 35.84\right); \quad P\left(\chi_{70}^2 \leq 35.84\right) < 0.005$$

d) As the sample size increases with all other values held constant, the probability $P(s^2 \geq 7.44)$ decreases because the right tail of the χ^2 distribution becomes relatively shorter.

e) As the sample size increases with all other values held constant, the probability $P(s^2 \leq 2.56)$ decreases because the left tail of the χ^2 distribution becomes relatively shorter.

8-73. a) The data appear to follow a normal distribution based on the normal probability plot since the data fall along a straight line.

b) It is important to check for normality of the distribution underlying the sample data since the confidence intervals to be constructed should have the assumption of normality for the results to be reliable (especially since the sample size is less than 30 and the central limit theorem does not apply).

c) No, with 95% confidence, we can not infer that the true mean could be 14.05 since this value is not contained within the given 95% confidence interval.

d) As with part b, to construct a confidence interval on the variance, the normality assumption must hold for the results to be reliable.

e) Yes, it is reasonable to infer that the variance could be 0.35 since the 95% confidence interval on the variance contains this value.

f) i) & ii) No, doctors and children would represent two completely different populations not represented by the population of Canadian Olympic hockey players. Since doctors nor children were the target of this study or part of the sample taken, the results should not be extended to these groups.

8-74. a) The data appear to follow a normal distribution based on the normal probability plot since the data fall along a straight line.

b) $\bar{x} = 25.12 \quad s = 8.42 \quad n = 9$; Use the t-distribution to construct the confidence intervals; $\alpha = 1 - 0.99$

$t_{\alpha,n-1} = t_{0.01,8} = 2.896$

$$\bar{x} - t_{\alpha,n-1}\left(\frac{s}{\sqrt{n}}\right) \leq \mu$$

$$25.12 - 2.896\left(\frac{8.42}{\sqrt{9}}\right) \leq \mu$$

$$17 \leq \mu$$

with 99% confidence, we believe the true mean compressive strength is at least 17 Mpa.

c) $t_{\alpha/2,n-1} = t_{0.01,8} = 2.896$

$$\bar{x} - t_{\alpha,n-1}\left(\frac{s}{\sqrt{n}}\right) \le \mu \le \bar{x} + t_{\alpha,n-1}\left(\frac{s}{\sqrt{n}}\right)$$
$$25.12 - 2.896\left(\frac{8.42}{\sqrt{9}}\right) \le \mu \le 25.12 + 2.896\left(\frac{8.42}{\sqrt{9}}\right)$$
$$17.0 \le \mu \le 33.25$$

With 98% confidence, we believe the true mean compressive strength is between 17.0 and 33.25 Mpa.

The lower endpoint of the one-sided confidence interval is the same as that of the two-sided confidence interval due to the level of confidence used. In both cases the probability in the left tail is 0.01.

d) $\chi^2_{1-\alpha,n-1} = \chi^2_{0.99,8} = 1.65$

$$\sigma^2 \le \frac{(n-1)s^2}{\chi^2_{1-\alpha.n-1}}$$
$$\sigma^2 \le \frac{8(8.42)^2}{1.65}$$
$$\sigma^2 \le 343.7$$

With 99% confidence, we believe the true variance of compressive strength is at most 343.7(Mpa)^2.

e) $\chi^2_{\alpha/2,8} = 20.09 \quad \chi^2_{1-(\alpha/2),8} = 1.65$ with $\alpha = 0.02$

$$\frac{(n-1)s^2}{\chi^2_{\alpha/2,8}} \le \sigma^2 \le \frac{(n-1)s^2}{\chi^2_{1-(\alpha/2),8}}$$
$$\frac{8(8.42)^2}{20.09} \le \sigma^2 \le \frac{8(8.42)^2}{1.65}$$
$$28.23 \le \sigma^2 \le 343.74$$

With 98% confidence, we believe the true variance of compressive strength is between 28.23 and 343.74 (Mpa)^2. The upper endpoint of the one-sided confidence interval is the same as that of the two-sided confidence interval due to the level of confidence. In both cases the probability in the right tail is 0.01.

f) Change 40.2 to 20.4

$\bar{x} = 23$ and $s = 6.3$

98% confidence on μ:

$$23 - 2.896\left(\frac{6.3}{\sqrt{9}}\right) \le \mu \le 23 + 2.896\left(\frac{6.3}{\sqrt{9}}\right)$$
$$16.92 \le \mu \le 29.08$$

With 98% confidence, we believe the true mean compressive strength is between 16.92 and 29.08 Mpa.

98% confidence interval on σ^2:

$$\frac{8(6.3)^2}{20.09} \le \sigma^2 \le \frac{8(6.3)^2}{1.65}$$
$$15.80 \le \sigma^2 \le 192.44$$

With 98% confidence, we believe the true variance of compressive strength is between 15.80 and 192.44 (Mpa)^2.

Comparison of intervals:

Confidence interval on μ: The confidence interval now covers a region slightly less than the original interval, that is the length of the second interval is shorter than the original.
Confidence interval on σ^2 : The confidence interval covers a region slightly less than the original; the length of the second interval is shorter than the original.
Effects: We see by correcting the value, the length of the intervals have become shorter and the mean and variance have decreased. In particular, the sample standard deviation as decreased by about 25% causing the confidence intervals to decrease substantially.

g) Change 25.8 to 24.8
$\bar{x} = 25$ and s = 8.41
98% confidence on μ:

$$25 - 2.896\left(\frac{8.41}{\sqrt{9}}\right) \le \mu \le 25 + 2.896\left(\frac{8.41}{\sqrt{9}}\right)$$

$$16.88 \le \mu \le 33.12$$

With 98% confidence, we believe the true mean compressive strength is between 16.88 and 33.12.

98% confidence interval on σ^2 :

$$\frac{8(8.41)^2}{20.09} \le \sigma^2 \le \frac{8(8.41)^2}{1.65}$$

$$28.16 \le \sigma^2 \le 342.92$$

With 98% confidence, we believe the true variance of compressive strength is between 28.16 and 342.92 $(\text{Mpa})^2$.

Comparison of intervals parts f and g:

Confidence interval on μ: There is very little difference when the data value is changed only slightly.
Confidence interval on σ^2 : The confidence interval has changed very little with a slight change in one data value.
Effects: The sample mean, sample variance, and confidence intervals have changed very little.

h) There are two cases that the above exercises illustrate: 1) when a value is changed and the new value lies far from the sample mean; and 2) when a value is changed and the new value lies near the sample mean. When the situation is case 1, the variance has changed dramatically resulting in a smaller confidence interval. The mean is less sensitive to this change. When the situation is case 2, the mean and variance have changed less than in case 1. Notice the widths of the confidence intervals for the original data, case 1, and case 2 in the table below.

	Width of Confidence Interval		
Parameter of Interest	**Original Data**	**Case 1**	**Case 2**
Mean, μ	16.25	12.16	16.24
Variance, σ^2	315.51	176.64	314.76

8-75. With σ = 8, the 95% confidence interval on the mean has length of at most 5; the error is then E = 2.5.

a) $n = \left(\frac{z_{0.025}}{2.5}\right)^2 8^2 = \left(\frac{1.96}{2.5}\right)^2 64 = 39.34 = 40$

b) $n = \left(\frac{z_{0.025}}{2.5}\right)^2 6^2 = \left(\frac{1.96}{2.5}\right)^2 36 = 22.13 = 23$

As the standard deviation decreases, with all other values held constant, the sample size necessary to maintain the acceptable level of significance and the length of the interval, decreases.

c) We would want to have a relatively large sample size, n ≥ 30. With a sample size of at least 30, the central limit theorem can apply.

8-76. Sample Mean = $\hat{p}$ Sample Variance = $\frac{\hat{p}(1-\hat{p})}{n}$

	Sample Size, n	Sampling Distribution	Sample Mean	Sample Variance
a.	50	Normal	p	$\frac{p(1-p)}{50}$
b.	80	Normal	p	$\frac{p(1-p)}{80}$
c.	100	Normal	p	$\frac{p(1-p)}{100}$

d) As the sample size increases, the variance of the sampling distribution decreases.

8-77. $z_{\alpha/2} = 1.96$ $E = z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

a) $n = 50$ $\hat{p} = 0.1$

$$E = 1.96\sqrt{\frac{0.1(1-0.1)}{50}} = 0.083$$

b) $n = 80$ $\hat{p} = 0.1$

$$E = 1.96\sqrt{\frac{0.1(1-0.1)}{80}} = 0.066$$

c) $n = 100$ $\hat{p} = 0.1$

$$E = 1.96\sqrt{\frac{0.1(1-0.1)}{100}} = 0.059$$

d) As the sample size increases and all other values held constant, the error decreases.

e) $z_{\alpha/2} = 2.575$ $E = z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

n = 50

$$E = 2.575\sqrt{\frac{0.1(1-0.1)}{50}} = 0.109$$

n = 80

$$E = 2.575\sqrt{\frac{0.1(1-0.1)}{80}} = 0.086$$

n = 100

$$E = 2.575\sqrt{\frac{0.1(1-0.1)}{100}} = 0.077$$

As the sample size increase and all other values held constant, the error decreases.

f) When the confidence level is increased, the error in estimating the true value of p will also increase. This can be seen by comparing values between parts d and e. For n = 50 we see that E = 0.083 with a 95% level of confidence, while E = 0.109 with a 99% level of confidence for the same sample size.

8-78. $\sigma = 12$, $\delta = 205 - 200 = 5$, $\frac{\alpha}{2} = 0.025$, $z_{0.025} = 1.96$,

a) n = 20: $\beta = \Phi\left(1.96 - \frac{5\sqrt{20}}{12}\right) = \Phi(0.163) = 0.564$

b) n = 50: $\beta = \Phi\left(1.96 - \frac{5\sqrt{50}}{12}\right) = \Phi(-0.986) = 1 - \Phi(0.986) = 1 - 0.839 = 0.161$

c) n = 100: $\beta = \Phi\left(1.96 - \frac{5\sqrt{100}}{12}\right) = \Phi(-2.207) = 1 - \Phi(2.207) = 1 - 0.9884 = 0.116$

d) β, which is the probability of a Type II error, decreases as the sample size increases because the variance of the sample mean decreases. Consequently, the probability of observing a sample mean in the acceptance region centered about the incorrect value of 200 ml/h decreases with larger n.

8-79. $\sigma = 14,\ \delta = 205 - 200 = 5,\ \frac{\alpha}{2} = 0.025,\ z_{0.025} = 1.96,$

a) n = 20: $\beta = \Phi\left(1.96 - \frac{5\sqrt{20}}{14}\right) = \Phi(0.36) = 0.64058$

b) n = 50: $\beta = \Phi\left(1.96 - \frac{5\sqrt{50}}{14}\right) = \Phi(-0.565) = 1 - \Phi(0.57) = 1 - 0.71566 = 0.28434$

c) n = 100: $\beta = \Phi\left(1.96 - \frac{5\sqrt{100}}{14}\right) = \Phi(-1.61) = 1 - \Phi(1.61) = 1 - 0.94630 = 0.0537$

d) The probability of a Type II error increases with an increase of the standard deviation.

8-80. $\sigma = 8,\ \delta = 204 - 200 = -4,\ \frac{\alpha}{2} = 0.025,\ z_{0.025} = 1.96.$

a) n = 20: $\beta = \Phi\left(1.96 - \frac{4\sqrt{20}}{8}\right) = \Phi(-0.28) = 1 - \Phi(0.28) = 1 - 0.61026 = 0.38974$

Therefore, power = $1 - \beta = 0.61026$

b) n = 50: $\beta = \Phi\left(1.96 - \frac{4\sqrt{50}}{8}\right) = \Phi(-2.58) = 1 - \Phi(2.58) = 1 - 0.99506 = 0.00494$

Therefore, power = $1 - \beta = 0.995$

c) n = 100: $\beta = \Phi\left(1.96 - \frac{4\sqrt{100}}{8}\right) = \Phi(-3.04) = 1 - \Phi(3.04) = 1 - 0.99882 = 0.00118$

Therefore, power = $1 - \beta = 0.9988$

d) As sample size increases, and all other values are held constant, the power increases because the variance of the sample mean decreases. Consequently, the probability of a Type II error decreases, which implies the power increases.

8-81. $H_0 : \mu = 85$ $\sigma = 16$ the true mean is 86

a) $\beta = P(\bar{x} \le 85 \text{ when } \mu = 86) = P\left(\frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \le \frac{85 - 86}{16/\sqrt{n}}\right) = P(z \le -0.0625\sqrt{n})$

n = 25: $\beta = P(z \le -0.0625\sqrt{25}) = 0.3783$

n = 100: $\beta = P(z \le -0.0625\sqrt{100}) = 0.2643$

n = 400: $\beta = P(z \le -0.0625\sqrt{400}) = 0.1056$

n = 2500: $\beta = P(z \le -0.0625\sqrt{2500}) = 0.0009$

b) P-value = $P(\bar{x} \ge 86) = P\left(\frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \ge \frac{86 - 85}{16/\sqrt{n}}\right) = P(z \ge 0.0625\sqrt{n})$

n = 25: P-value = $P(z \ge 0.0625\sqrt{25}) = 0.3783$

n = 100: P-value = $P(z \ge 0.0625\sqrt{100}) = 0.2643$

n = 400: P-value = $P(z \ge 0.0625\sqrt{400}) = 0.1056$

n = 2500: P-value = $P(z \ge 0.0625\sqrt{2500}) = 0.0009$; only sample that is statistically significant at $\alpha = 0.01$.

c) As the sample size increases, the probability of committing an error, Type I or Type II, decreases.

8-82. a) Rejecting a null hypothesis provides *stronger evidence* than not rejecting a null hypothesis. Therefore, place what we are trying to prove in the alternative hypothesis.

Assume the data follow a normal distribution.

b) 1) the parameter of interest is the mean weld strength, μ.
2) $H_0 : \mu = 150$
3) $H_1 : \mu > 150$
4) Not given
5) The test statistic is:

$$t_0 = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$$

6) Since no critical value is given, we will calculate the P-value
7) $\bar{x} = 153.7$

$$t_0 = 1.46$$

P-value = $P(t \geq 1.46) = 0.05 < p - \text{value} < 0.10$

8) There is somewhat significant evidence to support the claim that the weld strength exceeds 150 psi. If we used $\alpha = 0.01$ or 0.05, we would not reject the null hypothesis, thus the claim would not be supported. If we used $\alpha = 0.10$, we would reject the null in favor of the alternative and conclude the weld strength exceeds 150 psi.

8-83. $H_0 : p = 0.5$ versus $H_1 : p \neq 0.5$

a) Find the Power, $1-\beta$ for 95% confidence

$$\beta = \Phi\left(\frac{p_0 - p + z_{\alpha/2}\sqrt{\frac{p_0(1-p_0)}{n}}}{\sqrt{\frac{p(1-p)}{n}}}\right) - \Phi\left(\frac{p_0 - p - z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}}{\sqrt{\frac{p(1-p)}{n}}}\right)$$

$$= \Phi\left(\frac{0.5 - 0.6 + 1.96\sqrt{\frac{0.5(1-0.5)}{n}}}{\sqrt{\frac{0.6(1-0.6)}{n}}}\right) - \Phi\left(\frac{0.5 - 0.6 - 1.96\sqrt{\frac{0.6(1-0.6)}{n}}}{\sqrt{\frac{0.6(1-0.6)}{n}}}\right)$$

n = 100: $\beta = \Phi(-0.041) - \Phi(-4.04) = 0.48405 - 0 = 0.48405$
$1-\beta = 1 - 0.48405 = 0.51595$

n = 150: $\beta = \Phi(-0.50) - \Phi(-4.5) = (1 - 0.69146) - 0 = 0.30856$
$1-\beta = 1 - 0.30856 = 0.69146$

n = 300: $\beta = \Phi(-1.54) - \Phi(-5.5) = (1 - 0.93822) - 0 = 0.06178$
$1-\beta = 1 - 0.06178 = 0.93822$

Power increases as sample size increases, when all other values are held constant.

b) Find the Power, $1-\beta$ for 99% confidence

n = 100: $\beta = \Phi(0.59) - \Phi(-4.67) = 0.7224 - 0 = 0.7224$
$1-\beta = 1 - 0.7224 = 0.2776$

n = 150: $\beta = \Phi(0.13) - \Phi(-5.12) = 0.55172 - 0 = 0.55172$
$1-\beta = 1 - 0.55172 = 0.44828$

$n = 300$: $\beta = \Phi(-0.91) - \Phi(-6.16) = 1 - 0.81859 - 0 = 0.18141$

$1 - \beta = 1 - 0.18141 = 0.81859$

The power of the test is greater for the larger values of α, since the larger α value results in a smaller acceptance region.

c) $n = 100$ $p = 0.8$ $\alpha = 0.05$

$$\beta = \Phi\left(\frac{p_0 - p + z_{\alpha/2}\sqrt{\frac{p_0(1-p_0)}{n}}}{\sqrt{\frac{p(1-p)}{n}}}\right) - \Phi\left(\frac{p_0 - p - z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}}{\sqrt{\frac{p(1-p)}{n}}}\right)$$

$$= \Phi\left(\frac{.05 - 0.8 + 1.96\sqrt{\frac{0.5(1-0.5)}{n}}}{\sqrt{\frac{0.8(1-0.8)}{n}}}\right) - \Phi\left(\frac{0.5 - 0.8 - 1.96\sqrt{\frac{0.8(1-0.8)}{n}}}{\sqrt{\frac{0.8(1-0.8)}{n}}}\right)$$

$$= \Phi(-5.05) - \Phi(-9.95) = 0$$

Power = 1 – 0 = 1.

As the difference between the hypothesized value of p and the true value of p increases, the more powerful the test.

d) $\alpha = 0.01$, $\beta = 0.05$, $z_{\alpha/2} = 2.575$, $z_\beta = 1.645$

$p = 0.6$:

$$n = \left(\frac{2.575(\sqrt{0.5(0.5)}) + 1.645(\sqrt{0.6(0.4)})}{0.5 - 0.6}\right)^2 = 438.2$$

$n = 439$

$p = 0.8$:

$$n = \left(\frac{2.575(\sqrt{0.5(0.5)}) + 1.645(\sqrt{0.8(0.2)})}{0.5 - 0.8}\right)^2 = 48.69$$

$n = 49$

A smaller sample size is required when the true proportion lies further from the hypothesized value, since not as large of a sample will be necessary to detect this difference.

8-84. a) 1) the parameter of interest is the standard deviation, σ

2) $H_0 : \sigma^2 = 400$

3) $H_1 : \sigma^2 < 400$

4) Not given

5) The test statistic is: $\chi_0^2 = \frac{(n-1)s^2}{\sigma^2}$

6) Since no critical value is given, we will calculate the p-value

7) $n = 10$, $s = 15.7$

$$\chi_0^2 = \frac{9(15.7)^2}{400} = 5.546$$

P-value = $P(\chi^2 < 5.546)$; $\quad 0.1 < P - \text{value} < 0.5$

8) The P-value is greater than any acceptable significance level, α, therefore we do not reject the null hypothesis. There is insufficient evidence to support the claim that the standard deviation is less than 20 microamps.

b) 7) n = 51, s = 20

$$\chi_0^2 = \frac{50(15.7)^2}{400} = 30.81$$

P-value = $P\left(\chi^2 < 30.81\right)$; $\quad 0.01 < P - \text{value} < 0.025$

8) The P-value is less than 0.05, therefore we reject the null hypothesis and conclude that the standard deviation is significantly less than 20 microamps.

c) Increasing the sample size increases the test statistic χ_0^2 and therefore decreases the P-value, providing more evidence against the null hypothesis.

8-85. n = 6

a) 1) The parameter of interest is the standard deviation, σ.
2) $H_0: \sigma^2 = 1.0$
3) $H_1: \sigma^2 \neq 1.0$
4) $\alpha = 0.01$
5) The test statistic is: $\chi_0^2 = \dfrac{(n-1)s^2}{\sigma^2}$
6) Reject H_0 if $\chi_0^2 < \chi^2_{1-\alpha/2,n-1}$ where $\chi^2_{0.995,5} = 0.41$ or $\chi_0^2 > \chi^2_{\alpha/2,n-1}$ where $\chi^2_{0.005,5} = 16.75$
7) n = 6, s = 0.319

$$\chi_0^2 = \frac{5(0.319)^2}{1.0} = 0.5088$$

8) Since 0.41 < 0.5088 < 16.75, accept the null hypothesis and conclude the true variance of fatty acid for diet margarine is not significantly different from 1.0 at α = 0.05.

b) n = 51
2) $H_0: \sigma^2 = 1.0$
3) $H_1: \sigma^2 \neq 1.0$
4) $\alpha = 0.01$
5) The test statistic is: $\chi_0^2 = \dfrac{(n-1)s^2}{\sigma^2}$
6) Reject H_0 if $\chi_0^2 < \chi^2_{1-\alpha/2,n-1}$ where $\chi^2_{0.995,50} = 27.99$ or $\chi_0^2 > \chi^2_{\alpha/2,n-1}$ where $\chi^2_{0.005,50} = 79.49$
7) n = 51, s = 0.319

$$\chi_0^2 = \frac{50(0.319)^2}{1.0} = 5.09$$

8) Since 5.09 < 27.99, reject the null hypothesis and conclude the true variance of fatty acid for diet margarine is significantly different from 1.0 at α = 0.05.

c) The increased sample size changes the degrees of freedom of the test statistic and therefore the acceptance region; the larger sample size actually decreases the acceptable difference between the sample variance and actual.

8-86. Assume the data follow a normal distribution.

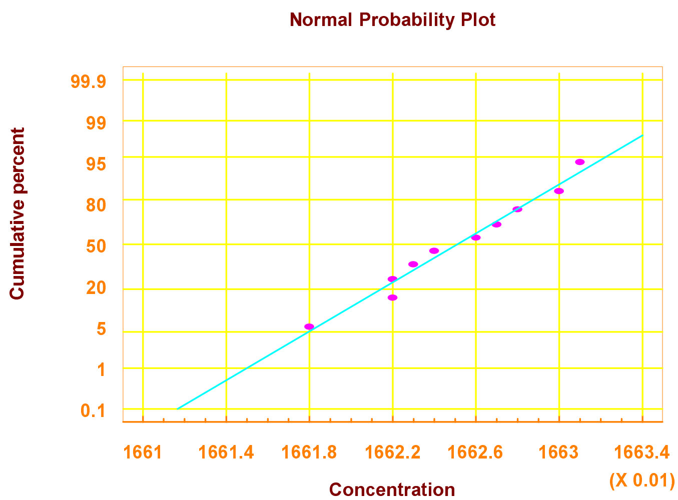
a) 1) The parameter of interest is the standard deviation, σ.
2) $H_0 : \sigma^2 = (0.00002)^2$
3) $H_1 : \sigma^2 < (0.00002)^2$
4) $\alpha = 0.01$
5) The test statistic is: $\chi_0^2 = \frac{(n-1)s^2}{\sigma^2}$
6) $\chi^2_{0.99,7} = 1.24$ reject H_0 if $\chi_0^2 < 1.24$
7) $s = 0.00001$ and $\alpha = 0.01$

$$\chi_0^2 = \frac{7(0.00001)^2}{(0.00002)^2} = 1.75$$

1.75 > 1.24, do not reject the null hypothesis; that is, there is insufficient evidence to conclude the standard deviation is at most 0.00002 mm.

b) Although the sample standard deviation is less than the hypothesized value of 0.00002, it is *not significantly less* (when $\alpha = 0.01$) than 0.00002 to conclude the standard deviation is at most 0.00002 mm. The value of 0.00001 could have occurred as a result of sampling variation.

8-87.



According to the normal probability plot, we can assume the underlying distribution is normal. This is evident by the fact that the data fall along a straight line. The assumption of normality should be satisfied in order to perform a hypothesis test using a χ^2 test statistic.

1) The parameter of interest is the standard deviation, σ.
2) $H_0 : \sigma^2 = 16$
3) $H_1 : \sigma^2 < 16$
4) Not given
5) The test statistic is: $\chi_0^2 = \frac{(n-1)s^2}{\sigma^2}$
6) Since no critical value is given, we will calculate the p-value
7) $n = 10$

$s^2 = 0.004$

$$\chi_0^2 = \frac{9(0.004)^2}{16} = 0.000009$$

P-value = $P(\chi^2 \le 0.000009)$; P-value < 0.005

8) A P-value of less than 0.005 is highly significant evidence to conclude the standard deviation is less than 4 g/l.

8-88. a) 1) The parameter of interest is the true proportion, p.

2) $H_0 : p = 0.01$

3) $H_1 : p < 0.01$

4) $\alpha = 0.01$

5) The test statistic is: $z_0 = \dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1 - p_0)}{n}}}$

6) $-z_{0.01} = -2.33$, reject H_0 if $z_0 < -2.33$

7) $\hat{p} = \dfrac{8}{1200} = 0.0067$

$$z_0 = \frac{0.0067 - 0.01}{\sqrt{\dfrac{0.01(1 - 0.01)}{1200}}} = -1.15$$

8) $-1.15 > -2.33$, do not reject the null hypothesis and conclude that there is insufficient evidence to support the claim that the true proportion is less than 1%.

$$\text{P-value} = P(z \le -1.15) = 1 - 0.87493 = 0.12507$$

b) Although the sample proportion is less than the hypothesized value of 0.01, it is *not significantly less* than 0.01 to conclude that the proportion is at less than 0.01. The value of 0.0067 could have occurred as a result of sampling variation.

8-89. $\hat{p} = \dfrac{8}{1600} = 0.005$

a) 99% confidence interval: $z_{\alpha/2} = 2.575$

$$\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} \le p \le \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}(1 - \hat{p})}{n}}$$

$0.00045 \le p \le 0.0095$

With 99% confidence, we believe the true proportion of aircraft that have wiring errors lies between 0.00045 and 0.0095.

b) $n = \left(\dfrac{z_{\alpha/2}}{E}\right)^2 \hat{p}(1 - \hat{p}) = \left(\dfrac{2.575}{0.008}\right)^2 (0.005)(1 - 0.005) = 515.427$

$n = 516$

c) Use $\hat{p} = 0.5$

$n = \left(\dfrac{z_{\alpha/2}}{E}\right)^2 \hat{p}(1 - \hat{p}) = \left(\dfrac{2.575}{0.008}\right)^2 (0.5)(1 - 0.5) = 25{,}900$

$n = 25{,}900$

d) Preliminary information can significantly decrease the computed needed sample size. Without this information, we must use p = 0.5 in the computations resulting the worst case size.

8-90. $\hat{p} = \dfrac{117}{484} = 0.242$

a) 90% confidence interval; $z_{\alpha/2} = 1.645$

$$\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} \le p \le \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}(1 - \hat{p})}{n}}$$

$0.210 \le p \le 0.274$

With 90% confidence, we believe the true proportion of new engineering graduates who were planning to continue studying for an advanced degree lies between 0.210 and 0.274.

b) 95% confidence interval; $z_{\alpha/2} = 1.96$

$$\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \le p \le \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$0.204 \le p \le 0.280$

With 95% confidence, we believe the true proportion of new engineering graduates who were planning to continue studying for an advanced degree lies between 0.204 and 0.280.

c) Comparison of parts a and b:

The 95% confidence interval is larger than the 90% confidence interval. Higher confidence always yields larger intervals, all other values held constant.

d) Yes, since both intervals contain the value 0.25, thus the true proportion cannot be considered significantly different from 0.25.

8-91. Create a table for the number of nonconforming coil springs (value) and the observed number of times the number appeared. One possible table is:

Value	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Obs	0	0	0	1	4	3	4	6	4	3	0	3	3	2	1	1	0	2	1	2

The value of p must be estimated. Let the estimate be denoted by $\hat{p}_{sample}$

$$\text{sample mean} = \frac{0(0)+1(0)+2(0)+\cdots+19(2)}{40} = 9.325$$

$$\hat{p}_{sample} = \frac{\text{sample mean}}{n} = \frac{9.325}{50} = 0.1865$$

Value	Observed	Expected
0	0	0.00165
1	0	0.01889
2	0	0.10608
3	1	0.38911
4	4	1.04816
5	3	2.21073
6	4	3.80118
7	6	5.47765
8	4	6.74985
9	3	7.22141
10	0	6.78777
11	3	5.65869
12	3	4.21619
13	2	2.82541
14	1	1.71190
15	1	0.94191
16	0	0.47237
17	2	0.21659
18	1	0.09103
19	2	0.03515

Since several of the expected values are less than 3, some cells must be combined resulting in the following table:

Value	Observed	Expected
0-5	8	3.77462
6	4	3.80118
7	6	5.47765
8	4	6.74985
9	3	7.22141
10	0	6.78777
11	3	5.65869
12	3	4.21619
≥13	9	6.29436

The degrees of freedom are k – p – 1 = 9 – 1 – 1 = 7

a) 1) The variable of interest is the form of the distribution for the number of nonconforming coil springs.
2) H_0: The form of the distribution is binomial
3) H_1: The form of the distribution is not binomial
4) $\alpha = 0.05$
5) The test statistic is

$$\chi_0^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$$

6) Reject H_0 if $\chi_0^2 > \chi_{0.05,7}^2 = 14.07$
7)

$$\chi_0^2 = \frac{(8 - 3.77462)^2}{3.77462} + \frac{(4 - 3.8.011)^2}{3.8011} + \cdots + \frac{(9 - 6.29436)^2}{6.29436} = 17.929$$

8) Since 17.929 > 14.07 reject H_0. We are able to conclude the distribution of nonconforming springs is not binomial at $\alpha = 0.05$.

b) P-value = 0.0123 (found using Minitab)

8-92. Create a table for the number of errors in a string of 1000 bits (value) and the observed number of times the number appeared. One possible table is:

Value	0	1	2	3	4	5
Obs	3	7	4	5	1	0

The value of p must be estimated. Let the estimate be denoted by $\hat{p}_{sample}$

$$\text{sample mean} = \frac{0(3) + 1(7) + 2(4) + 3(5) + 4(1) + 5(0)}{20} = 1.7$$

$$\hat{p}_{sample} = \frac{\text{sample mean}}{n} = \frac{1.7}{1000} = 0.0017$$

Value	0	1	2	3	4	5
Observed	3	7	4	5	1	0
Expected	3.64839	6.21282	5.28460	2.99371	1.27067	0.43103

Since several of the expected values are less than 3, some cells must be combined resulting in the following table:

Value	0	1	2	≥3
Observed	3	7	4	6
Expected	3.64839	6.21282	5.28460	4.69541

The degrees of freedom are k – p – 1 = 4 – 1 – 1 = 2

a) 1) The variable of interest is the form of the distribution for the number of errors in a string of 1000 bits.
2) H_0: The form of the distribution is binomial
3) H_1: The form of the distribution is not binomial
4) $\alpha = 0.05$
5) The test statistic is

$$\chi_0^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$$

6) Reject H_0 if $\chi_0^2 > \chi_{0.05,2}^2 = 5.99$
7)

$$\chi_0^2 = \frac{(3-3.64839)^2}{3.64839} + \cdots + \frac{(6-4.69541)^2}{4.69541} = 0.88971$$

8) Since 0.88971 < 9.49 do not reject H_0. We are unable to reject the null hypothesis that the distribution of the number of errors is binomial at $\alpha = 0.05$.

b) P-value = 0.6409 (found using Minitab)

Mind-Expanding Exercises

8-93. a)

$$\begin{aligned} T_r = {} & X_1 + \\ & X_1 + X_2 - X_1 + \\ & X_1 + X_2 - X_1 + X_3 - X_2 + \\ & \ldots + \\ & X_1 + X_2 - X_1 + X_3 - X_2 + \ldots + X_r - X_{r-1} + \\ & (n-r)(X_1 + X_2 - X_1 + X_3 - X_2 + \ldots + X_r - X_{r-1}) \end{aligned}$$

Because X_1 is the minimum lifetime of n items, $E(X_1) = \frac{1}{n\lambda}$.

Then, $X_2 - X_1$ is the minimum lifetime of (n-1) items from the memoryless property of the exponential and $E(X_2 - X_1) = \frac{1}{(n-1)\lambda}$.

Similarly, $E(X_k - X_{k-1}) = \frac{1}{(n-k+1)\lambda}$. Then,

$$E(T_r) = \frac{n}{n\lambda} + \frac{n-1}{(n-1)\lambda} + \ldots + \frac{n-r+1}{(n-r+1)\lambda} = \frac{r}{\lambda} \text{ and } E\left(\frac{T_r}{r}\right) = \frac{1}{\lambda} = \mu$$

b) $P(\chi^2_{1-\frac{\alpha}{2},2r} < 2\lambda T_r < \chi^2_{\frac{\alpha}{2},2r}) = 1-\alpha$

$$= P\left(\frac{\chi^2_{1-\frac{\alpha}{2},2r}}{2T_r} < \lambda < \frac{\chi^2_{\frac{\alpha}{2},2r}}{2T_r}\right)$$

Then a confidence interval for $\mu = \frac{1}{\lambda}$ is $\left(\frac{2T_r}{\chi^2_{\frac{\alpha}{2},2r}}, \frac{2T_r}{\chi^2_{1-\frac{\alpha}{2},2r}}\right)$

c) n = 20 , r = 10 , and the observed value of T_r is 199 + 10(29) = 489.

A 95% confidence interval for $\frac{1}{\lambda}$ is $\left(\frac{2(489)}{34.17}, \frac{2(489)}{9.59}\right) = (28.62, 101.98)$

8-94. $\alpha_1 = \int_{z_{\alpha_1}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 1 - \int_{-\infty}^{z_{\alpha_1}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$

Therefore, $1-\alpha_1 = \Phi(z_{\alpha_1})$.

To minimize L we need to minimize $\Phi^{-1}(1-\alpha_1) + \Phi(1-\alpha_2)$ subject to $\alpha_1 + \alpha_2 = \alpha$. Therefore, we need to minimize $\Phi^{-1}(1-\alpha_1) + \Phi(1-\alpha+\alpha_1)$.

$$\frac{\partial}{\partial \alpha_1} \Phi^{-1}(1-\alpha_1) = -\sqrt{2\pi}\ e^{\frac{z^2_{\alpha_1}}{2}}$$

$$\frac{\partial}{\partial \alpha_1} \Phi^{-1}(1-\alpha+\alpha_1) = \sqrt{2\pi}\ e^{\frac{z^2_{\alpha-\alpha_1}}{2}}$$

Upon setting the sum of the two derivatives equal to zero, we obtain $e^{\frac{z^2_{\alpha-\alpha_1}}{2}} = e^{\frac{z^2_{\alpha_1}}{2}}$. This is solved by $z_{\alpha_1} = z_{\alpha-\alpha_1}$. Consequently, $\alpha_1 = \alpha - \alpha_1$, $2\alpha_1 = \alpha$ and $\alpha_1 = \alpha_2 = \frac{\alpha}{2}$.

8-95. $E(X_{n+1}) = \overline{X}$, $E(X_{n+1} - \overline{X}) = E(X_{n+1}) - E(\overline{X}) = \overline{X} - \overline{X} = 0$

$$V(X_{n+1} - \overline{X}) = V(X_{n+1}) + V(\overline{X}) = V(X_{n+1}) + V\left(\frac{\sum X_i}{n}\right)$$

$$= \sigma^2 + \frac{1}{n^2} V(X_1 + X_2 + \ldots + X_n)$$

$$= \sigma^2 + \frac{1}{n^2}\left[V(X_1) + V(X_2) + \ldots + V(X_n)\right]$$

$$= \sigma^2 + \frac{1}{n^2}\left[\overbrace{\sigma^2 + \sigma^2 + \ldots + \sigma^2}^{n}\right]$$

$$= \sigma^2 + \frac{1}{n^2}\left[n\sigma^2\right]$$

$$= \sigma^2 + \frac{\sigma^2}{n}$$

$$= \sigma^2\left(1 + \frac{1}{n}\right)$$

The distribution of $\dfrac{X_{n+1} - \overline{X}}{S\sqrt{1 + \frac{1}{n}}}$ is a t-distribution with n - 1 degrees of freedom.

A 100(1−α) percent prediction interval on X_{n+1} can be developed as:

$$-t_{\alpha/2,n-1} \le \frac{X_{n+1} - \overline{X}}{S\sqrt{1 + \frac{1}{n}}} \le t_{\alpha/2,n-1}$$

$$-t_{\alpha/2,n-1} S\sqrt{1 + \frac{1}{n}} \le X_{n+1} - \overline{X} \le t_{\alpha/2,n-1} S\sqrt{1 + \frac{1}{n}}$$

$$\overline{X} - t_{\alpha/2,n-1} S\sqrt{1 + \frac{1}{n}} \le X_{n+1} \le \overline{X} + t_{\alpha/2,n-1} S\sqrt{1 + \frac{1}{n}}$$

95% prediction interval on X_{16} is

$$\overline{X} - t_{\alpha/2,n-1} S\sqrt{1 + \frac{1}{n}} \le X_{n+1} \le \overline{X} + t_{\alpha/2,n-1} S\sqrt{1 + \frac{1}{n}}$$

$$8.234 - 2.145(0.0253)\sqrt{1 + \frac{1}{15}} \le X_{16} \le 8.234 + 2.145(0.0253)\sqrt{1 + \frac{1}{15}}$$

$$8.18 \le X_{16} \le 8.29$$

The prediction interval is wider because it depends on both the error from the fitted model (error from using $\overline{X}$ as the estimate) and error associated with future observations.

CHAPTER 9

Section 9-2

9-1. a) 1) The parameter of interest is the difference in fill volume, $\mu_1 - \mu_2$

2) $H_0: \mu_1 - \mu_2 = 0$ or $\mu_1 = \mu_2$

3) $H_1: \mu_1 - \mu_2 \neq 0$ or $\mu_1 \neq \mu_2$

4) $\alpha = 0.05$

5) The test statistic is

$$z_0 = \frac{(\bar{x}_1 - \bar{x}_2) - \Delta_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

6) Reject H_0 if $z_0 < -z_{\alpha/2} = -1.96$ or $z_0 > z_{\alpha/2} = 1.96$

7) $\bar{x}_1 = 16.015$ $\bar{x}_2 = 16.005$ $\delta = 0$

$\sigma_1 = 0.02$ $\sigma_2 = 0.025$

$n_1 = 10$ $n_2 = 10$

$$z_0 = \frac{(16.015 - 16.005) - 0}{\sqrt{\frac{(0.02)^2}{10} + \frac{(0.025)^2}{10}}} = 0.99$$

8) since -1.96 < 0.99 < 1.96, do not reject the null hypothesis and conclude there is no evidence that the two machine fill volumes differ at $\alpha = 0.05$.

b) P-value $= 2(1 - \Phi(0.99)) = 2(1 - 0.8389) = 0.3222$

c) Power $= 1 - \beta$, where

$$\beta = \Phi\left(z_{\alpha/2} - \frac{\Delta - \Delta_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}\right) - \Phi\left(-z_{\alpha/2} - \frac{\Delta - \Delta_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}\right)$$

$$= \Phi\left(1.96 - \frac{0.08}{\sqrt{\frac{(0.02)^2}{10} + \frac{(0.025)^2}{10}}}\right) - \Phi\left(-1.96 - \frac{0.08}{\sqrt{\frac{(0.02)^2}{10} + \frac{(0.025)^2}{10}}}\right)$$

$$= \Phi(1.96 - 7.9) - \Phi(-1.96 - 7.9) = \Phi(-5.94) - \Phi(-9.86)$$

$$= 0 - 0$$

$$= 0$$

Power = 1 – 0 = 1

d) $$\left(\bar{x}_1 - \bar{x}_2\right) - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \le \mu_1 - \mu_2 \le \left(\bar{x}_1 - \bar{x}_2\right) + z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$$\left(16.015 - 16.005\right) - 1.96\sqrt{\frac{(0.02)^2}{10} + \frac{(0.025)^2}{10}} \le \mu_1 - \mu_2 \le \left(16.015 - 16.005\right) + 1.96\sqrt{\frac{(0.02)^2}{10} + \frac{(0.025)^2}{10}}$$

$$-0.0098 \le \mu_1 - \mu_2 \le 0.0298$$

With 95% confidence, we believe the true difference in the mean fill volumes is between –0.0098 and 0.0298. Since 0 is contained in this interval, we can conclude there is no significant difference between the means.

e) Assume the sample sizes are to be equal, use $\alpha = 0.05$, $\beta = 0.05$, and $\Delta = 0.08$

$$n \cong \frac{\left(z_{\alpha/2} + z_\beta\right)^2\left(\sigma_1^2 + \sigma_2^2\right)}{\delta^2} = \frac{\left(1.96 + 1.645\right)^2\left((0.02)^2 + (0.025)^2\right)}{(0.08)^2} = 2.08, \quad n = 3, \text{ use } n_1 = n_2 = 3$$

9-2. 1) The parameter of interest is the difference in breaking strengths, $\mu_1 - \mu_2$ and $\Delta_0 = 10$

2) $H_0 : \mu_1 - \mu_2 = 10$ or $\mu_1 = \mu_2$

3) $H_1 : \mu_1 - \mu_2 > 10$ or $\mu_1 > \mu_2$

4) $\alpha = 0.05$

5) The test statistic is

$$z_0 = \frac{(\overline{x}_1 - \overline{x}_2) - \Delta_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

6) Reject H_0 if $z_0 > z_\alpha = 1.645$

7) $\overline{x}_1 = 162.5$ $\overline{x}_2 = 155.0$ $\delta = 10$

$\sigma_1 = 1.0$ $\sigma_2 = 1.0$

$n_1 = 10$ $n_2 = 12$

$$z_0 = \frac{(162.5 - 155.0) - 10}{\sqrt{\frac{(1.0)^2}{10} + \frac{(1.0)^2}{12}}} = -5.84$$

8) Since -5.84 < 1.645 do not reject the null hypothesis and conclude there is insufficient evidence to support the use of plastic 1 at $\alpha = 0.05$.

9-3. $\beta = \Phi\left(1.645 - \frac{(12)}{\sqrt{\frac{1}{10} + \frac{1}{12}}}\right) = 0$, Power = 1 - 0 = 1.

9-4. a) 1) The parameter of interest is the difference in mean burning rate, $\mu_1 - \mu_2$

2) $H_0 : \mu_1 - \mu_2 = 0$ or $\mu_1 = \mu_2$

3) $H_1 : \mu_1 - \mu_2 \neq 0$ or $\mu_1 \neq \mu_2$

4) $\alpha = 0.05$

5) The test statistic is

$$z_0 = \frac{(\overline{x}_1 - \overline{x}_2) - \Delta_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

6) Reject H_0 if $z_0 < -z_{\alpha/2} = -1.96$ or $z_0 > z_{\alpha/2} = 1.96$

7) $\overline{x}_1 = 18$ $\overline{x}_2 = 24$ $\delta = 0$

$\sigma_1 = 3$ $\sigma_2 = 3$

$n_1 = 20$ $n_2 = 20$

$$z_0 = \frac{(18 - 24) - 0}{\sqrt{\frac{(3)^2}{20} + \frac{(3)^2}{20}}} = -6.32$$

8) Since −6.32 < −1.96 reject the null hypothesis and conclude the mean burning rates do not differ significantly at $\alpha = 0.05$.

b) P-value $= 2(1 - \Phi(6.32)) = 2(1 - 1) = 0$

c) $\beta = \Phi\left(z_{\alpha/2} - \frac{\Delta \; - \Delta_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \right) - \Phi\left(-z_{\alpha/2} - \frac{\Delta - \Delta_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \right)$

$= \Phi\left(1.96 - \frac{2.5}{\sqrt{\frac{(3)^2}{20} + \frac{(3)^2}{20}}} \right) - \Phi\left(-1.96 - \frac{2.5}{\sqrt{\frac{(3)^2}{20} + \frac{(3)^2}{20}}} \right)$

$= \Phi(1.96 - 2.64) - \Phi(-1.96 - 2.64) = \Phi(-0.68) - \Phi(-4.6)$

$= 0.24825 - 0$

$= 0.24825$

d) $\left(\overline{x}_1 - \overline{x}_2\right) - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \le \mu_1 - \mu_2 \le \left(\overline{x}_1 - \overline{x}_2\right) + z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$

$(18 - 24) - 1.96\sqrt{\frac{(3)^2}{20} + \frac{(3)^2}{20}} \le \mu_1 - \mu_2 \le (18 - 24) + 1.96\sqrt{\frac{(3)^2}{20} + \frac{(3)^2}{20}}$

$-7.86 \le \mu_1 - \mu_2 \le -4.14$

We are 95% confident that the mean burning rate for solid fuel propellant 2 exceeds that of propellant 1 by between 4.14 and 7.86 cm/s.

9-5. $\overline{x}_1 = 30.87 \quad \overline{x}_2 = 30.68$

$\sigma_1 = 0.10 \quad \sigma_2 = 0.15$

$n_1 = 12 \quad n_2 = 10$

a) 90% two-sided confidence interval:

$\left(\overline{x}_1 - \overline{x}_2\right) - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \le \mu_1 - \mu_2 \le \left(\overline{x}_1 - \overline{x}_2\right) + z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$

$(30.87 - 30.68) - 1.645\sqrt{\frac{(0.10)^2}{12} + \frac{(0.15)^2}{10}} \le \mu_1 - \mu_2 \le (30.87 - 30.68) + 1.645\sqrt{\frac{(0.10)^2}{12} + \frac{(0.15)^2}{10}}$

$0.0987 \le \mu_1 - \mu_2 \le 0.2813$

We are 90% confident that the mean fill volume for machine 1 exceeds that of machine 2 by between 0.0987 and 0.2813 fl. oz.

b) 95% two-sided confidence interval:

$\left(\overline{x}_1 - \overline{x}_2\right) - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \le \mu_1 - \mu_2 \le \left(\overline{x}_1 - \overline{x}_2\right) + z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$

$(30.87 - 30.68) - 1.96\sqrt{\frac{(0.10)^2}{12} + \frac{(0.15)^2}{10}} \le \mu_1 - \mu_2 \le (30.87 - 30.68) + 1.96\sqrt{\frac{(0.10)^2}{12} + \frac{(0.15)^2}{10}}$

$0.081 \le \mu_1 - \mu_2 \le 0.299$

We are 95% confident that the mean fill volume for machine 1 exceeds that of machine 2 by between 0.081 and 0.299 fl. oz.

Comparison of parts a and b:

As the level of confidence increases, the interval width also increases (with all other values held constant).

c) 95% upper-sided confidence interval:

$$\mu_1 - \mu_2 \le (\overline{x}_1 - \overline{x}_2) + z_\alpha \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$$\mu_1 - \mu_2 \le (30.87 - 30.68) + 1.645\sqrt{\frac{(0.10)^2}{12} + \frac{(0.15)^2}{10}}$$

$$\mu_1 - \mu_2 \le 0.2813$$

With 95% confidence, we believe the fill volume for machine 1 exceeds the fill volume of machine 2 by no more than 0.2813 fl. oz.

9-6. a) 1) The parameter of interest is the difference in mean fill volume, $\mu_1 - \mu_2$

2) $H_0: \mu_1 - \mu_2 = 0$ or $\mu_1 = \mu_2$

3) $H_1: \mu_1 - \mu_2 \ne 0$ or $\mu_1 \ne \mu_2$

4) $\alpha = 0.05$

5) The test statistic is

$$z_0 = \frac{(\overline{x}_1 - \overline{x}_2) - \Delta_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

6) Reject H_0 if $z_0 < -z_{\alpha/2} = -1.96$ or $z_0 > z_{\alpha/2} = 1.96$

7) $\overline{x}_1 = 30.87$ $\overline{x}_2 = 30.68$ $\Delta_0 = 0$

$\sigma_1 = 0.10$ $\sigma_2 = 0.15$

$n_1 = 12$ $n_2 = 10$

$$z_0 = \frac{(30.87 - 30.68) - 0}{\sqrt{\frac{(0.10)^2}{12} + \frac{(0.15)^2}{10}}} = 3.42$$

8) Since 3.42 > 1.96 reject the null hypothesis and conclude the mean fill volumes of machine 1 and machine 2 differ significantly at $\alpha = 0.05$.

b) P-value = $2(1 - \Phi(3.42)) = 2(1 - 0.99969) = 0.00062$

c) Assume the sample sizes are to be equal, use $\alpha = 0.05$, $\beta = 0.10$, and $\Delta = 0.20$

$$n \cong \frac{(z_{\alpha/2} + z_\beta)^2(\sigma_1^2 + \sigma_2^2)}{(\Delta - \Delta_0)^2} = \frac{(1.96 + 1.28)^2((0.10)^2 + (0.15)^2)}{(-0.20)^2} = 8.53, \quad n = 9, \text{ use } n_1 = n_2 = 9$$

9-7. $\overline{x}_1 = 89.6$ $\overline{x}_2 = 92.5$

$\sigma_1^2 = 1.5$ $\sigma_2^2 = 1.2$

$n_1 = 15$ $n_2 = 20$

a) 95% confidence interval:

$$(\overline{x}_1 - \overline{x}_2) - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \le \mu_1 - \mu_2 \le (\overline{x}_1 - \overline{x}_2) + z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$$(89.6 - 92.5) - 1.96\sqrt{\frac{1.5}{15} + \frac{1.2}{20}} \le \mu_1 - \mu_2 \le (89.6 - 92.5) + 1.96\sqrt{\frac{1.5}{15} + \frac{1.2}{20}}$$

$$-3.684 \le \mu_1 - \mu_2 \le -2.116$$

With 95% confidence, we believe the mean road octane number for formulation 2 exceeds that of formulation 1 by between 2.116 and 3.684.

b) 1) The parameter of interest is the difference in mean road octane number, $\mu_1 - \mu_2$ and $\Delta_0 = 0$

2) $H_0: \mu_1 - \mu_2 = 0$ or $\mu_1 = \mu_2$

3) $H_1: \mu_1 - \mu_2 < 0$ or $\mu_1 < \mu_2$

4) $\alpha = 0.05$

5) The test statistic is

$$z_0 = \frac{(\overline{x}_1 - \overline{x}_2) - \Delta_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

6) Reject H_0 if $z_0 < -z_\alpha = -1.645$

7) $\overline{x}_1 = 89.6 \quad \overline{x}_2 = 92.5$

$\sigma_1^2 = 1.5 \quad \sigma_2^2 = 1.2$

$n_1 = 15 \quad n_2 = 20$

$$z_0 = \frac{(89.6 - 92.5) - 0}{\sqrt{\frac{(1.5)^2}{15} + \frac{(1.2)^2}{20}}} = -7.254$$

8) Since $-7.25 < -1.645$ reject the null hypothesis and conclude the mean road octane number for formulation 2 exceeds that of formulation 1 using $\alpha = 0.05$.

c) P-value = $P(z \le -7.25) = 1 - P(z \le 7.25) = 1 - 1 = 0$

9-8. 99% level of confidence, E = 4, and $z_{0.005} = 2.575$

$$n \cong \left(\frac{z_{0.005}}{E}\right)^2 \left(\sigma_1^2 + \sigma_2^2\right) = \left(\frac{2.575}{4}\right)^2 (9 + 9) = 7.46, \text{ n = 8, use } n_1 = n_2 = 8$$

9-9. 95% level of confidence, E = 1, and $z_{0.025} = 1.96$

$$n \cong \left(\frac{z_{0.025}}{E}\right)^2 \left(\sigma_1^2 + \sigma_2^2\right) = \left(\frac{1.96}{1}\right)^2 (1.5 + 1.2) = 10.37, \text{ n = 11, use } n_1 = n_2 = 11$$

9-10.

Case 1: Before Process Change	Case 2: After Process Change
μ_1 = mean batch viscosity before change	μ_2 = mean batch viscosity after change
$\overline{x}_1 = 750.2$	$\overline{x}_2 = 756.88$
$\sigma_1 = 20$	$\sigma_2 = 20$
$n_1 = 15$	$n_2 = 8$

90% confidence on $\mu_1 - \mu_2$, the difference in mean batch viscosity before and after process change:

$$\left(\overline{x}_1 - \overline{x}_2\right) - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \le \mu_1 - \mu_2 \le \left(\overline{x}_1 - \overline{x}_2\right) + z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$$(750.2 - 756.88) - 1.645\sqrt{\frac{(20)^2}{15} + \frac{(20)^2}{8}} \le \mu_1 - \mu_2 \le \left(750.2 - 756/88\right) + 1.645\sqrt{\frac{(20)^2}{15} + \frac{(20)^2}{8}}$$

$$-21.08 \le \mu_1 - \mu_2 \le 7.72$$

We are 90% confident that the difference in mean batch viscosity before and after the process change lies within –21.08 and 7.72. Since 0 is contained in this interval we can conclude with 90% confidence that the mean batch viscosity was unaffected by the process change.

9-11. Catalyst 1 / Catalyst 2

Catalyst 1	Catalyst 2
$\overline{x}_1 = 65.22$	$\overline{x}_2 = 68.42$
$\sigma_1 = 3$	$\sigma_2 = 3$
$n_1 = 10$	$n_2 = 10$

a) 95% confidence interval on $\mu_1 - \mu_2$, the difference in mean active concentration

$$\left(\overline{x}_1 - \overline{x}_2\right) - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \le \mu_1 - \mu_2 \le \left(\overline{x}_1 - \overline{x}_2\right) + z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$$(65.22 - 68.42) - 1.96\sqrt{\frac{(3)^2}{10} + \frac{(3)^2}{10}} \le \mu_1 - \mu_2 \le \left(65.22 - 68.42\right) + 1.96\sqrt{\frac{(3)^2}{10} + \frac{(3)^2}{10}}$$

$$-5.83 \le \mu_1 - \mu_2 \le -0.57$$

We are 95% confident that the mean active concentration of catalyst 2 exceeds that of catalyst 1 by between 0.57 and 5.83 g/l.

b) Yes, since the 95% confidence interval did not contain the value 0, we would conclude the mean active concentration depends on the choice of catalyst.

9-12. a) 1) The parameter of interest is the difference in mean batch viscosity before and after the process change, $\mu_1 - \mu_2$

2) $H_0: \mu_1 - \mu_2 = 10$

3) $H_1: \mu_1 - \mu_2 < 10$

4) $\alpha = 0.10$

5) The test statistic is

$$z_0 = \frac{(\overline{x}_1 - \overline{x}_2) - \Delta_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

6) Reject H_0 if $z_0 < -z_\alpha$ where $z_{0.1} = -1.28$

7) $\overline{x}_1 = 750.2$ $\quad \overline{x}_2 = 756.88$ $\quad \Delta_0 = 10$

$\sigma_1 = 20$ $\quad \sigma_2 = 20$

$n_1 = 15$ $\quad n_2 = 8$

$$z_0 = \frac{(750.2 - 756.88) - 10}{\sqrt{\frac{(20)^2}{15} + \frac{(20)^2}{8}}} = -1.90$$

8) Since $-1.90 < -1.28$ reject the null hypothesis and conclude the process change has increased the mean by less than 10.

b) P-value = $P(z \le -1.90) = 1 - P(z \le 1.90) = 1 - 0.97128 = 0.02872$

c) Parts a and b above give evidence that the mean batch viscosity change is less than 10. This conclusion is also seen by the confidence interval given in a previous problem since the interval did not contain the value 10. Since the upper endpoint is 7.72, then this also gives evidence that the difference is less than 10.

9-13. 1) The parameter of interest is the difference in mean active concentration, $\mu_1 - \mu_2$

2) $H_0: \mu_1 - \mu_2 = 0$ or $\mu_1 = \mu_2$

3) $H_1: \mu_1 - \mu_2 \neq 0$ or $\mu_1 \neq \mu_2$

4) $\alpha = 0.05$

5) The test statistic is

$$z_0 = \frac{(\overline{x}_1 - \overline{x}_2) - \Delta_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

6) Reject H_0 if $z_0 < -z_{\alpha/2} = -1.96$ or $z_0 > z_{\alpha/2} = 1.96$

7) $\overline{x}_1 = 750.2$ $\overline{x}_2 = 756.88$ $\delta = 0$

$\sigma_1 = 20$ $\sigma_2 = 20$

$n_1 = 15$ $n_2 = 8$

$$z_0 = \frac{(750.2 - 756.88) - 0}{\sqrt{\frac{(20)^2}{15} + \frac{(20)^2}{8}}} = -2.385$$

8) Since $-2.385 < -1.96$ reject the null hypothesis and conclude the mean active concentrations do differ significantly at $\alpha = 0.05$.

P-value $= 2(1 - \Phi(2.385)) = 2(1 - 0.99146) = 0.0171$

The conclusions reached by the confidence interval of the previous problem and the test of hypothesis conducted here are the same. A two-sided confidence interval can be thought of as representing the "acceptance region" of a hypothesis test, given that the level of significance is the same for both procedures. Thus if the value δ falls outside the confidence interval, it is the same result as rejecting the null hypothesis.

9-14.

$$\beta = \Phi\left(1.96 - \frac{(5)}{\sqrt{\frac{3^2}{10} + \frac{3^2}{10}}}\right) - \Phi\left(-1.96 - \frac{(5)}{\sqrt{\frac{3^2}{10} + \frac{3^2}{10}}}\right)$$

$$= \Phi(0.68) - \Phi(-3.30) = 0.75175 - 0.00048$$

$$= 0.75127$$

Power $= 1 - \beta = 1 - 0.75127 = 0.24873$. it would appear that the sample sizes will have to be increased to detect the difference of 5. Calculate the value of n using α and β.

Section 9-3

9-15. a) 1) The parameter of interest is the difference in mean rod diameter, $\mu_1 - \mu_2$

2) $H_0: \mu_1 - \mu_2 = 0$ or $\mu_1 = \mu_2$

3) $H_1: \mu_1 - \mu_2 \neq 0$ or $\mu_1 \neq \mu_2$

4) $\alpha = 0.05$

5) The test statistic is

$$t_0 = \frac{(\overline{x}_1 - \overline{x}_2) - \Delta_0}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

6) Reject the null hypothesis if $t_0 < -t_{\alpha/2, n_1+n_2-2}$ where $-t_{0.025,30} = -2.042$ or $t_0 > t_{\alpha/2, n_1+n_2-2}$ where $t_{0.025,30} = 2.042$

7)) $\bar{x}_1 = 8.73$ $\quad \bar{x}_2 = 8.68$ $\quad \Delta_0 = 0$

$s_1^2 = 0.35$ $\quad s_2^2 = 0.40$

$n_1 = 15$ $\quad n_2 = 17$

$$s_p = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}} = \sqrt{\frac{14(0.35)+16(0.40)}{30}} = 0.614$$

$$t_0 = \frac{(8.73-8.68)-0}{0.614\sqrt{\frac{1}{15}+\frac{1}{17}}} = 0.230$$

8) Since −2.042 < 0.230 < 2.042, do not reject the null hypothesis and conclude the two machines do not produce rods with significantly different mean diameters at $\alpha = 0.05$.

b) P-value = $2P(t > 0.230) > 2(0.40)$, P-value > 0.80

c) 95% confidence interval: $t_{0.025,30} = 2.042$

$$(\bar{x}_1 - \bar{x}_2) - t_{\alpha/2,n_1+n_2-2}(s_p)\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} \le \mu_1 - \mu_2 \le (\bar{x}_1 - \bar{x}_2) + t_{\alpha/2,n_1+n_2-2}(s_p)\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$

$$(8.73-8.68) - 2.042(0.614)\sqrt{\frac{1}{15}+\frac{1}{17}} \le \mu_1 - \mu_2 \le (8.73-8.68) + 2.042(0.643)\sqrt{\frac{1}{15}+\frac{1}{17}}$$

$$-0.415 \le \mu_1 - \mu_2 \le 0.515$$

Since zero is contained in this interval, we are 95% confident that machine 1 and machine 2 do not produce rods whose diameters are significantly different.

9-16. Assume the populations follow normal distributions and $\sigma_1^2 = \sigma_2^2$. The assumption of equal variances may be permitted in this case since it is known that the t-test and confidence intervals involving the t-distribution are robust to this assumption of equal variances when sample sizes are equal.

Case 1: AFCC	Case 2: ATC
μ_1 = mean foam expansion for AFCC	μ_2 = mean foam expansion for ATC
$\bar{x}_1 = 4.7$	$\bar{x}_2 = 6.9$
$s_1 = 0.6$	$s_2 = 0.8$
$n_1 = 5$	$n_2 = 5$

95% confidence interval: $t_{0.025,8} = 2.306$

$$s_p = \sqrt{\frac{5(0.60)+5(0.80)}{8}} = 0.7071$$

$$(\bar{x}_1 - \bar{x}_2) - t_{\alpha/2,n_1+n_2-2}(s_p)\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} \le \mu_1 - \mu_2 \le (\bar{x}_1 - \bar{x}_2) + t_{\alpha/2,n_1+n_2-2}(s_p)\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$

$$(4.7-6.9) - 2.306\sqrt{\frac{1}{5}+\frac{1}{5}} \le \mu_1 - \mu_2 \le (4.7-6.9) + 2.306\sqrt{\frac{1}{5}+\frac{1}{5}}$$

$$-3.23 \le \mu_1 - \mu_2 \le -1.17$$

Yes, with 95% confidence, we believe the mean foam expansion for ATC exceeds that of AFCC by between 1.17 and 2.32.

9-17. a) 1) The parameter of interest is the difference in mean catalyst yield, $\mu_1 - \mu_2$

2) $H_0: \mu_1 - \mu_2 = 0$ or $\mu_1 = \mu_2$

3) $H_1: \mu_1 - \mu_2 < 0$ or $\mu_1 < \mu_2$

4) $\alpha = 0.01$

5) The test statistic is

$$t_0 = \frac{(\bar{x}_1 - \bar{x}_2) - \Delta_0}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

6) Reject the null hypothesis if $t_0 < -t_{\alpha, n_1+n_2-2}$ where $-t_{0.005,25} = -2.787$

7) $\bar{x}_1 = 86 \quad \bar{x}_2 = 89 \quad \Delta_0 = 0$

$s_1 = 3 \quad s_2 = 2$

$n_1 = 12 \quad n_2 = 15$

$$s_p = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}} = \sqrt{\frac{12(3)^2 + 15(2)^2}{25}} = 2.298$$

$$t_0 = \frac{(86-89)-0}{2.298\sqrt{\frac{1}{12} + \frac{1}{15}}} = -3.37$$

8) Since $-3.37 < -2.787$, reject the null hypothesis and conclude that the mean yield of catalyst 2 significantly exceeds that of catalyst 1 at $\alpha = 0.01$.

b) 99% confidence interval: $t_{0.005,19} = 2.861$

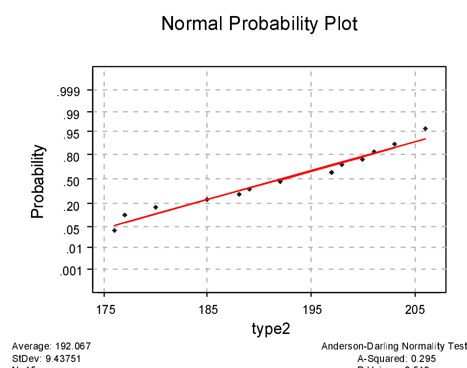
$$(\bar{x}_1 - \bar{x}_2) - t_{\alpha/2, n_1+n_2-2}(s_p)\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \le \mu_1 - \mu_2 \le (\bar{x}_1 - \bar{x}_2) + t_{\alpha/2, n_1+n_2-2}(s_p)\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

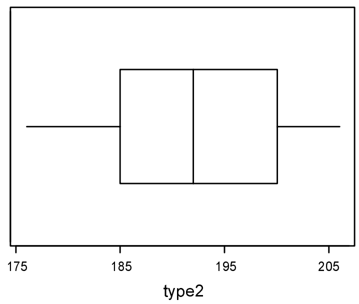
$$(86-89) - 2.787(2.298)\sqrt{\frac{1}{12} + \frac{1}{15}} \le \mu_1 - \mu_2 \le (86-89) + 2.787(2.298)\sqrt{\frac{1}{12} + \frac{1}{15}}$$

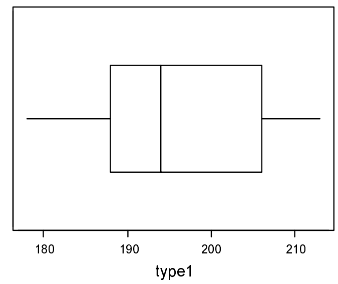
$$-5.480 \le \mu_1 - \mu_2 \le -0.519$$

We are 95% confident that the mean yield of catalyst 2 exceeds that of catalyst 1 by between 0.519 and 5.480.

9-18. a) According to the normal probability plots, the assumption of normality appears to be met since the data fall approximately along a straight line. The equality of variances does not appear to be severely violated either since the slopes are approximately the same for both samples.







b) 1) The parameter of interest is the difference in deflection temperature under load, $\mu_1 - \mu_2$

2) $H_0: \mu_1 - \mu_2 = 0$ or $\mu_1 = \mu_2$

3) $H_1: \mu_1 - \mu_2 < 0$ or $\mu_1 < \mu_2$

4) $\alpha = 0.05$

5) The test statistic is

$$t_0 = \frac{(\overline{x}_1 - \overline{x}_2) - \Delta_0}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

6) Reject the null hypothesis if $t_0 < -t_{\alpha, n_1+n_2-2}$ where $-t_{0.05,28} = -1.701$

7) Type 1   Type 2

$\overline{x}_1 = 196.4$ $\overline{x}_2 = 192.067$ $\Delta_0 = 0$

$s_1 = 10.48$ $s_2 = 9.44$

$n_1 = 15$ $n_2 = 15$

$$s_p = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}} = \sqrt{\frac{14(10.48)^2 + 14(9.44)^2}{28}} = 9.97$$

$$t_0 = \frac{(196.4 - 192.067) - 0}{9.97\sqrt{\frac{1}{15} + \frac{1}{15}}} = 1.19$$

8) Since $1.19 > -1.701$ do not reject the null hypothesis and conclude the mean deflection temperature under load for type 2 does not significantly exceed the mean deflection temperature under load for type 1 at the 0.05 level of significance.

c) P-value $= 2P(t > 1.19)$ $0.75 <$ p-value < 0.90

d) $\Delta = 5$ Use s_p as an estimate of σ:

$$d = \frac{\mu_2 - \mu_1}{2s_p} = \frac{5}{2(9.97)} = 0.251$$

Using Chart V g) with $\beta = 0.10$, $d = 0.251$ we get $n \cong 100$. So, $n_1 = n_2 = 100$; Therefore, the sample sizes of 15 are inadequate.

9-19. a) 1) The parameter of interest is the difference in mean etch rate, $\mu_1 - \mu_2$

2) $H_0: \mu_1 - \mu_2 = 0$ or $\mu_1 = \mu_2$

3) $H_1: \mu_1 - \mu_2 \neq 0$ or $\mu_1 \neq \mu_2$

4) $\alpha = 0.05$

5) The test statistic is

$$t_0 = \frac{(\bar{x}_1 - \bar{x}_2) - \Delta_0}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

6) Reject the null hypothesis if $t_0 < -t_{\alpha/2,n_1+n_2-2}$ where $-t_{0.025,18} = -2.101$ or $t_0 > t_{\alpha/2,n_1+n_2-2}$ where $t_{0.025,18} = 2.101$

7) $\bar{x}_1 = 9.97$ $\quad \bar{x}_2 = 10.4$ $\quad \Delta_0 = 0$

$s_1 = 0.422$ $\quad s_2 = 0.231$

$n_1 = 10$ $\quad n_2 = 10$

$$s_p = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}} = \sqrt{\frac{10(0.422)^2 + 10(0.231)^2}{18}} = 0.340$$

$$t_0 = \frac{(9.97 - 10.4) - 0}{0.340\sqrt{\frac{1}{10} + \frac{1}{10}}} = -2.82$$

8) Since $-2.82 < -2.101$ reject the null hypothesis and conclude the two machines mean etch rates do significantly differ at $\alpha = 0.05$.

b) P-value $= 2P\left(t < -2.82\right)$ $\quad 2(0.005) < \text{P-value} < 2(0.010) = 0.010 < \text{P-value} < 0.020$

c) 95% confidence interval: $t_{0.025,18} = 2.101$

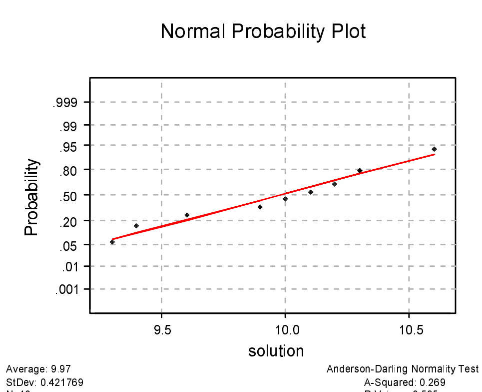
$$\left(\bar{x}_1 - \bar{x}_2\right) - t_{\alpha/2,n_1+n_2-2}(s_p)\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \le \mu_1 - \mu_2 \le \left(\bar{x}_1 - \bar{x}_2\right) + t_{\alpha/2,n_1+n_2-2}(s_p)\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

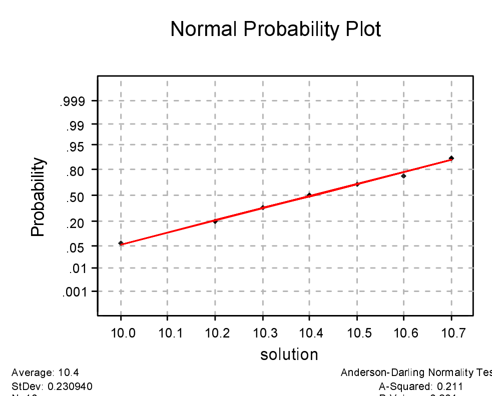
$$(9.97 - 10.4) - 2.101\sqrt{\frac{1}{10} + \frac{1}{10}} \le \mu_1 - \mu_2 \le \left(9.97 - 10.4\right) + 2.101\sqrt{\frac{1}{10} + \frac{1}{10}}$$

$$-0.749 \le \mu_1 - \mu_2 \le -0.115$$

We are 95% confident that the mean etch rate for solution 2 exceeds that for solution 1 by between 0.1105 and 0.749.

d) According to the normal probability plots, the assumption of normality appears to be met since the data fall approximately along a straight line. The equality of variances does not appear to be severely violated either since the slopes are approximately the same for both samples.





9-20. a) 1) The parameter of interest is the difference in mean impact strength, $\mu_1 - \mu_2$

2) $H_0: \mu_1 - \mu_2 = 0$ or $\mu_1 = \mu_2$

3) $H_1: \mu_1 - \mu_2 < 0$ or $\mu_1 < \mu_2$

4) $\alpha = 0.05$

5) The test statistic is

$$t_0 = \frac{(\bar{x}_1 - \bar{x}_2) - \Delta_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

6) Reject the null hypothesis if $t_0 < -t_{\alpha,\nu}$ where $t_{0.05,25} = -1.708$ since

$$\nu = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1+1} + \frac{\left(\frac{s_2^2}{n_2}\right)}{n_2+1}} - 2 = 27.43 - 2 = 25.43$$

$\nu \cong 25$
(truncated)

7) $\bar{x}_1 = 290$ $\quad \bar{x}_2 = 321$

$s_1 = 12$ $\quad s_2 = 22$

$n_1 = 10$ $\quad n_2 = 16$

$$t_0 = \frac{(290 - 321) - 0}{\sqrt{\frac{(12)^2}{10} + \frac{(22)^2}{16}}} = -4.64$$

8) Since $-4.64 < -1.708$ reject the null hypothesis and conclude that supplier 2 provides gears with higher mean impact strength at the 0.05 level of significance.

b) P-value = $P(t < -4.64)$: P-value < 0.0005

c) 1) The parameter of interest is the difference in mean impact strength, $\mu_2 - \mu_1$

2) $H_0: \mu_2 - \mu_1 = 25$

3) $H_1: \mu_2 - \mu_1 > 25$ or $\mu_2 > \mu_1 + 25$

4) $\alpha = 0.05$

5) The test statistic is

$$t_0 = \frac{(\bar{x}_2 - \bar{x}_1) - \delta}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

6) Reject the null hypothesis if $t_0 > t_{\alpha,\nu} = 1.708$ where

$$\nu = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1+1} + \frac{\left(\frac{s_2^2}{n_2}\right)}{n_2+1}} - 2 = 27.43 - 2 = 25.43$$

$\nu \cong 25$

7) $\bar{x}_1 = 290$ $\quad \bar{x}_2 = 321$ $\quad \Delta_0 = 25$ $\quad s_1 = 12$ $\quad s_2 = 22$ $\quad n_1 = 10$ $\quad n_2 = 16$

$$t_0 = \frac{(321 - 290) - 25}{\sqrt{\frac{(12)^2}{10} + \frac{(22)^2}{16}}} = 0.898$$

8) Since $0.898 < 1.708$, do not reject the null hypothesis and conclude that the mean impact strength from supplier 2 is not at least 25 ft-lb higher that supplier 1 using $\alpha = 0.05$.

9-21. Using the information provided in Exercise 9-20, and $t_{0.025,25} = 2.06$, we find a 95% confidence interval on the difference, $\mu_2 - \mu_1$:

$$(\bar{x}_2 - \bar{x}_1) - t_{0.025,25}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \le \mu_2 - \mu_1 \le (\bar{x}_2 - \bar{x}_1) + t_{0.025,25}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$$31 - 2.06(6.682) \le \mu_2 - \mu_1 \le 31 + 2.06(6.682)$$

$$17.235 \le \mu_2 - \mu_1 \le 44.765$$

Since the 95% confidence interval represents the differences that $\mu_2 - \mu_1$ could take on with 95% confidence, we can conclude that Supplier 2 does provide gears with a higher mean impact strength than Supplier 1. This is visible from the interval (17.235, 44.765) since zero is not contained in the interval and the differences are all positive, meaning that $\mu_2 - \mu_1 > 0$.

9-22. a) 1) The parameter of interest is the difference in mean speed, $\mu_1 - \mu_2$

2) $H_0: \mu_1 - \mu_2 = 0$ or $\mu_1 = \mu_2$

3) $H_1: \mu_1 - \mu_2 > 0$ or $\mu_1 > \mu_2$

4) $\alpha = 0.10$

5) The test statistic is

$$t_0 = \frac{(\bar{x}_1 - \bar{x}_2) - \Delta_0}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

6) Reject the null hypothesis if $t_0 > t_{\alpha, n_1+n_2-2}$ where $t_{0.10,14} = 1.345$

7) Case 1: 25 mil — Case 2: 20 mil

$\bar{x}_1 = 1.15$ $\bar{x}_2 = 1.06$

$s_1 = 0.11$ $s_2 = 0.09$

$n_1 = 8$ $n_2 = 8$

$$s_p = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}} = \sqrt{\frac{7(0.11)^2 + 7(0.09)^2}{14}} = 0.1005$$

$$t_0 = \frac{(1.15 - 1.06) - 0}{0.1005\sqrt{\frac{1}{8} + \frac{1}{8}}} = 1.79$$

8) Since 1.79 > 1.345 reject the null hypothesis and conclude reducing the film thickness from 25 mils to 20 mils significantly increases the mean speed of the film at the 0.10 level of significance (Note: since increase in film speed will result in *lower* values of observations).

b) P-value = $P(t > 1.79)$ $0.025 < \text{P-value} < 0.05$

c) 90% confidence interval: $t_{0.05,14} = 1.761$

$$(\bar{x}_1 - \bar{x}_2) - t_{\alpha/2,n_1+n_2-2}(s_p)\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \le \mu_1 - \mu_2 \le (\bar{x}_1 - \bar{x}_2) + t_{\alpha/2,n_1+n_2-2}(s_p)\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$(1.15 - 1.06) - 1.761\sqrt{\frac{1}{8} + \frac{1}{8}} \le \mu_1 - \mu_2 \le (1.15 - 1.06) + 1.761\sqrt{\frac{1}{8} + \frac{1}{8}}$$

$$0.005 \le \mu_1 - \mu_2 \le 0.1785$$

We are 90% confident the mean speed of the film at 20 mil exceeds the mean speed for the film at 25 mil by between 0.0015 and 0.1785 $\mu J/in^2$.

9-23. 1) The parameter of interest is the difference in mean melting point, $\mu_1 - \mu_2$

2) $H_0: \mu_1 - \mu_2 = 0$ or $\mu_1 = \mu_2$

3) $H_1: \mu_1 - \mu_2 \neq 0$ or $\mu_1 \neq \mu_2$

4) $\alpha = 0.02$

5) The test statistic is

$$t_0 = \frac{(\bar{x}_1 - \bar{x}_2) - \Delta_0}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

6) Reject the null hypothesis if $t_0 < -t_{\alpha/2,n_1+n_2-2}$ where $-t_{0.01,40} = -2.423$ or $t_0 > t_{\alpha/2,n_1+n_2-2}$ where $t_{0.01,40} = 2.423$

7) $\bar{x}_1 = 421$ $\bar{x}_2 = 426$ $\Delta_0 = 0$ $s_p = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}}$

$s_1 = 4$ $s_2 = 3$ $= \sqrt{\frac{20(4)^2 + 20(3)^2}{40}} = 2.915$

$n_1 = 21$ $n_2 = 21$

$$t_0 = \frac{(421 - 426) - 0}{2.915\sqrt{\frac{1}{20} + \frac{1}{20}}} = -5.424$$

8) Since $-5.424 < -2.423$ reject the null hypothesis and conclude that the data do not support the claim that both alloys have the same melting point at $\alpha = 0.02$

P-value = $2P\left(t < -5.424\right)$ P-value < 0.0010

9-24. $d = \frac{|\mu_1 - \mu_2|}{2\sigma} = \frac{3}{2(4)} = 0.375$

Using the appropriate chart in the Appendix, with $\beta = 0.10$ and $\alpha = 0.05$ we have: $n^* = 75$, so

$n = \frac{n^* + 1}{2} = 38$, $n_1 = n_2 = 38$

9-25. a) 1) The parameter of interest is the difference in mean wear amount, $\mu_1 - \mu_2$.

2) $H_0: \mu_1 - \mu_2 = 0$ or $\mu_1 = \mu_2$

3) $H_1: \mu_1 - \mu_2 \neq 0$ or $\mu_1 \neq \mu_2$

4) $\alpha = 0.05$

5) The test statistic is

$$t_0 = \frac{(\bar{x}_1 - \bar{x}_2) - \Delta_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

6) Reject the null hypothesis if $t_0 < -t_{0.025,27}$ where $-t_{0.025,27} = -2.052$ or $t_0 > t_{0.025,27}$ where $t_{0.025,27} = 2.052$ since

$$\nu = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1+1} + \frac{\left(\frac{s_2^2}{n_2}\right)}{n_2+1}} - 2 = 29.24 - 2 = 27.43$$

$\nu \cong 27$

(truncated)

7) $\bar{x}_1 = 20 \quad \bar{x}_2 = 15 \qquad \Delta_0 = 0$

$s_1 = 2 \quad s_2 = 8$

$n_1 = 25 \quad n_2 = 25$

$$t_0 = \frac{(20-15)-0}{\sqrt{\frac{(2)^2}{25}+\frac{(8)^2}{25}}} = 3.03$$

8) Since 3.03 > 2.052 reject the null hypothesis and conclude that the data support the claim that the two companies produce material with significantly different wear at the 0.05 level of significance.

b) P-value = 2P(t > 3.03), 2(0.0025) < P-value < 2(0.005)

0.005 < P-value < 0.010

c) 1) The parameter of interest is the difference in mean wear amount, $\mu_1 - \mu_2$

2) $H_0 : \mu_1 - \mu_2 = 0$

3) $H_1 : \mu_1 - \mu_2 > 0$

4) $\alpha = 0.05$

5) The test statistic is

$$t_0 = \frac{(\bar{x}_1 - \bar{x}_2) - \Delta_0}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}}$$

6) Reject the null hypothesis if $t_0 > t_{0.05,27}$ where $t_{0.05,27} = 1.703$ since

$$\nu = \frac{\left(\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1+1}+\frac{\left(\frac{s_2^2}{n_2}\right)}{n_2+1}} - 2 = 29.24 - 2 = 27.43$$

$\nu \cong 27$

7) $\bar{x}_1 = 20 \quad \bar{x}_2 = 15$

$s_1 = 2 \quad s_2 = 8 \qquad \Delta_0 = 0$

$n_1 = 25 \quad n_2 = 25$

$$t_0 = \frac{(20-15)-0}{\sqrt{\frac{(2)^2}{25}+\frac{(8)^2}{25}}} = 3.03$$

8) Since 3.03 > 1.703 reject the null hypothesis and conclude that the data support the claim that the material from company 1 has a higher mean wear than the material from company 2 using a 0.05 level of significance.

9-26. 1) The parameter of interest is the difference in mean coating thickness, $\mu_1 - \mu_2$.

2) $H_0: \mu_1 - \mu_2 = 0$

3) $H_1: \mu_1 - \mu_2 > 0$

4) $\alpha = 0.01$

5) The test statistic is

$$t_0 = \frac{(\bar{x}_1 - \bar{x}_2) - \delta}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

6) Reject the null hypothesis if $t_0 > t_{0.01,19}$ where $t_{0.01,19} = 2.539$ since

$$\nu = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 + 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 + 1}} - 2 = 21.13 - 2 = 19.13$$

$\nu \cong 19$
(truncated)

7) $\bar{x}_1 = 103.5 \quad \bar{x}_2 = 99.7$

$s_1 = 10.2 \quad s_2 = 20.1$

$n_1 = 11 \quad n_2 = 13$

$$t_0 = \frac{(103.5 - 99.7) - 0}{\sqrt{\frac{(10.2)^2}{11} + \frac{(20.1)^2}{13}}} = 0.597$$

8) Since 0.597 < 2.539, do not reject the null hypothesis and conclude that increasing the temperature does not significantly reduce the mean coating thickness at $\alpha = 0.01$.

P-value = P(t > 0.602), 0.25 < P-value < 0.40

9-27. If $\alpha = 0.01$, construct a 99% lower one-sided confidence interval on the difference to answer this question.

$t_{0.01,19} = 2.539$

$$\left(\bar{x}_1 - \bar{x}_2\right) - t_{\alpha,\nu}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \le \mu_1 - \mu_2$$

$$(103.5 - 99.7) - 2.539\sqrt{\frac{(10.2)^2}{12} + \frac{(20.1)^2}{13}} \le \mu_1 - \mu_2$$

$-12.21 \le \mu_1 - \mu_2$.

Since the interval contains 0, we are 99% confident there is no difference in the mean coating thickness between the two temperatures; that is, raising the process temperature does not significantly reduce the mean coating thickness.

9-28. 95% confidence interval:

$t_{0.025,27} = 2.052$

$$\left(\bar{x}_1 - \bar{x}_2\right) - t_{\alpha,\nu}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \le \mu_1 - \mu_2 \le \left(\bar{x}_1 - \bar{x}_2\right) + t_{\alpha,\nu}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$$(20 - 15) - 2.052\sqrt{\frac{(2)^2}{25} + \frac{(8)^2}{25}} \le \mu_1 - \mu_2 \le \left(20 - 15\right) + 2.052\sqrt{\frac{(2)^2}{25} + \frac{(8)^2}{25}}$$

$1.616 \le \mu_1 - \mu_2 \le 8.384$

95% lower one-sided confidence interval:

$t_{0.05,27} = 1.703$

$$\left(\bar{x}_1-\bar{x}_2\right)-t_{\alpha,\nu}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}} \le \mu_1-\mu_2$$

$$\left(20-15\right)-1.703\sqrt{\frac{(2)^2}{25}+\frac{(8)^2}{25}} \le \mu_1-\mu_2$$

$$2.19 \le \mu_1-\mu_2$$

For part a):

We are 95% confident the mean abrasive wear from company 1 exceeds the mean abrasive wear from company 2 by between 1.616 and 8.384 mg/1000.

For part c):

We are 95% confident the mean abrasive wear from company 1 exceeds the mean abrasive wear from company 2 by at least 2.19mg/1000.

Section 9-4

9-29. $\bar{d} = 0.2736 \quad s_d = 0.1356, n = 9$

95% confidence interval:

$$\bar{d}-t_{\alpha/2,n-1}\left(\frac{s_d}{\sqrt{n}}\right) \le \mu_d \le \bar{d}+t_{\alpha/2,n-1}\left(\frac{s_d}{\sqrt{n}}\right)$$

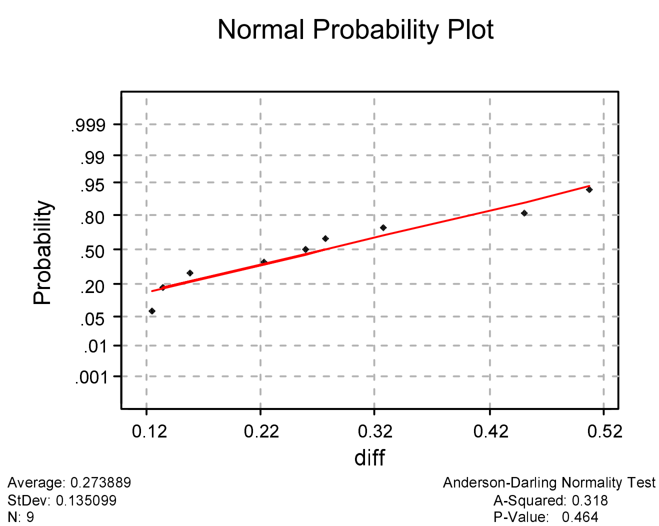
$$0.2736-2.306\left(\frac{0.1356}{\sqrt{9}}\right) \le \mu_d \le 0.2736+2.306\left(\frac{0.1356}{\sqrt{9}}\right)$$

$$0.1694 \le \mu_d \le 0.3778$$

With 95% confidence, we believe the mean shear strength of Karlsruhe method exceeds the mean shear strength of the Lehigh method by between 0.1714 and 0.3758. Since 0 is not included in this interval, the interval is consistent with rejecting the null hypothesis that the means are the same.

The 95% confidence interval is directly related to a test of hypothesis with 0.05 level of significance, and the conclusions reached are identical.

9-30. It is only necessary for the differences to be normally distributed for the paired t-test to be appropriate and reliable.



9-31. 1) The parameter of interest is the difference between the mean parking times, μ_d.

2) $H_0: \mu_d = 0$

3) $H_1: \mu_d \neq 0$

4) $\alpha = 0.10$

5) The test statistic is

$$t_0 = \frac{\bar{d}}{s_d / \sqrt{n}}$$

6) Reject the null hypothesis if $t_0 < -t_{0.05,13}$ where $-t_{0.05,13} = -1.771$ or $t_0 > t_{0.05,13}$ where $t_{0.05,13} = 1.771$

7) $\bar{d} = 1.21$

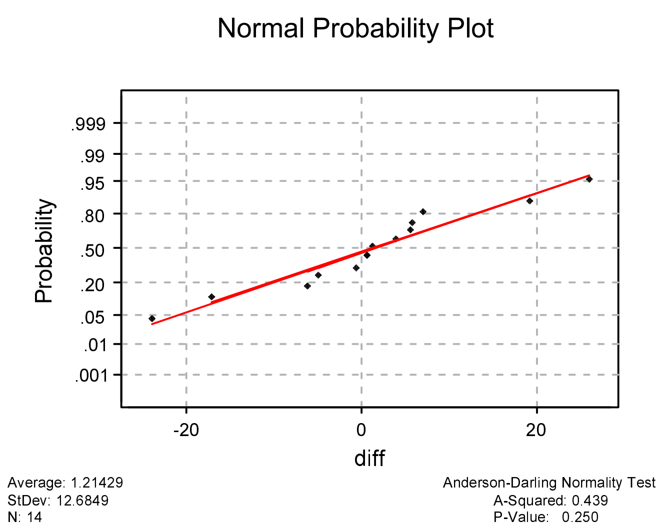
$s_d = 12.68$

$n = 14$

$$t_0 = \frac{1.21}{12.68 / \sqrt{14}} = 0.357$$

8) Since $-1.771 < 0.357 < 1.771$ do not reject the null and conclude the data do not support the claim that the two cars have different mean parking times at the 0.10 level of significance. The result is consistent with the confidence interval constructed since 0 is included in the 90% confidence interval.

9-32. According to the normal probability plots, the assumption of normality does not appear to be violated since the data fall approximately along a straight line.



9-33. $\bar{d} = 868.375$ $s_d = 1290$, $n = 8$ where d_i = brand 1 - brand 2

99% confidence interval:

$$\bar{d} - t_{\alpha/2,n-1}\left(\frac{s_d}{\sqrt{n}}\right) \le \mu_d \le \bar{d} + t_{\alpha/2,n-1}\left(\frac{s_d}{\sqrt{n}}\right)$$

$$868.375 - 3.499\left(\frac{1290}{\sqrt{8}}\right) \le \mu_d \le 868.375 + 3.499\left(\frac{1290}{\sqrt{8}}\right)$$

$$-727.46 \le \mu_d \le 2464.21$$

Since zero is contained within this interval, we are 99% confident there is no significant difference between the two brands of tire.

9-34. a) $\bar{d} = 0.667$ $s_d = 2.964$, $n = 12$

95% confidence interval:

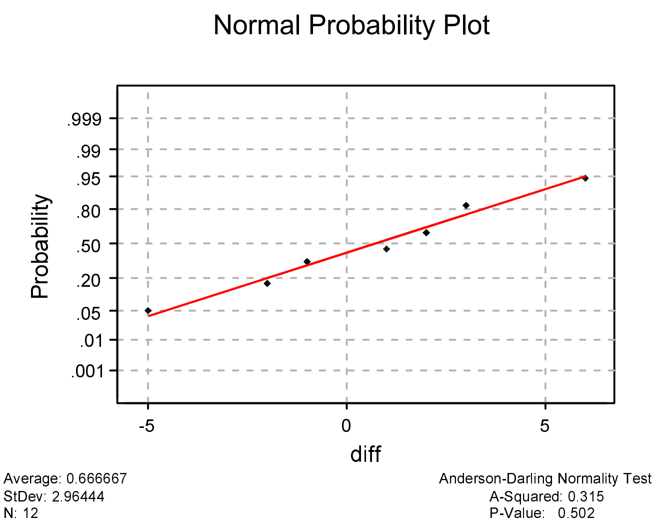
$$\bar{d} - t_{\alpha/2,n-1}\left(\frac{s_d}{\sqrt{n}}\right) \le \mu_d \le \bar{d} + t_{\alpha/2,n-1}\left(\frac{s_d}{\sqrt{n}}\right)$$

$$0.667 - 2.201\left(\frac{2.964}{\sqrt{12}}\right) \le \mu_d \le 0.667 + 2.201\left(\frac{2.964}{\sqrt{12}}\right)$$

$$-1.216 \le \mu_d \le 2.55$$

Since zero is contained within this interval, we are 95% confident there is no significant indication that one design language is preferable.

b) According to the normal probability plots, the assumption of normality does not appear to be violated since the data fall approximately along a straight line.



9-35. 1) The parameter of interest is the difference in blood cholesterol level, μ_d where d_i = Before – After.

2) $H_0 : \mu_d = 0$

3) $H_1 : \mu_d > 0$

4) $\alpha = 0.05$

5) The test statistic is

$$t_0 = \frac{\bar{d}}{s_d / \sqrt{n}}$$

6) Reject the null hypothesis if $t_0 > t_{0.05,14}$ where $t_{0.05,14} = 1.761$

7) $\bar{d} = 26.867$

$s_d = 19.04$

$n = 15$

$$t_0 = \frac{26.867}{19.04 / \sqrt{15}} = 5.465$$

8) Since $5.465 > 1.761$ reject the null and conclude the data support the claim that the mean difference in cholesterol levels is significantly less after fat diet and aerobic exercise program at the 0.05 level of significance.

9-36. a) 1) The parameter of interest is the mean difference in natural vibration frequencies, μ_d where d_i = finite element – Equivalent Plate.

2) $H_0: \mu_d = 0$

3) $H_1: \mu_d \neq 0$

4) $\alpha = 0.01$

5) The test statistic is

$$t_0 = \frac{\bar{d}}{s_d / \sqrt{n}}$$

6) Reject the null hypothesis if $t_0 < -t_{0.005,6}$ where $-t_{0.005,6} = -3.707$ or $t_0 > t_{0.005,6}$ where $t_{0.005,6}$ = 3.707

7) $\bar{d} = -5.49$
$s_d = 5.924$
$n = 7$

$$t_0 = \frac{-5.49}{5.924 / \sqrt{7}} = -2.45$$

8) Since $-3.707 < -2.45 < 3.707$, do not reject the null and conclude the data suggest that the two methods do not produce significantly different mean values for natural vibration frequency at the 0.01 level of significance.

b) 99% confidence interval:

$$\bar{d} - t_{\alpha/2,n-1}\left(\frac{s_d}{\sqrt{n}}\right) \le \mu_d \le \bar{d} + t_{\alpha/2,n-1}\left(\frac{s_d}{\sqrt{n}}\right)$$

$$-5.49 - 3.707\left(\frac{5.924}{\sqrt{7}}\right) \le \mu_d \le -5.49 + 3.707\left(\frac{5.924}{\sqrt{7}}\right)$$

$$-13.790 \le \mu_d \le 2.810$$

With 99% confidence, we believe that the mean difference between the natural vibration frequency from the equivalent plate method and the natural vibration frequency from the finite element method is between −13.790 and 2.810 cycle/s.

9-37. 1) The parameter of interest is the difference in mean weight, μ_d where d_i =Weight Before – Weight After.

2) $H_0: \mu_d = 0$

3) $H_1: \mu_d > 0$

4) $\alpha = 0.05$

5) The test statistic is

$$t_0 = \frac{\bar{d}}{s_d / \sqrt{n}}$$

6) Reject the null hypothesis if $t_0 > t_{0.05,9}$ where $t_{0.05,9} = 1.833$

7) $\bar{d} = 17$
$s_d = 6.41$
$n = 10$

$$t_0 = \frac{17}{6.41 / \sqrt{10}} = 8.387$$

8) Since $8.387 > 1.833$ reject the null and conclude there is evidence to conclude that the mean weight loss is significantly greater than 0; that is, the data support the claim that this particular diet modification program is significantly effective in reducing weight at the 0.05 level of significance.

9-38. 1) The parameter of interest is the mean difference in impurity level, μ_d
where d_i = Test 1 − Test 2.

2) $H_0: \mu_d = 0$

3) $H_1: \mu_d \neq 0$

4) $\alpha = 0.05$

5) The test statistic is

$$t_0 = \frac{\bar{d}}{s_d / \sqrt{n}}$$

6) Reject the null hypothesis if $t_0 < -t_{0.025,7}$ where $-t_{0.025,7} = -2.365$ or $t_0 > t_{0.025,7}$ where $t_{0.025,7} = 2.365$

7) $\bar{d} = -0.2125$
$s_d = 0.1727$
$n = 8$

$$t_0 = \frac{-0.2125}{0.1727 / \sqrt{8}} = -3.48$$

8) Since −3.48 < 2.365 reject the null and conclude the tests give significantly different impurity levels at α=0.05

9-39. 1) The parameter of interest is the difference in mean weight loss, μ_d
where d_i = Before − After.

2) $H_0: \mu_d = 10$

3) $H_1: \mu_d > 10$

4) $\alpha = 0.05$

5) The test statistic is

$$t_0 = \frac{\bar{d} - \Delta_0}{s_d / \sqrt{n}}$$

6) Reject the null hypothesis if $t_0 > t_{0.05,9}$ where $t_{0.05,9} = 1.833$

7) $\bar{d} = 17$
$s_d = 6.41$
$n = 10$

$$t_0 = \frac{17 - 10}{6.41 / \sqrt{10}} = 3.45$$

8) Since 3.45 > 1.833 reject the null and conclude there is evidence to support the claim that this particular diet modification program is effective in producing a mean weight loss of at least 10 lbs at the 0.05 level of significance.

9-40. Use s_d as an estimate for σ:

$$n = \left(\frac{(z_\alpha + z_\beta)\sigma_d}{10}\right)^2 = \left(\frac{(1.833 + 1.29)6.41}{10}\right)^2 = 4.007\ ,\ \ n = 5$$

Yes, the sample size of 10 is adequate for this test.

Section 9-5

9-41. a) $f_{0.25,5,10} = 1.59$

b) $f_{0.10,24,9} = 2.28$

c) $f_{0.05,8,15} = 2.64$

d) $f_{0.75,5,10} = \dfrac{1}{f_{0.25,10,5}} = \dfrac{1}{1.89} = 0.529$

e) $f_{0.90,24,9} = \dfrac{1}{f_{0.10,9,24}} = \dfrac{1}{1.91} = 0.524$

f) $f_{0.95,8,15} = \dfrac{1}{f_{0.05,15,8}} = \dfrac{1}{3.22} = 0.311$

9-42. a) $f_{0.25,7,15} = 1.47$ d) $f_{0.75,7,15} = \frac{1}{f_{0.25,15,7}} = \frac{1}{1.68} = 0.595$

b) $f_{0.10,10,12} = 2.19$ e) $f_{0.90,10,12} = \frac{1}{f_{0.10,12,10}} = \frac{1}{2.28} = 0.439$

c) $f_{0.01,20,10} = 4.41$ f) $f_{0.99,20,10} = \frac{1}{f_{0.01,10,20}} = \frac{1}{3.37} = 0.297$

9-43. 1) The parameters of interest are the variances of concentration, σ_1^2, σ_2^2

2) $H_0 : \sigma_1^2 = \sigma_2^2$

3) $H_1 : \sigma_1^2 \neq \sigma_2^2$

4) $\alpha = 0.05$

5) The test statistic is

$$f_0 = \frac{s_1^2}{s_2^2}$$

6) Reject the null hypothesis if $f_0 < f_{0.975,9,15}$ where $f_{0.975,9,15} = 0.265$ or $f_0 > f_{0.025,9,15}$ where $f_{0.025,9,15} = 3.12$

7) $n_1 = 10$ $n_2 = 16$
$s_1 = 4.7$ $s_2 = 5.8$

$$f_0 = \frac{(4.7)^2}{(5.8)^2} = 0.657$$

8) Since $0.265 < 0.657 < 3.12$ do not reject the null hypothesis and conclude there is insufficient evidence to indicate the two population variances differ significantly at the 0.05 level of significance.

9-44. 1) The parameters of interest are the etch rate variances, σ_1^2, σ_2^2.

2) $H_0 : \sigma_1^2 = \sigma_2^2$

3) $H_1 : \sigma_1^2 \neq \sigma_2^2$

4) $\alpha = 0.05$

5) The test statistic is

$$f_0 = \frac{s_1^2}{s_2^2}$$

6) Reject the null hypothesis if $f_0 < f_{0.975,9,9} = 0.248$ or $f_0 > f_{0.025,9,9} = 4.03$

7) $n_1 = 10$ $n_2 = 10$
$s_1 = 10.48$ $s_2 = 9.44$

$$f_0 = \frac{(10.48)^2}{(9.44)^2} = 1.232$$

8) Since $0.248 < 1.232 < 4.03$ do not reject the null hypothesis and conclude the etch rate variances do not differ at the 0.05 level of significance.

9-45. With $\lambda = \sqrt{2} = 1.4, \beta = 0.10,$ and $\alpha = 0.05$, we find from Chart VI o that $n_1 = n_2 = 100$. Therefore, the samples of size 10 would not be adequate.

9-46. a) 90% confidence interval for the ratio of variances:

$$\left(\frac{s_1^2}{s_2^2}\right) f_{1-\alpha/2,n_1-1,n_2-1} \le \frac{\sigma_1^2}{\sigma_2^2} \le \left(\frac{s_1^2}{s_2^2}\right) f_{\alpha/2,n_1-1,n_2-1}$$

$$\left(\frac{(0.35)^2}{(0.40)^2}\right) 0.412 \le \frac{\sigma_1^2}{\sigma_2^2} \le \left(\frac{(0.35)^2}{(0.40)^2}\right) 2.33$$

$$0.3605 \le \frac{\sigma_1^2}{\sigma_2^2} \le 2.039$$

Since the interval contains 1, we are 90% confident the variances for the rod diameters are not significantly different.

b) 95% confidence interval:

$$\left(\frac{s_1^2}{s_2^2}\right) f_{1-\alpha/2,n_1-1,n_2-1} \le \frac{\sigma_1^2}{\sigma_2^2} \le \left(\frac{s_1^2}{s_2^2}\right) f_{\alpha/2,n_1-1,n_2-1}$$

$$\left(\frac{(0.35)^2}{(0.40)^2}\right) 0.345 \le \frac{\sigma_1^2}{\sigma_2^2} \le \left(\frac{(0.35)^2}{(0.40)^2}\right) 2.75$$

$$0.302 \le \frac{\sigma_1^2}{\sigma_2^2} \le 2.406$$

We are 95% confident the variances for the rod diameters are not significantly different. The 95% confidence interval is wider than the 90% confidence interval.

c) 90% lower-sided confidence interval:

$$\left(\frac{s_1^2}{s_2^2}\right) f_{1-\alpha,n_1-1,n_2-1} \le \frac{\sigma_1^2}{\sigma_2^2}$$

$$\left(\frac{(0.35)^2}{(0.40)^2}\right) 0.503 \le \frac{\sigma_1^2}{\sigma_2^2}$$

$$0.440 \le \frac{\sigma_1^2}{\sigma_2^2}$$

With 90% confidence, we believe the variances for the rod diameters are not significantly different from each other.

9-47. a) 90% confidence interval for the ratio of variances:

$$\left(\frac{s_1^2}{s_2^2}\right) f_{1-\alpha/2,n_1-1,n_2-1} \le \frac{\sigma_1^2}{\sigma_2^2} \le \left(\frac{s_1^2}{s_2^2}\right) f_{\alpha/2,n_1-1,n_2-1}$$

$$\left(\frac{(0.6)^2}{(0.8)^2}\right) 0.156 \le \frac{\sigma_1^2}{\sigma_2^2} \le \left(\frac{(0.6)^2}{(0.8)^2}\right) 6.39$$

$$0.08775 \le \frac{\sigma_1^2}{\sigma_2^2} \le 3.594$$

Since the interval contains 1, we are 90% confident the catalyst variances are not significantly different.

b) 95% confidence interval:

$$\left(\frac{s_1^2}{s_2^2}\right) f_{1-\alpha/2,n_1-1,n_2-1} \le \frac{\sigma_1^2}{\sigma_2^2} \le \left(\frac{s_1^2}{s_2^2}\right) f_{\alpha/2,n_1-1,n_2-1}$$

$$\left(\frac{(0.6)^2}{(0.8)^2}\right) 0.104 \le \frac{\sigma_1^2}{\sigma_2^2} \le \left(\frac{(0.6)^2}{(0.8)^2}\right) 9.60$$

$$0.0585 \le \frac{\sigma_1^2}{\sigma_2^2} \le 5.4$$

We are 95% confident the catalyst variances are not significantly different. The 95% confidence interval is wider than the 90% confidence interval.

c) 90% lower-sided confidence interval:

$$\left(\frac{s_1^2}{s_2^2}\right) f_{1-\alpha,n_1-1,n_2-1} \le \frac{\sigma_1^2}{\sigma_2^2}$$

$$\left(\frac{(0.6)^2}{(0.8)^2}\right) 0.243 \le \frac{\sigma_1^2}{\sigma_2^2}$$

$$0.137 \le \frac{\sigma_1^2}{\sigma_2^2}$$

With 90% confidence, we believe the catalyst variances are not significantly different from each other.

9-48. 1) The parameters of interest are the thickness variances, σ_1^2, σ_2^2

2) $H_0 : \sigma_1^2 = \sigma_2^2$

3) $H_1 : \sigma_1^2 \neq \sigma_2^2$

4) $\alpha = 0.02$

5) The test statistic is

$$f_0 = \frac{s_1^2}{s_2^2}$$

6) Reject the null hypothesis if $f_0 < f_{0.99,7,7}$ where $f_{0.99,7,7} = 0.143$ or $f_0 > f_{0.01,7,7}$ where $f_{0.01,7,7} = 6.99$

7) $n_1 = 8$ $n_2 = 8$
$s_1 = 0.11$ $s_2 = 0.09$

$$f_0 = \frac{(0.11)^2}{(0.09)^2} = 1.49$$

8) Since 0.143 < 1.232 < 6.99 do not reject the null hypothesis and conclude the thickness variances do not significantly differ at the 0.02 level of significance.

9-49. 1) The parameters of interest are the strength variances, σ_1^2, σ_2^2

2) $H_0 : \sigma_1^2 = \sigma_2^2$

3) $H_1 : \sigma_1^2 \neq \sigma_2^2$

4) $\alpha = 0.05$

5) The test statistic is

$$f_0 = \frac{s_1^2}{s_2^2}$$

6) Reject the null hypothesis if $f_0 < f_{0.975,9,15}$ where $f_{0.975,9,15} = 0.265$ or $f_0 > f_{0.025,9,15}$ where $f_{0.025,9,15} = 3.12$

7) $n_1 = 10$ $n_2 = 16$
$s_1 = 12$ $s_2 = 22$

$$f_0 = \frac{(12)^2}{(22)^2} = 0.297$$

8) Since 0.265 < 0.297 < 3.12 do not reject the null hypothesis and conclude the population variances do not significantly differ at the 0.05 level of significance.

9-50. 1) The parameters of interest are the melting variances, σ_1^2, σ_2^2

2) $H_0: \sigma_1^2 = \sigma_2^2$

3) $H_1: \sigma_1^2 \neq \sigma_2^2$

4) $\alpha = 0.05$

5) The test statistic is

$$f_0 = \frac{s_1^2}{s_2^2}$$

6) Reject the null hypothesis if $f_0 < f_{0.975,20,20}$ where $f_{0.975,20,20} = 0.4065$ or $f_0 > f_{0.025,20,20}$ where $f_{0.025,20,20} = 2.46$

7) $n_1 = 21$ $n_2 = 21$
$s_1 = 4$ $s_2 = 3$

$$f_0 = \frac{(4)^2}{(3)^2} = 1.78$$

8) Since 0.4065 < 1.78 < 2.46 do not reject the null hypothesis and conclude the population variances do not significantly differ at the 0.05 level of significance.

9-51. 1) The parameters of interest are the thickness variances, σ_1^2, σ_2^2

2) $H_0: \sigma_1^2 = \sigma_2^2$

3) $H_1: \sigma_1^2 \neq \sigma_2^2$

4) $\alpha = 0.10$

5) The test statistic is

$$f_0 = \frac{s_1^2}{s_2^2}$$

6) Reject the null hypothesis if $f_0 < f_{0.95,10,12}$ where $f_{0.95,10,12} = 0.3436$ or $f_0 > f_{0.05,10,12}$ where $f_{0.05,10,12} = 2.75$

7) $n_1 = 11$ $n_2 = 13$
$s_1 = 10.2$ $s_2 = 20.1$

$$f_0 = \frac{(10.2)^2}{(20.1)^2} = 0.2575$$

8) Since 0.2575 < 0.3436 reject the null hypothesis and conclude the thickness variances are not equal at the 0.10 level of significance.

9-52. 1) The parameters of interest are the time to assemble standard deviations, σ_1, σ_2

2) $H_0: \sigma_1^2 = \sigma_2^2$

3) $H_1: \sigma_1^2 \neq \sigma_2^2$

4) $\alpha = 0.02$

5) The test statistic is

$$f_0 = \frac{s_1^2}{s_2^2}$$

6) Reject the null hypothesis if $f_0 < f_{1-\alpha/2,n_1-1,n_2-1} = 0..365$ or $f_0 > f_{\alpha/2,n_1-1,n_2-1} = 2.86$

7) $n_1 = 25$ $n_2 = 21$ $s_1 = 0.98$ $s_2 = 1.02$

$$f_0 = \frac{(1.02)^2}{(0.98)^2} = 0.923$$

8) Since 0.365 < 0.923 < 2.86 do not reject the null hypothesis and conclude there is no evidence to support the claim that men and women differ significantly in repeatability for this assembly task at the 0.02 level of significance.

9-53. 98% confidence interval:

$$\left(\frac{s_1^2}{s_2^2}\right) f_{1-\alpha/2, n_1-1, n_2-1} \le \frac{\sigma_1^2}{\sigma_2^2} \le \left(\frac{s_1^2}{s_2^2}\right) f_{\alpha/2, n_1-1, n_2-1}$$

$$(0.923)0.365 \le \frac{\sigma_1^2}{\sigma_2^2} \le (0.923)2.86$$

$$0.3369 \le \frac{\sigma_1^2}{\sigma_2^2} \le 2.640$$

Since the value 1 is contained within this interval, we are 98% confident there is no significant difference between the repeatability of mean and women for the assembly task.

9-54. For one population standard deviation being 50% larger than the other, then $\lambda = 2$. Using n =8, $\alpha = 0.01$ and Chart VIp, we find that $\beta \cong 0.85$. Therefore, we would say that $n = n_1 = n_2 = 8$ is adequate to detect this difference with high probability.

Section 9-6

9-55. 1) the parameters of interest are the proportion of defective parts, p_1 and p_2

2) $H_0 : p_1 = p_2$

3) $H_1 : p_1 \ne p_2$

4) $\alpha = 0.05$

5) Test statistic is

$$z_0 = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \quad \text{where}$$

$$\hat{p} = \frac{x_1 + x_2}{n_1 + n_2}$$

6) Reject the null hypothesis if $z_0 < -z_{0.025}$ where $-z_{0.025} = -1.96$ or $z_0 > z_{0.025}$ where $z_{0.025} = 1.96$

7) $n_1 = 300$ $\quad n_2 = 300$

$x_1 = 15$ $\quad x_2 = 8$

$\hat{p}_1 = 0.05$ $\quad \hat{p}_2 = 0.0267$ $\quad \hat{p} = \frac{15+8}{300+300} = 0.0383$

$$z_0 = \frac{0.05 - 0.0267}{\sqrt{0.0383(1-0.0383)\left(\frac{1}{300} + \frac{1}{300}\right)}} = 1.49$$

8) Since $-1.96 < 1.49 < 1.96$ do not reject the null hypothesis and conclude that yes the evidence indicates that there is not a significant difference in the fraction of defective parts produced by the two machines at the 0.05 level of significance.

P-value = $2(1-P(z < 1.49)) = 0.13622$

9-56. a) Power = 1 – β

$$\beta = \Phi\left(\frac{z_{\alpha/2}\sqrt{\overline{pq}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}-(p_1-p_2)}{\hat{\sigma}_{\hat{p}_1-\hat{p}_2}}\right)-\Phi\left(\frac{-z_{\alpha/2}\sqrt{\overline{pq}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}-(p_1-p_2)}{\hat{\sigma}_{\hat{p}_1-\hat{p}_2}}\right)$$

$$\overline{p} = \frac{300(0.05)+300(0.01)}{300+300} = 0.03 \qquad \overline{q} = 0.97$$

$$\hat{\sigma}_{\hat{p}_1-\hat{p}_2} = \sqrt{\frac{0.05(1-0.05)}{300}+\frac{0.01(1-0.01)}{300}} = 0.014$$

$$\beta = \Phi\left(\frac{1.96\sqrt{0.03(0.97)\left(\frac{1}{300}+\frac{1}{300}\right)}-(0.05-0.01)}{0.014}\right)-\Phi\left(\frac{-1.96\sqrt{0.03(0.97)\left(\frac{1}{300}+\frac{1}{300}\right)}-(0.05-0.01)}{0.014}\right)$$

$$= \Phi(-0.91)-\Phi(-4.81) = 0.18141-0 = 0.18141$$

Power = 1 – 0.18141 = 0.81859

$$\text{b) } n = \frac{\left(z_{\alpha/2}\sqrt{\frac{(p_1+p_2)(q_1+q_2)}{2}}+z_\beta\sqrt{p_1q_1+p_2q_2}\right)^2}{(p_1-p_2)^2}$$

$$= \frac{\left(1.96\sqrt{\frac{(0.05+0.01)(0.95+0.99)}{2}}+1.29\sqrt{0.05(0.95)+0.01(0.99)}\right)^2}{(0.05-0.01)^2} = 382.11$$

n = 383

9-57. a)

$$\beta = \Phi\left(\frac{z_{\alpha/2}\sqrt{\overline{pq}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}-(p_1-p_2)}{\hat{\sigma}_{\hat{p}_1-\hat{p}_2}}\right)-\Phi\left(\frac{-z_{\alpha/2}\sqrt{\overline{pq}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}-(p_1-p_2)}{\hat{\sigma}_{\hat{p}_1-\hat{p}_2}}\right)$$

$$\overline{p} = \frac{300(0.05)+300(0.02)}{300+300} = 0.035 \qquad \overline{q} = 0.965$$

$$\hat{\sigma}_{\hat{p}_1-\hat{p}_2} = \sqrt{\frac{0.05(1-0.05)}{300}+\frac{0.02(1-0.02)}{300}} = 0.015$$

$$\beta = \Phi\left(\frac{1.96\sqrt{0.035(0.965)\left(\frac{1}{300}+\frac{1}{300}\right)}-(0.05-0.02)}{0.015}\right)-\Phi\left(\frac{-1.96\sqrt{0.035(0.965)\left(\frac{1}{300}+\frac{1}{300}\right)}-(0.05-0.02)}{0.015}\right)$$

$$= \Phi(-0.04)-\Phi(-3.96) = 0.48405-0.00004 = 0.48401$$

Power = 1 – 0.48401 = 0.51599

$$\text{b) } n = \frac{\left(z_{\alpha/2}\sqrt{\frac{(p_1+p_2)(q_1+q_2)}{2}}+z_\beta\sqrt{p_1q_1+p_2q_2}\right)^2}{(p_1-p_2)^2}$$

$$= \frac{\left(1.96\sqrt{\frac{(0.05+0.02)(0.95+0.98)}{2}}+1.29\sqrt{0.05(0.95)+0.02(0.98)}\right)^2}{(0.05-0.02)^2} = 790.67$$

n = 791

9-58. 1) the parameters of interest are the proportion of residents in favor of an increase, p_1 and p_2

2) $H_0 : p_1 = p_2$

3) $H_1 : p_1 \neq p_2$

4) $\alpha = 0.05$

5) Test statistic is

$$z_0 = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \quad \text{where}$$

$$\hat{p} = \frac{x_1 + x_2}{n_1 + n_2}$$

6) Reject the null hypothesis if $z_0 < -z_{0.025}$ where $-z_{0.025} = -1.96$ or $z_0 > z_{0.025}$ where $z_{0.025} = 1.96$

7) $n_1 = 500 \qquad n_2 = 400$

$x_1 = 385 \qquad x_2 = 267$

$\hat{p}_1 = 0.77 \qquad \hat{p}_2 = 0.6675 \qquad \hat{p} = \frac{385 + 267}{500 + 400} = 0.724$

$$z_0 = \frac{0.77 - 0.6675}{\sqrt{0.724(1-0.724)\left(\frac{1}{500} + \frac{1}{400}\right)}} = 3.42$$

8) Since 3.42 > 1.96 reject the null hypothesis and conclude that yes the data do indicate a significant difference in the proportions of support for increasing the speed limit between residents of the two counties at the 0.05 level of significance.

P-value = $2(1-P(z < 3.42)) = 0.00062$

9-59. 95% confidence interval on the difference:

$$(\hat{p}_1 - \hat{p}_2) - z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} \le p_1 - p_2 \le (\hat{p}_1 - \hat{p}_2) + z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

$$(0.05 - 0.0267) - 1.96\sqrt{\frac{0.05(1-0.05)}{300} + \frac{0.0267(1-0.0267)}{300}} \le p_1 - p_2 \le (0.05 - 0.0267) + 1.96\sqrt{\frac{0.05(1-0.05)}{300} + \frac{0.0267(1-0.0267)}{300}}$$

$$-0.0074 \le p_1 - p_2 \le 0.054$$

Since this interval contains the value zero, we are 95% confident there is no significant difference in the fraction of defective parts produced by the two machines and that the difference in proportions is between –0.0074 and 0.054.

9-60. 95% confidence interval on the difference:

$$(\hat{p}_1 - \hat{p}_2) - z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} \le p_1 - p_2 \le (\hat{p}_1 - \hat{p}_2) + z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

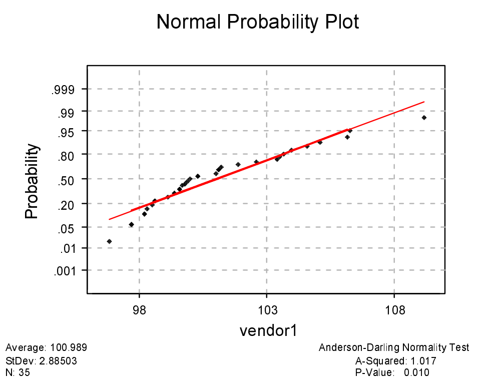
$$(0.77 - 0.6675) - 1.96\sqrt{\frac{0.77(1-0.77)}{500} + \frac{0.6675(1-0.6675)}{400}} \le p_1 - p_2 \le (0.77 - 0.6675) + 1.96\sqrt{\frac{0.77(1-0.77)}{500} + \frac{0.6675(1-0.6675)}{400}}$$

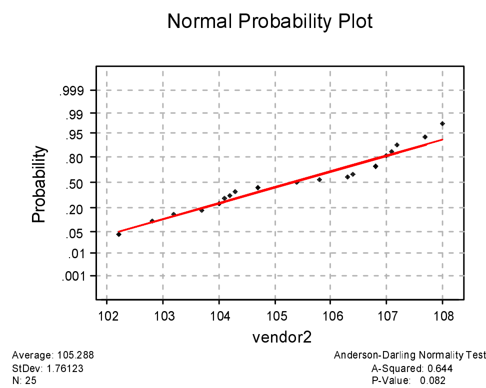
$$0.0377 \le p_1 - p_2 \le 0.1673$$

Since this interval does not contain the value zero, we are 95% confident there is a significant difference in the proportions of support for increasing the speed limit between residents of the two counties and that the difference in proportions is between 0.0377 and 0.1673.

Supplemental Exercises

9-61. a) The assumption of normality is necessary to test the claim. According to the normal probability plots, the assumption of normality does not appear to be violated. This is evident from the fact that the data appear to fall along a straight line.





b) 1) the parameters of interest are the variances of resistance of products, σ_1^2, σ_2^2

2) $H_0 : \sigma_1^2 = \sigma_2^2$

3) $H_1 : \sigma_1^2 \neq \sigma_2^2$

4) $\alpha = 0.05$

5) The test statistic is

$$f_0 = \frac{s_1^2}{s_2^2}$$

6) Reject H_0 if $f_0 < f_{0.975,24,34}$ where $f_{0.975,24,34} = \frac{1}{f_{0.025,34,24}} = \frac{1}{2.18} = 0.459$

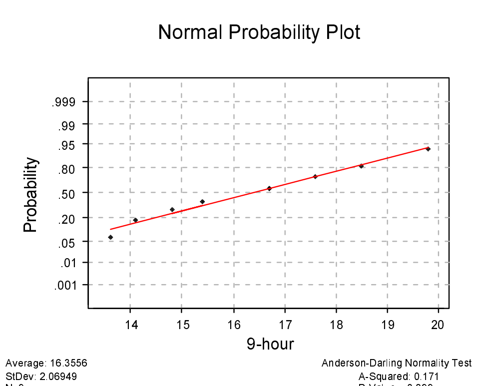
or $f_0 > f_{0.025,24,34}$ where $f_{0.025,24,34} = 2.07$

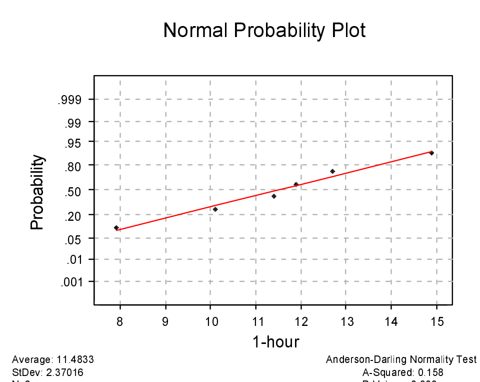
7) $s_1 = 1.53$ $\quad s_2 = 1.96$
$n_1 = 25$ $\quad n_2 = 35$

$$f_0 = \frac{(1.53)^2}{(1.96)^2} = 0.609$$

8) Since 0.061 < 0.459, reject H_0 and conclude the variances are significantly different at $\alpha = 0.05$.

9-62. a) Normality and equality of variances appears to be reasonable, see normal probability plot. The data appear to fall along a straight line and the slopes appear to be the same.





b) $\overline{x}_1 = 16.36$ $\overline{x}_2 = 11.486$
$s_1 = 2.07$ $s_2 = 2.37$
$n_1 = 9$ $n_2 = 6$

99% confidence interval: $t_{\alpha/2,n_1+n_2-2} = t_{0.005,13}$ where $t_{0.005,13} = 3.012$

$$s_p = \sqrt{\frac{8(2.07)^2 + 5(2.37)^2}{13}} = 2.19$$

$$\left(\overline{x}_1 - \overline{x}_2\right) - t_{\alpha/2,n_1+n_2-2}\left(s_p\right)\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \le \mu_1 - \mu_2 \le \left(\overline{x}_1 - \overline{x}_2\right) + t_{\alpha/2,n_1+n_2-2}\left(s_p\right)\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$\left(16.36 - 11.486\right) - 3.012\left(2.19\right)\sqrt{\frac{1}{9} + \frac{1}{6}} \le \mu_1 - \mu_2 \le \left(16.36 - 11.486\right) + 3.012\left(2.19\right)\sqrt{\frac{1}{9} + \frac{1}{6}}$$

$$1.40 \le \mu_1 - \mu_2 \le 8.36$$

c) Yes, we are 99% confident the results from the first test condition exceed the results of the second test condition by between 1.40 and 8.36 ($\times 10^6$ PA).

9-63. a) 95% confidence interval for σ_1^2 / σ_2^2

95% confidence interval on $\frac{\sigma_1^2}{\sigma_2^2}$:

$$f_{0.975,8,5} = \frac{1}{f_{0.025,5,8}} = \frac{1}{4.82} = 0.2075, \qquad f_{0.025,8,5} = 6.76$$

$$\frac{s_1^2}{s_2^2} f_{0.975,8,5} \le \frac{\sigma_1^2}{\sigma_2^2} \le \frac{s_1^2}{s_2^2} f_{0.025,8,5}$$

$$\left(\frac{4.285}{5.617}\right)(0.2075) \le \frac{\sigma_1^2}{\sigma_2^2} \le \left(\frac{4.285}{5.617}\right)(6.76)$$

$$0.1583 \le \frac{\sigma_1^2}{\sigma_2^2} \le 5.157$$

b) Since the value 1 is contained within this interval, with 95% confidence, the population variances do not differ significantly and can be assumed to be equal.

9-64. a) 1) The parameter of interest is the mean weight loss, μ_d
where d_i = Initial Weight − Final Weight.
2) $H_0 : \mu_d = 3$
3) $H_1 : \mu_d > 3$
4) $\alpha = 0.05$
5) The test statistic is

$$t_0 = \frac{\overline{d} - \Delta_0}{s_d / \sqrt{n}}$$

6) Reject H_0 if $t_0 > t_{\alpha,n-1}$ where $t_{0.05,7} = 1.895$.
7) $\overline{d} = 4.125$
$s_d = 1.246$
$n = 8$

$$t_0 = \frac{4.125 - 3}{1.246 / \sqrt{8}} = 2.554$$

8) Since 2.554 < 1.895, reject the null hypothesis and conclude the average weight loss is significantly greater than 3 at $\alpha = 0.05$.

b) 2) $H_0 : \mu_d = 3$
3) $H_1 : \mu_d > 3$
4) $\alpha = 0.01$

5) The test statistic is

$$t_0 = \frac{\bar{d} - \Delta_0}{s_d / \sqrt{n}}$$

6) Reject H_0 if $t_0 > t_{\alpha,n-1}$ where $t_{0.01,7} = 2.998$.

7) $\bar{d} = 4.125$
$s_d = 1.246$
$n = 8$

$$t_0 = \frac{4.125 - 3}{1.246 / \sqrt{8}} = 2.554$$

8) Since 2.554 <2.998, do not reject the null hypothesis and conclude the average weight loss is not significantly greater than 3 at $\alpha = 0.01$.

c) 2) $H_0 : \mu_d = 5$

3) $H_1 : \mu_d > 5$

4) $\alpha = 0.05$

5) The test statistic is

$$t_0 = \frac{\bar{d} - \Delta_0}{s_d / \sqrt{n}}$$

6) Reject H_0 if $t_0 > t_{\alpha,n-1}$ where $t_{0.05,7} = 1.895$.

7) $\bar{d} = 4.125$
$s_d = 1.246$
$n = 8$

$$t_0 = \frac{4.125 - 5}{1.246 / \sqrt{8}} = -1.986$$

8) Since −1.986 < 1.895, do not reject the null hypothesis and conclude the average weight loss is not significantly greater than 5 at $\alpha = 0.05$.

Using $\alpha = 0.01$

2) $H_0 : \mu_d = 5$

3) $H_1 : \mu_d > 5$

4) $\alpha = 0.01$

5) The test statistic is

$$t_0 = \frac{\bar{d} - \Delta_0}{s_d / \sqrt{n}}$$

6) Reject H_0 if $t_0 > t_{\alpha,n-1}$ where $t_{0.01,7} = 2.998$.

7) $\bar{d} = 4.125$
$s_d = 1.246$
$n = 8$

$$t_0 = \frac{4.125 - 5}{1.246 / \sqrt{8}} = -1.986$$

8) Since −1.986 < 2.998, do not reject the null hypothesis and conclude the average weight loss is not significantly greater than 5 at $\alpha = 0.01$.

9-65. $$\left(\bar{x}_1 - \bar{x}_2\right) - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \le \mu_1 - \mu_2 \le \left(\bar{x}_1 - \bar{x}_2\right) + z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

a) 90% confidence interval: $z_{\alpha/2} = 1.65$

$$\left(88 - 91\right) - 1.65\sqrt{\frac{5^2}{20} + \frac{4^2}{20}} \le \mu_1 - \mu_2 \le \left(88 - 91\right) + 1.65\sqrt{\frac{5^2}{20} + \frac{4^2}{20}}$$

$$-5.362 \le \mu_1 - \mu_2 \le -0.638$$

Yes, with 90% confidence, the data indicate that the mean breaking strength of the yarn of manufacturer 2 exceeds that of manufacturer 1 by between 5.362 and 0.638.

b) 98% confidence interval: $z_{\alpha/2} = 2.33$

$$\left(88-91\right)-2.33\sqrt{\frac{5^2}{20}+\frac{4^2}{20}} \le \mu_1-\mu_2 \le \left(88-91\right)+2.33\sqrt{\frac{5^2}{20}+\frac{4^2}{20}}$$

$-6.340 \le \mu_1 - \mu_2 \le 0.340$

Yes, we are 98% confident manufacturer 2 produces yarn with higher breaking strength by between 0.340 and 6.340 psi.

c) The results of parts a) and b) are different because the confidence level or z-value used is different.. Which one is used depends upon the level of confidence considered acceptable.

9-66. a) $\alpha = 0.10$ $\quad z_{\alpha/2} = 1.65$

$$n \cong \frac{\left(z_{\alpha/2}\right)^2\left(\sigma_1^2+\sigma_2^2\right)}{(E)^2} \cong \frac{\left(1.65\right)^2(25+16)}{(1.5)^2} = 49.61, \quad n = 50$$

b) $\alpha = 0.10$ $\quad z_{\alpha/2} = 2.33$

$$n \cong \frac{\left(z_{\alpha/2}\right)^2\left(\sigma_1^2+\sigma_2^2\right)}{(E)^2} \cong \frac{\left(2.33\right)^2(25+16)}{(1.5)^2} = 98.93, \quad n = 99$$

c) As the confidence level increases, sample size will also increase.

d) $\alpha = 0.10$ $\quad z_{\alpha/2} = 1.65$

$$n \cong \frac{\left(z_{\alpha/2}\right)^2\left(\sigma_1^2+\sigma_2^2\right)}{(E)^2} \cong \frac{\left(1.65\right)^2(25+16)}{(0.75)^2} = 198.44, \quad n = 199$$

e) $\alpha = 0.10$ $\quad z_{\alpha/2} = 2.33$

$$n \cong \frac{\left(z_{\alpha/2}\right)^2\left(\sigma_1^2+\sigma_2^2\right)}{(E)^2} \cong \frac{\left(2.33\right)^2(25+16)}{(0.75)^2} = 395.70, \quad n =396$$

f) As the error decreases, the required sample size increases.

9-67. a) 1) The parameters of interest are the proportions of children who contract polio, p_1 , p_2

2) $H_0 : p_1 = p_2$

3) $H_1 : p_1 \neq p_2$

4) $\alpha = 0.05$

5) The test statistic is

$$z_0 = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$

6) Reject H_0 if $z_0 < -z_{\alpha/2}$ or $z_0 > z_{\alpha/2}$ where $z_{\alpha/2} = 1.96$

7) $\hat{p}_1 = \frac{x_1}{n_1} = \frac{110}{201299} = 0.00055$ (Placebo) $\qquad \hat{p} = \frac{x_1 + x_2}{n_1 + n_2} = 0.000356$

$\hat{p}_2 = \frac{x_2}{n_2} = \frac{33}{200745} = 0.00016$ (Vaccine)

$$z_0 = \frac{0.00055 - 0.00016}{\sqrt{0.000356(1-0.000356)\left(\frac{1}{201299}+\frac{1}{200745}\right)}} = 6.55$$

8) Since 6.55 > 1.96 reject H_0 and conclude the proportion of children who contracted polio is significantly different at $\alpha = 0.05$.

b) $\alpha = 0.01$

Reject H_0 if $z_0 < -z_{\alpha/2}$ or $z_0 > z_{\alpha/2}$ where $z_{\alpha/2}$ =2.33

$z_0 = 6.55$

Since 6.55 > 2.33, reject H_0 and conclude the proportion of children who contracted polio is different at $\alpha = 0.05$.

c) The conclusions are the same since z_0 is so large it exceeds $z_{\alpha/2}$ in both cases.

9-68. $\hat{p}_1 = \frac{x_1}{n_1} = \frac{387}{1500} = 0.258$ $\hat{p}_2 = \frac{x_2}{n_2} = \frac{310}{1200} = 0.2583$

$$\left(\hat{p}_1 - \hat{p}_2\right) \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

a) $z_{\alpha/2} = z_{0.025} = 1.96$

$$\left(0.258 - 0.2583\right) \pm 1.96\sqrt{\frac{0.258(0.742)}{1500} + \frac{0.2583(0.7417)}{1200}}$$

$$-0.0335 \le p_1 - p_2 \le 0.0329$$

Since zero is contained in this interval, we are 95% confident there is no significant difference between the proportion of unlisted numbers in the two cities.

b) $z_{\alpha/2} = z_{0.05} = 1.65$

$$\left(0.258 - 0.2583\right) \pm 1.65\sqrt{\frac{0.258(0.742)}{1500} + \frac{0.2583(0.7417)}{1200}}$$

$$-0.0282 \le p_1 - p_2 \le 0.0276$$

Again, the proportion of unlisted numbers in the two cities do not differ.

c) $\hat{p}_1 = \frac{x_1}{n_1} = \frac{774}{3000} = 0.258$

$\hat{p}_2 = \frac{x_2}{n_2} = \frac{620}{2400} = 0.2583$

95% confidence interval:

$$\left(0.258 - 0.2583\right) \pm 1.96\sqrt{\frac{0.258(0.742)}{3000} + \frac{0.2583(0.7417)}{2400}}$$

$$-0.0238 \le p_1 - p_2 \le 0.0232$$

90% confidence interval:

$$\left(0.258 - 0.2583\right) \pm 1.65\sqrt{\frac{0.258(0.742)}{3000} + \frac{0.2583(0.7417)}{2400}}$$

$$-0.0203 \le p_1 - p_2 \le 0.0200$$

Increasing the sample size decreased the error and width of the confidence intervals, but does not change the conclusions drawn. The conclusion remains that there is no significant difference.

9-69. a) 1) The parameters of interest are the proportions of those residents who wear a seat belt regularly, p_1 , p_2

2) $H_0 : p_1 = p_2$

3) $H_1 : p_1 \neq p_2$

4) $\alpha = 0.05$

5) The test statistic is

$$z_0 = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

6) Reject H_0 if $z_0 < -z_{\alpha/2}$ or $z_0 > z_{\alpha/2}$ where $z_{0.025} = 1.96$

7) $\hat{p}_1 = \frac{x_1}{n_1} = \frac{165}{200} = 0.825$ $\hat{p} = \frac{x_1 + x_2}{n_1 + n_2} = 0.807$

$\hat{p}_2 = \frac{x_2}{n_2} = \frac{198}{250} = 0.792$

$$z_0 = \frac{0.825 - 0.792}{\sqrt{0.807(1-0.807)\left(\frac{1}{200} + \frac{1}{250}\right)}} = 0.715$$

8) Since $-1.96 < 0.715 < 1.96$ do not reject H_0 and conclude that evidence is insufficient to claim that there is a difference in seat belt usage $\alpha = 0.05$.

b) $\alpha = 0.10$

Reject H_0 if $z_0 < -z_{\alpha/2}$ or $z_0 > z_{\alpha/2}$ where $z_{0.05} = 1.65$

$z_0 = 0.715$

Since $-1.65 < 0.715 < 1.65$, do not reject H_0 and conclude that evidence is insufficient to claim that there is a difference in seat belt usage $\alpha = 0.10$.

c) The conclusions are the same, but with different levels of confidence.

d) $n_1 = 400$, $n_2 = 500$

$\alpha = 0.05$

Reject H_0 if $z_0 < -z_{\alpha/2}$ or $z_0 > z_{\alpha/2}$ where $z_{0.025} = 1.96$

$$z_0 = \frac{0.825 - 0.792}{\sqrt{0.807(1-0.807)\left(\frac{1}{400}+\frac{1}{500}\right)}} = 1.012$$

Since $-1.96 < 1.012 < 1.96$ do not reject H_0 and conclude that evidence is insufficient to claim that there is a difference in seat belt usage $\alpha = 0.05$.

$\alpha = 0.10$

Reject H_0 if $z_0 < -z_{\alpha/2}$ or $z_0 > z_{\alpha/2}$ where $z_{0.05} = 1.65$

$z_0 = 1.012$

Since $-1.65 < 1.012 < 1.65$, do not reject H_0 and conclude that evidence is insufficient to claim that there is a difference in seat belt usage $\alpha = 0.10$.

As the sample size increased, the test statistic has also increased, since the denominator of z_0 decreased. However, the decrease (or sample size increase) was not enough to change our conclusion.

9-70. a) Yes, there could be some bias in the results due to the telephone survey.

b) If it could be shown that these populations are similar to the respondents, the results may be extended.

9-71. a) 1) The parameters of interest are the proportion of lenses that are unsatisfactory after tumble-polishing, p_1, p_2

2) $H_0 : p_1 = p_2$

3) $H_1 : p_1 \neq p_2$

4) $\alpha = 0.01$

5) The test statistic is

$$z_0 = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$

6) Reject H_0 if $z_0 < -z_{\alpha/2}$ or $z_0 > z_{\alpha/2}$ where $z_{\alpha/2} = 2.58$

7) x_1 =number of defective lenses

$$\hat{p}_1 = \frac{x_1}{n_1} = \frac{47}{300} = 0.1567 \qquad \hat{p} = \frac{x_1 + x_2}{n_1 + n_2} = 0.2517$$

$$\hat{p}_2 = \frac{x_2}{n_2} = \frac{104}{300} = 0..3467$$

$$z_0 = \frac{0.1567 - 0.3467}{\sqrt{0.2517(1-0.2517)\left(\frac{1}{300}+\frac{1}{300}\right)}} = -5.36$$

8) Since $-5.36 < -2.58$ reject H_0 and conclude there is strong evidence to support the claim that the two polishing fluids are different.

b) The conclusions are the same whether we analyze the data using the proportion unsatisfactory or proportion satisfactory. The proportion of defectives are different for the two fluids.

9-72.

$$n = \frac{\left(2.575\sqrt{\frac{(0.9+0.6)(0.1+0.4)}{2}} + 1.28\sqrt{0.9(0.1)+0.6(0.4)}\right)^2}{(0.9-0.6)^2}$$

$$= \frac{5.346}{0.09} = 59.4$$

$n = 60$

9-73. The parameter of interest is $\mu_1 - 2\mu_2$

$H_0: \mu_1 = 2\mu_2 \quad \rightarrow \quad H_0: \mu_1 - 2\mu_2 = 0$

$H_1: \mu_1 > 2\mu_2 \quad\quad\quad H_1: \mu_1 - 2\mu_2 > 0$

Let n_1 = size of sample 1 $\quad \overline{X}_1$ estimate for μ_1

Let n_2 = size of sample 2 $\quad \overline{X}_2$ estimate for μ_2

$\overline{X}_1 - 2\overline{X}_2$ is an estimate for $\mu_1 - 2\mu_2$

The variance is $V(\overline{X}_1 - 2\overline{X}_2) = V(\overline{X}_1) + V(2\overline{X}_2) = \frac{\sigma_1^2}{n_1} + \frac{4\sigma_2^2}{n_2}$

The test statistic for this hypothesis would then be:

$$Z_0 = \frac{(\overline{X}_1 - 2\overline{X}_2) - 0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{4\sigma_2^1}{n_2}}}$$

We would reject the null hypothesis if $z_0 > z_{\alpha/2}$ for a given level of significance.
The P-value would be $P(Z \geq z_0)$.

9-74. $H_0 : \mu_1 = \mu_2$

$H_1 : \mu_1 \neq \mu_2$

$n_1 = n_2 = n$

$\beta = 0.10$

$\alpha = 0.05$

Assume normal distribution and $\sigma_1^2 = \sigma_2^2 = \sigma^2$

$\mu_1 = \mu_2 + \sigma$

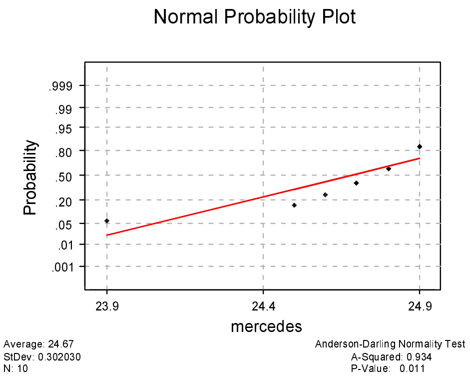
$$d = \frac{|\mu_1 - \mu_2|}{2\sigma} = \frac{\sigma}{2\sigma} = \frac{1}{2}$$

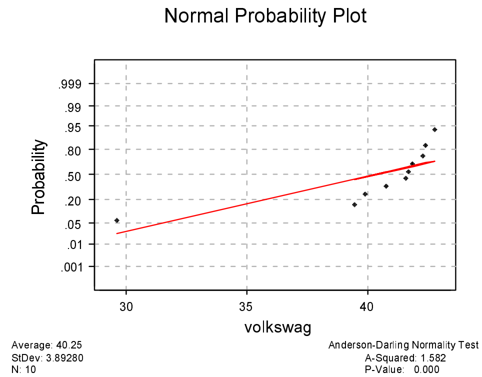
From Chart V e), $n^* = 40$

$$n = \frac{n^* + 1}{2} = \frac{40+1}{2} = 20.5$$

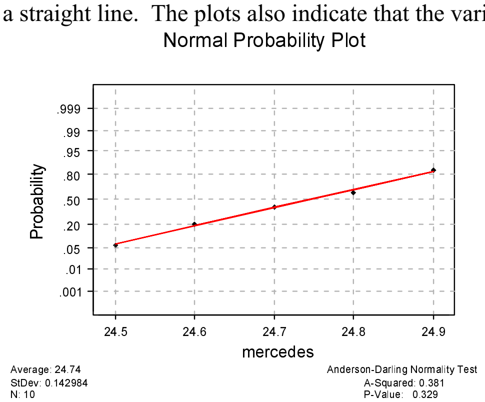
$n_1 = n_2 = 21$

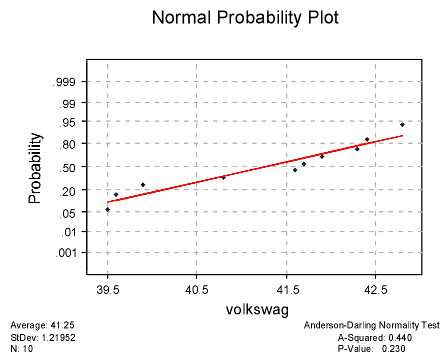
9-75. a) No.





b) The normal probability plots indicate that the data follow normal distributions since the data appear to fall along a straight line. The plots also indicate that the variances could be equal since the slopes appear to be the same.





c) By correcting the data points, it is more apparent the data follow normal distributions. Apparently, one observation can cause an analyst to reject the normality assumption.

d) 95% confidence interval on the ratio of the variances, σ_V^2 / σ_M^2

$s_V^2 = 1.49 \qquad f_{9,9,0.025} = 4.03$

$$s_M^2 = 0.0204 \qquad f_{9,9,0.975} = \frac{1}{f_{9,9,0.025}} = \frac{1}{4.03} = 0.248$$

$$\left(\frac{s_V^2}{s_M^2}\right) f_{9,9,0.975} < \frac{\sigma_V^2}{\sigma_M^2} < \left(\frac{s_V^2}{s_M^2}\right) f_{9,9,0.025}$$

$$\left(\frac{1.49}{0.0204}\right) 0.248 < \frac{\sigma_V^2}{\sigma_M^2} < \left(\frac{1.49}{0.0204}\right) 4.03$$

$$18.124 < \frac{\sigma_V^2}{\sigma_M^2} < 294.35$$

Since the interval covers a range larger than 1 and not including 1, we are 95% confident that there is evidence to support the claim that the variability in mileage performance is greater for a Volkswagen than for a Mercedes.

9-76. 1) the parameters of interest are the variances in mileage performance, σ_1^2, σ_2^2

2) $H_0 : \sigma_1^2 = \sigma_2^2$ Where Volkswagen is represented by variance 1, Mercedes by variance 2.

3) $H_1 : \sigma_1^2 \neq \sigma_2^2$

4) $\alpha = 0.05$

5) The test statistic is

$$f_0 = \frac{s_1^2}{s_2^2}$$

6) Reject H_0 if $f_0 < f_{0.975,9,9}$ where $f_{0.975,9,9} = \frac{1}{f_{0.025,9,9}} = \frac{1}{4.03} = 0.248$

or $f_0 > f_{0.025,9,9}$ where $f_{0.025,9,9} = 4.03$

7) $s_1 = 1.22$ $\quad s_2 = 0.143$
$n_1 = 10$ $\quad n_2 = 10$

$$f_0 = \frac{(1.22)^2}{(0.143)^2} = 72.78$$

8) Since 72.78 > 4.03, reject H_0 and conclude that there is a significant difference between Volkswagen and Mercedes in terms of mileage variability. Same conclusions reached in 9-75d.

9-77. a) $\alpha = 0.05$, $z_{\alpha/2} = 1.96$, $E = 0.5$

$$n = \left(\frac{z_{\alpha/2}}{E}\right)^2 (s_1^2 + s_2^2)$$

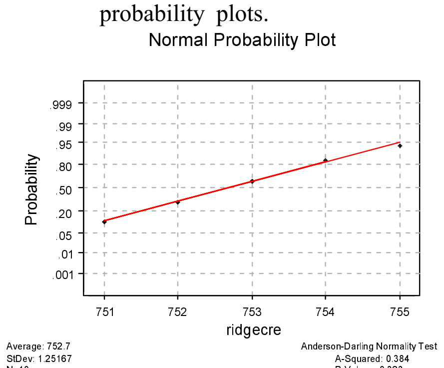
$$n = \left(\frac{1.96}{0.5}\right)^2 (0.6^2 + 0.8^2)$$

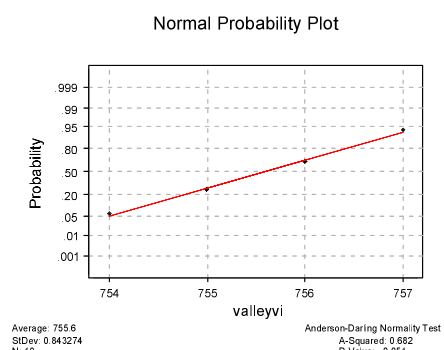
$$= 15.37$$

Need n = 16 to reject the null hypothesis that the two agents differ by 0.5 with probability of at least 0.95

b) The original size of n = 5 in Exercise 9-16 was not appropriate to detect the difference since it is necessary for a sample size of 16 to reject the null hypothesis that the two agents differ by 0.5 with probability of at least 0.95.

9-78. a) Underlying distributions appear to be normal since the data fall along a straight line on the normal probability plots.





A test of hypothesis can be conducted to test for equality of variances:

1) the parameters of interest are the variances of the fill volumes, σ_1^2, σ_2^2

2) $H_0 : \sigma_1^2 = \sigma_2^2$

3) $H_1 : \sigma_1^2 \neq \sigma_2^2$

4) $\alpha = 0.05$

5) The test statistic is

$$f_0 = \frac{s_1^2}{s_2^2}$$

6) Reject H_0 if $f_0 < f_{0.975,9,9}$ where $f_{0.975,9,9} = \dfrac{1}{f_{0.025,9,9}} = \dfrac{1}{4.03} = 0.248$

or $f_0 > f_{0.025,9,9}$ where $f_{0.025,9,9} = 4.03$

7) $s_1 = 1.252$ $s_2 = 0.843$
$n_1 = 10$ $n_2 = 10$

$$f_0 = \frac{(1.252)^2}{(0.843)^2} = 2.206$$

8) Since 0.248 < 2.206 < 4.03, do not reject H_0 and conclude the variances are not significantly different at $\alpha = 0.05$.

b) 1) The parameter of interest is the difference in mean volumes, $\mu_1 - \mu_2$

2) $H_0: \mu_1 - \mu_2 = 0$ or $\mu_1 = \mu_2$

3) $H_1: \mu_1 - \mu_2 \neq 0$ or $\mu_1 \neq \mu_2$

4) $\alpha = 0.05$

5) The test statistic is

$$t_0 = \frac{(\overline{x}_1 - \overline{x}_2) - \delta}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

6) Reject H_0 if $t_0 < -t_{\alpha/2,\nu}$ or $z_0 > t_{\alpha/2,\nu}$ where $t_{\alpha/2,\nu} = t_{0.025,18} = 2.101$

7) $\overline{x}_1 = 752.7$ $\overline{x}_2 = 755.6$ $s_p = \sqrt{\dfrac{9(1.252)^2 + 9(0.843)^2}{18}} = 1.07$

$s_1 = 1.252$ $s_2 = 0.843$

$n_1 = 10$ $n_2 = 10$

$$t_0 = \frac{(752.7 - 755.6) - 0}{1.07\sqrt{\frac{1}{10} + \frac{1}{10}}} = -6.06$$

8) Since –6.06 < –2.101, reject H_0 and conclude there is a significant difference between mean fill volumes.

9-79. $\alpha = 0.05$, $z_{\alpha/2} = 1.96$, E = 2,

$$n = \left(\frac{z_{\alpha/2}}{E}\right)^2 (s_1^2 + s_2^2)$$

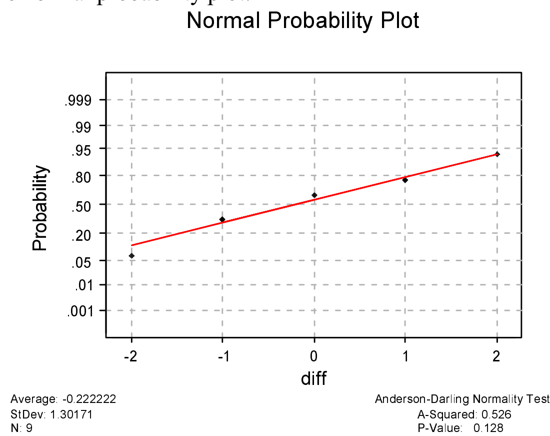
$$n = \left(\frac{1.96}{2}\right)^2 (1.252^2 + 0.843^2)$$

= 2.188

Use n = 3

Since it is necessary that the sample size be at least 3 to detect that the true difference is as much as 2 fluid ounces, the sample sizes of 10 from each vineyard do provide good detection capability when $\alpha = 0.05$.

9-80. a) The assumption of normality appears to be valid. This is evident by the fact that the data lie along a straight line in the normal probability plot.



b) 1) The parameter of interest is the mean difference in tip hardness, μ_d

2) $H_0 : \mu_d = 0$

3) $H_1 : \mu_d \neq 0$

4) No significance level, calculate P-value

5) The test statistic is

$$t_0 = \frac{\overline{d}}{s_d / \sqrt{n}}$$

6) Reject H_0 if the P-value is significantly small.

7) $\overline{d} = -0.222$

$s_d = 1.30$

$n = 9$

$$t_0 = \frac{-0.222}{1.30 / \sqrt{9}} = -0.512$$

8) P-value = 2P(T < -0.512) = 2P(T > 0.512)   2(0.25) < P-value < 2(0.40)

0.50 < P-value < 0.80

Since the P-value is larger than any acceptable level of significance, do not reject H_0 and conclude there is no difference in mean tip hardness.

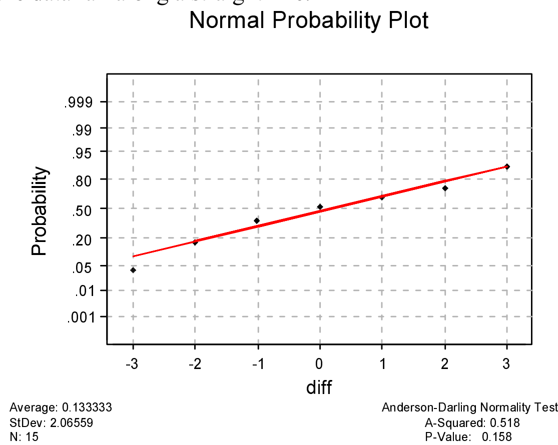
c) $\beta = 0.10$

$\mu_d = 1$

$$d = \frac{1}{\sigma_d} = \frac{1}{1.3} = 0.769$$

From Chart VI with $\alpha = 0.01$, n = 30 coupons

9-81. a) According to the normal probability plot the data appear to follow a normal distribution. This is evident by the fact that the data fall along a straight line.



b) 1) The parameter of interest is the mean difference in depth using the two gauges, μ_d

2) $H_0 : \mu_d = 0$

3) $H_1 : \mu_d \neq 0$

4) No significance level, calculate p-value

5) The test statistic is

$$t_0 = \frac{\bar{d}}{s_d / \sqrt{n}}$$

6) Reject H_0 if the P-value is significantly small.

7) $\bar{d} = 0.133$

$s_d = 2.065$

$n = 15$

$$t_0 = \frac{0.133}{2.065/\sqrt{15}} = 0.25$$

8) P-value = 2P(T > 0.25) $\quad\quad$ 2(0.40) < P-value

0.80 < P-value

Since the P-value is larger than any acceptable level of significance, do not reject H_0 and conclude there is no difference in mean depth measurements for the two gauges.

c) $\beta = 0.20$

$\mu_d = 1.65$

$$d = \frac{1.65}{2\sigma_d} = \frac{1.65}{2(2.065)} = 0.40$$

From Chart VI with $\alpha = 0.01$ and $1 - \beta = 0.80$, we find n = 20.

Mind-Expanding Exercises

9-82. The estimate of μ is given by $\bar{X}$. Therefore, $\bar{X} = \frac{1}{2}(\bar{X}_1 + \bar{X}_2) - \bar{X}_3$. The variance of $\bar{X}$ can be shown to be: $V(\bar{X}) = \frac{1}{4}\left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right) + \frac{\sigma_3^2}{n_3}$. Using s_1, s_2, and s_3 as estimates for σ_1 σ_2 and σ_3 respectively.

a) A 100(1-α)% confidence interval on μ is then:

$$\bar{X} - Z_{\alpha/2}\sqrt{\frac{1}{4}\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right) + \frac{s_3^2}{n_3}} \leq \mu \leq \bar{X} + Z_{\alpha/2}\sqrt{\frac{1}{4}\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right) + \frac{s_3^2}{n_3}}$$

b) A 95% confidence interval for μ is

$$\left(\frac{1}{2}(4.6+5.2)-6.1\right)-1.96\sqrt{\frac{1}{4}\left(\frac{0.7^2}{100}+\frac{0.6^2}{120}\right)+\frac{0.8^2}{130}} \le \mu \le \left(\frac{1}{2}(4.6+5.2)-6.1\right)+1.96\sqrt{\frac{1}{4}\left(\frac{0.7^2}{100}+\frac{0.6^2}{120}\right)+\frac{0.8^2}{130}}$$

$$-1.2-0.163 \le \mu \le -1.2+0.163$$

$$-1.363 \le \mu \le -1.037$$

Since zero is not contained in this interval, and because the possible differences (-1.363, -1.037) are negative, we can conclude that there is sufficient evidence to indicate that pesticide three is more effective.

9-83. The $V(\overline{X}_1-\overline{X}_2)=\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}$ and suppose this is to equal a constant k. Then, we are to minimize $C_1n_1+C_2n_2$ subject to $\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}=k$. Using a Lagrange multiplier, we minimize by setting the partial derivatives of $f(n_1,n_2,\lambda)=C_1n_1+C_2n_2+\lambda\left(\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}-k\right)$ with respect to n_1, n_2 and λ equal to zero.

These equations are

$$\frac{\partial}{\partial n_1}f(n_1,n_2,\lambda)=C_1-\frac{\lambda\sigma_1^2}{n_1^2}=0 \qquad (1)$$

$$\frac{\partial}{\partial n_2}f(n_1,n_2,\lambda)=C_2-\frac{\lambda\sigma_2^2}{n_2^2}=0 \qquad (2)$$

$$\frac{\partial}{\partial \lambda}f(n_1,n_2,\lambda)=\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}=k \qquad (3)$$

Upon adding equations (1) and (2), we obtain $C_1+C_2-\lambda\left(\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}\right)=0$

Substituting from equation (3) enables us to solve for λ to obtain $\frac{C_1+C_2}{k}=\lambda$

Then, equations (1) and (2) are solved for n_1 and n_2 to obtain

$$n_1=\frac{\sigma_1^2(C_1+C_2)}{kC_1} \qquad n_2=\frac{\sigma_2^2(C_1+C_2)}{kC_2}$$

It can be verified that this is a minimum and that with these choices for n_1 and n_2.

$$V(\overline{X}_1-\overline{X}_2)=\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}.$$

9-84. Maximizing the probability of rejecting H_0 is equivalent to minimizing

$$P\left(-z_{\alpha/2}<\frac{\overline{x}_1-\overline{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}<z_{\alpha/2}\,|\,\mu_1-\mu_2=\delta\right)=P\left(-z_{\alpha/2}-\frac{\delta}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}<Z<z_{\alpha/2}-\frac{\delta}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}\right)$$

where z is a standard normal random variable. This probability is minimized by maximizing $\frac{\delta}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}$.

Therefore, we are to minimize $\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$ subject to $n_1 + n_2 = N$.

From the constraint, $n_2 = N - n_1$, and we are to minimize $f(n_1) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{N - n_1}$. Taking the derivative of $f(n_1)$ with respect to n_1 and setting it equal to zero results in the equation $\frac{-\sigma_1^2}{n_1^2} + \frac{\sigma_2^2}{(N - n_1)^2} = 0$.

Upon solving for n_1, we obtain $n_1 = \frac{\sigma_1 N}{\sigma_1 + \sigma_2}$ and $n_2 = \frac{\sigma_2 N}{\sigma_1 + \sigma_2}$

Also, it can be verified that the solution minimizes $f(n_1)$.

9-85. a) $\alpha = P(Z > z_\varepsilon \text{ or } Z < -z_{\alpha-\varepsilon})$ where Z has a standard normal distribution.

Then, $\alpha = P(Z > z_\varepsilon) + P(Z < -z_{\alpha-\varepsilon}) = \varepsilon + \alpha - \varepsilon = \alpha$

b) $\beta = P(-z_{\alpha-\varepsilon} < Z_0 < z_\varepsilon | \mu_1 = \mu_0 + \delta)$

$$\beta = P(-z_{\alpha-\varepsilon} < \frac{\bar{x} - \mu_0}{\sqrt{\sigma^2/n}} < z_\varepsilon | \mu_1 = \mu_0 + \delta)$$

$$= P(-z_{\alpha-\varepsilon} - \frac{\delta}{\sqrt{\sigma^2/n}} < Z < z_\varepsilon - \frac{\delta}{\sqrt{\sigma^2/n}})$$

$$= \Phi(z_\varepsilon - \frac{\delta}{\sqrt{\sigma^2/n}}) - \Phi(-z_{\alpha-\varepsilon} - \frac{\delta}{\sqrt{\sigma^2/n}})$$

9-86. The requested result can be obtained from data in which the pairs are very different. Example:

pair	1	2	3	4	5
sample 1	100	10	50	20	70
sample 2	110	20	59	31	80

$\bar{x}_1 = 50 \quad \bar{x}_2 = 60$

$s_1 = 36.74 \qquad s_2 = 36.54 \qquad s_{pooled} = 36.64$

Two-sample t-test : $t_0 = -0.43$ P-value = 0.68

$\bar{x}_d = -10 \qquad s_d = 0.707$

Paired t-test : $t_0 = -31.62$ P-value ≈ 0

9-87. $H_0: \sigma^2 = \sigma_0^2$

$H_1: \sigma^2 \neq \sigma_0^2$

$$\beta = P\left(\chi^2_{1-\alpha/2,n-1} < \frac{(n-1)S^2}{\sigma_0^2} < \chi^2_{\alpha/2,n-1} | \sigma^2 = \sigma_1^2 \neq \sigma_0^2\right)$$

$$= P\left(\frac{\sigma_0^2}{\sigma_1^2}\chi^2_{1-\alpha/2,n-1} < \frac{(n-1)S^2}{\sigma_1^2} < \frac{\sigma_0^2}{\sigma_1^2}\chi^2_{\alpha/2,n-1} | \sigma^2 = \sigma_1^2\right)$$

where $\frac{(n-1)S^2}{\sigma_1^2}$ has a chi-square distribution with n-1 degrees of freedom.

9-88. $H_0: \sigma_1^2 = \sigma_2^2$

$H_1: \sigma_1^2 \neq \sigma_2^2$

$$\beta = P\left(f^2_{1-\alpha/2, n_1-1, n_2-1} < \frac{S_1^2}{S_2^2} < f^2_{\alpha/2, n_1-1, n_2-1} \,\Big|\, \frac{\sigma_1^2}{\sigma_2^2} = \delta \neq 1 \right)$$

$$= P\left(\frac{\sigma_2^2}{\sigma_1^2} f_{1-\alpha/2, n_1-1, n_2-1} < \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} < \frac{\sigma_2^2}{\sigma_1^2} f_{\alpha/2, n_1-1, n_2-1} \,\Big|\, \frac{\sigma_1^2}{\sigma_2^2} = \delta \right)$$

where $\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}$ has an F distribution with $n_1 - 1$ and $n_2 - 1$ degrees of freedom.

CHAPTER 10

Section 10-2

10-1. a) $y_i = \beta_0 + \beta_1 x_1 + \varepsilon_i$

$S_{xx} = 157.42 - \frac{43^2}{14}$

$= 25.348571$

$S_{xy} = 1697.80 - \frac{43(572)}{14}$

$= -59.057143$

$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = \frac{-59.057143}{25.348571} = -2.330$

$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = \frac{572}{14} - (-2.3298017)(\frac{43}{14}) = 48.013$

b) $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$

$\hat{y} = 48.012962 - 2.3298017(4.3) = 37.99$

c) $\hat{y} = 48.012962 - 2.3298017(3.7) = 39.39$

d) $e_i = y_i - \hat{y}_i = 46.1 - 39.39 = 6.71$

10-2. a) $y_0 = \beta_0 + \beta_1 x_1$

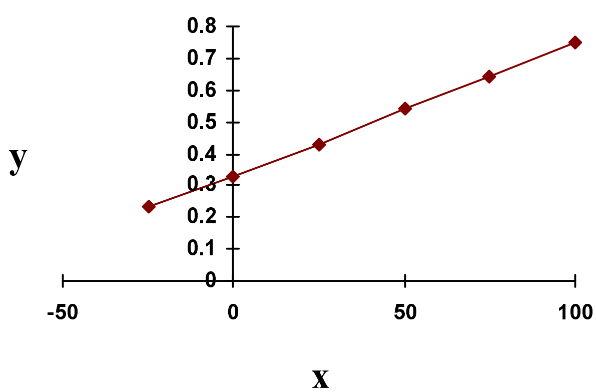
$S_{xx} = 143215.8 - \frac{1478^2}{20} = 33991.6$

$S_{xy} = 1083.67 - \frac{(1478)(12.75)}{20} = 141.445$

$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = \frac{141.445}{33991.6} = 0.0041612$

$\hat{\beta}_0 = \frac{12.75}{20} - (0.0041612)(\frac{1478}{20}) = 0.3299892$

$\hat{y} = 0.3299892 + 0.0041612x$



b) $\hat{y} = 0.3299892 + 0.0041612(85) = 0.683689$

c) $\hat{y} = 0.3299892 + 0.0041612(90) = 0.7044949$

d) $\hat{\beta}_1 = 0.00416$

10-3. a) $\hat{y} = 0.3299892 + 0.0041612(\frac{9}{5}x + 32)$

$\hat{y} = 0.3299892 + 0.0074902x + 0.1331584$

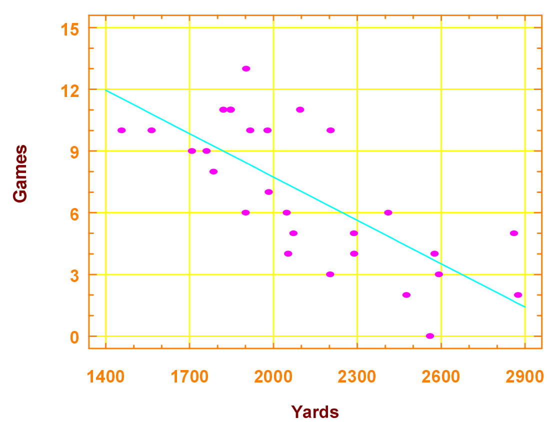
$\hat{y} = 0.4631476 + 0.0074902x$

b) $\hat{\beta}_1 = 0.00749$

10-4. a)

```
Regression Analysis - Linear model: Y = a+bX
Dependent variable: Games                          Independent variable: Yards
--------------------------------------------------------------------------------
                                Standard            T             Prob.
Parameter        Estimate        Error            Value           Level
Intercept         21.7883       2.69623           8.081          .00000
Slope          -7.0251E-3     1.25965E-3       -5.57703          .00001
--------------------------------------------------------------------------------
                            Analysis of Variance
Source              Sum of Squares     Df  Mean Square    F-Ratio  Prob. Level
Model                    178.09231      1    178.09231    31.1032       .00001
Residual                 148.87197     26      5.72585
--------------------------------------------------------------------------------
Total (Corr.)            326.96429     27
Correlation Coefficient = -0.738027          R-squared =  54.47 percent
Stnd. Error of Est. = 2.39287
```

Regression of Games on Yards



If the calculations were to be done by hand use Equations (10-7) and (10-8).

b) $\hat{y} = 21.7883 - 0.0070251(1800) = 9.143$

c) $-0.0070251(-100) = 0.70251$ games won.

d) $\dfrac{1}{0.0070251} = 142.35$ yds decrease required.

e) $\hat{y} = 21.7883 - 0.0070251(1917) = 8.321$

$e_i = y_i - \hat{y}_i$

$= 10 - 8.321 = 1.679$

10-5. a)

```
Regression Analysis - Linear model: Y = a+bX
Dependent variable: SalePrice                      Independent variable: Taxes
--------------------------------------------------------------------------------
                                Standard            T             Prob.
Parameter        Estimate        Error            Value           Level
Intercept         13.3202       2.57172         5.17948          .00003
Slope             3.32437      0.390276           8.518          .00000
--------------------------------------------------------------------------------
                            Analysis of Variance
Source              Sum of Squares     Df  Mean Square    F-Ratio  Prob. Level
Model                    636.15569      1    636.15569    72.5563       .00000
Residual                 192.89056     22      8.76775
--------------------------------------------------------------------------------
Total (Corr.)            829.04625     23
Correlation Coefficient = 0.875976           R-squared =  76.73 percent
Stnd. Error of Est. = 2.96104
```

If the calculations were to be done by hand use Equations (10-7) and (10-8).

$\hat{y} = 13.3202 + 3.32437x$

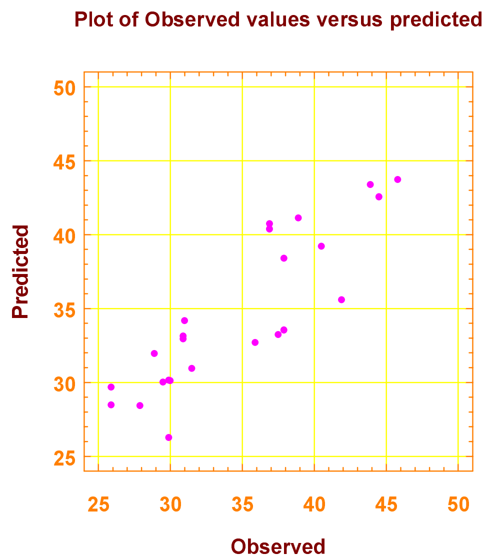
b) $\hat{y} = 13.3202 + 3.32437(7.5) = 38.253$

c) $\hat{y} = 13.3202 + 3.32437(5.8980) = 32.9273$

$\hat{y} = 32.9273$

$e = y - \hat{y} = -2.02732$

d) All the points would lie along the 45% axis line. That is, the regression model would estimate the values exactly. At this point, the graph of observed vs. predicted looks like a reasonable fit.



10-6. a)

```
Regression Analysis - Linear model: Y = a+bX
Dependent variable: Usage                  Independent variable: Temperature
--------------------------------------------------------------------------------
                               Standard              T             Prob.
Parameter       Estimate         Error             Value           Level
Intercept        -6.3355        1.66765          -3.79906         .00349
Slope            9.20836       0.0337744          272.643         .00000
--------------------------------------------------------------------------------
                           Analysis of Variance
Source              Sum of Squares     Df  Mean Square    F-Ratio  Prob. Level
Model                    280583.12      1    280583.12    74334.4        .00000
Residual                 37.746089     10     3.774609
--------------------------------------------------------------------------------
Total (Corr.)            280620.87     11
Correlation Coefficient = 0.999933          R-squared =  99.99 percent
Stnd. Error of Est. = 1.94284
```

If the calculations were to be done by hand use Equations (10-7) and (10-8).

$\hat{y} = -6.3355 + 9.20836x$

b) $\hat{y} = -6.3355 + 9.20836(55) = 500.124$

c) If monthly temperature increases by $1^{\circ}F$, $\hat{y}$ increases by 9.20836.

d) $\hat{y} = -6.3355 + 9.20836(47) = 426.458$

$\hat{y} = 426.458$

$e = y - \hat{y} = 424.84 - 426.458 = -1.618$

10-7.

```
Predictor        Coef       StDev          T        P
Constant       33.535       2.614      12.83    0.000
x            -0.03540     0.01663      -2.13    0.047
S = 3.660      R-Sq = 20.1%     R-Sq(adj) = 15.7%
Analysis of Variance
Source        DF          SS          MS          F        P
Regression     1       60.69       60.69       4.53    0.047
Error         18      241.06       13.39
Total         19      301.75
```

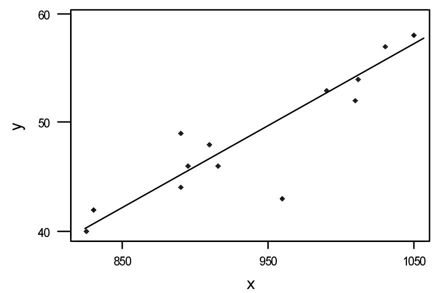
a) $\hat{y} = 33.5348 - 0.0353971x$

b) $\hat{y} = 33.5348 - 0.0353971(150) = 28.26$

c) $\hat{y} = 29.4995$

$e = y - \hat{y} = 31.0 - 29.4995 = 1.50048$

10-8. a)



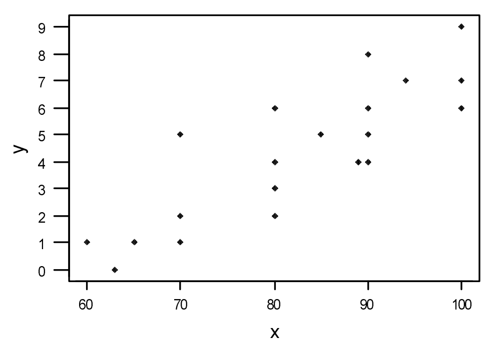

```
Predictor       Coef       StDev          T        P
Constant     -16.509       9.843      -1.68    0.122
x            0.06936     0.01045       6.64    0.000
S = 2.706       R-Sq = 80.0%     R-Sq(adj) = 78.2%
Analysis of Variance
Source       DF          SS          MS          F        P
Regression    1      322.50      322.50      44.03    0.000
Error        11       80.57        7.32
Total        12      403.08
```

$\hat{y} = -16.5093 + 0.0693554x$

b) $\hat{y} = 46.6041$ $e = y - \hat{y} = 1.39592$

c) $\hat{y} = -16.5093 + 0.0693554(950) = 49.38$

10-9. a)



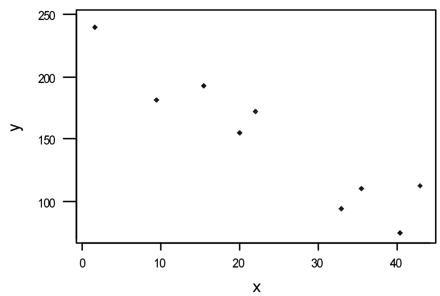
Yes, a linear regression would seem appropriate, but one or two points appear to be outliers.

```
Predictor       Coef       StDev          T        P
Constant      -9.813       2.135      -4.60    0.000
x            0.17148     0.02566       6.68    0.000
S = 1.408       R-Sq = 71.3%     R-Sq(adj) = 69.7%
Analysis of Variance
Source       DF          SS          MS          F        P
Regression    1      88.520      88.520      44.66    0.000
Error        18      35.680       1.982
Total        19     124.200
```

b) $\hat{y} = -9.8131 + 0.171484x$

c) $\hat{y} = 4.76301$ at x = 85

10-10. a)



Yes

Predictor	Coef	StDev	T	P
Constant	234.07	13.75	17.03	0.000
x	-3.5086	0.4911	-7.14	0.000

S = 19.96 R-Sq = 87.9% R-Sq(adj) = 86.2%

Analysis of Variance

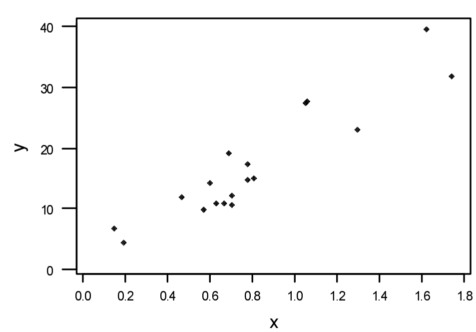
Source	DF	SS	MS	F	P
Regression	1	20329	20329	51.04	0.000
Error	7	2788	398		
Total	8	23117			

b) $\hat{y} = 234.071 - 3.50856x$

c) $\hat{y} = 234.071 - 3.50856(30) = 128.814$

d) $\hat{y} = 156.883 \quad e = 15.1175$

10-11. a)



Yes

Predictor	Coef	StDev	T	P
Constant	0.470	1.936	0.24	0.811
x	20.567	2.142	9.60	0.000

S = 3.716 R-Sq = 85.2% R-Sq(adj) = 84.3%

Analysis of Variance

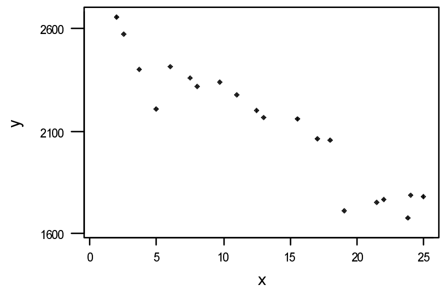
Source	DF	SS	MS	F	P
Regression	1	1273.5	1273.5	92.22	0.000
Error	16	220.9	13.8		
Total	17	1494.5			

b) $\hat{y} = 0.470467 + 20.5673x$

c) $\hat{y} = 0.470467 + 20.5673(1) = 21.038$

d) $\hat{y} = 10.1371 \quad e = 1.6629$

10-12. a)



Yes

Predictor	Coef	StDev	T	P
Constant	2625.39	45.35	57.90	0.000
x	-36.962	2.967	-12.46	0.000

S = 99.05   R-Sq = 89.6%   R-Sq(adj) = 89.0%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	1522819	1522819	155.21	0.000
Error	18	176602	9811		
Total	19	1699421			

b) $\hat{y} = 2625.39 - 36.9618x$

c) $\hat{y} = 2625.39 - 36.9618(20) = 1886.154$

d) If there were no error, the values would all lie along the 45° axis. The plot indicates age was reasonable regressor variable.

10-13. $\hat{\beta}_0 + \hat{\beta}_1\bar{x} = (\bar{y} - \hat{\beta}_1\bar{x}) + \hat{\beta}_1\bar{x} = \bar{y}$

10-14. a) The slopes of both regression models will be the same, but the intercept will be shifted.

b) $\hat{y} = 2132.41 - 36.9618x$

$\hat{\beta}_0 = 2625.39$	vs.	$\hat{\beta}_0^* = 2132.41$
$\hat{\beta}_1 = -36.9618$		$\hat{\beta}_1^* = -36.9618$

10-15. Let $x_i^* = x_i - \bar{x}$. Then, the model is $Y_i^* = \beta_0^* + \beta_1^* x_i^* + \varepsilon_i$.

Equations 10-7 and 10-8 can be applied to the new variables using the facts that $\sum_{i=1}^{n} x_i^* = \sum_{i=1}^{n} y_i^* = 0$. Then,

$\hat{\beta}_1^* = \hat{\beta}_1$ and $\hat{\beta}_0^* = 0$.

10-16. The least squares estimate minimizes $\sum (y_i - \beta x_i)^2$. Upon setting the derivative equal to zero, we obtain

$2\sum (y_i - \beta x_i)\ (-x_i) = 2[\sum y_i x_i - \beta \sum x_i^2] = 0$

Therefore, $\hat{\beta} = \dfrac{\sum y_i x_i}{\sum x_i^2}$.

10-17. $\hat{y} = 21.031461x$. The model seems very appropriate - an even better fit.

Section 10-5

10-18. a) 1) The parameter of interest is the regressor variable coefficient, β_1

2) $H_0 : \beta_1 = 0$

3) $H_1 : \beta_1 \neq 0$

4) $\alpha = 0.05$

5) The test statistic is

$$f_0 = \frac{MS_R}{MS_E} = \frac{SS_R / 1}{SS_E / (n-2)}$$

6) Reject H_0 if $f_0 > f_{\alpha,1,12}$ where $f_{0.05,1,12} = 4.75$

7) Using results from Exercise 6-1

$$SS_R = \hat{\beta}_1 S_{xy} = -2.3298017(-59.057143)$$
$$= 137.59$$
$$SS_E = S_{yy} - SS_R$$
$$= 159.71429 - 137.59143$$
$$= 22.123$$

$$f_0 = \frac{137.59}{22.123 / 12} = 74.63$$

8) Since 74.63 > 4.75 reject H_0 and conclude that compressive strength is a significant in predicting intrinsic permeability of concrete at $\alpha = 0.05$. We can therefore conclude model specifies a useful linear relationship between these two variables.

P − value ≅ 0.000002

b) $\hat{\sigma}^2 = MS_E = \frac{SS_E}{n-2} = \frac{22.123}{12} = 1.844$

10-19. a) 1) The parameter of interest is the regressor variable coefficient, β_1.

2) $H_0 : \beta_1 = 0$

3) $H_1 : \beta_1 \neq 0$

4) $\alpha = 0.05$

5) The test statistic is

$$f_0 = \frac{MS_R}{MS_E} = \frac{SS_R / 1}{SS_E / (n-2)}$$

6) Reject H_0 if $f_0 > f_{\alpha,1,18}$ where $f_{0.05,1,18} = 4.416$

7) Using the results from Exercise 6-2

$$SS_R = \hat{\beta}_1 S_{xy} = (0.0041612)(141.445)$$
$$= 0.5886$$
$$SS_E = S_{yy} - SS_R$$
$$= (8.86 - \tfrac{12.75^2}{20}) - 0.5886$$
$$= 0.143275$$

$$f_0 = \frac{0.5886}{0.143275 / 18} = 73.95$$

8) Since 73.95 > 4.416, reject H_0 and conclude the model specifies a useful relationship at $\alpha = 0.05$.

P − value ≅ 0.000001

b) $\hat{\sigma}^2 = MS_E = \frac{SS_E}{n-2} = \frac{0.143275}{18} = 0.00796$

10-20. a) Refer to ANOVA table of Exercise 10-4.

1) The parameter of interest is the regressor variable coefficient, β_1.

2) $H_0:\beta_1 = 0$

3) $H_1:\beta_1 \neq 0$

4) $\alpha = 0.01$

5) The test statistic is

$$f_0 = \frac{MS_R}{MS_E} = \frac{SS_R / 1}{SS_E / (n-2)}$$

6) Reject H_0 if $f_0 > f_{\alpha,1,26}$ where $f_{0.01,1,26} = 7.724$

7) Using the results of Exercise 10-4

$$f_0 = \frac{MS_R}{MS_E} = 31.1032$$

8) Since 31.1032 > 7.724 reject H_0 and conclude the model is useful at $\alpha = 0.01$. P – value = 0.000007

b) $\hat{\sigma}^2 = MS_E = 5.72585$

c) 2) $H_0:\beta_1 = 0$

3) $H_1:\beta_1 \neq 0$

4) $\alpha = 0.01$

5) The test statistic is $t_0 = \dfrac{\hat{\beta}_1}{se(\hat{\beta}_1)}$

6) Reject H_0 if $t_0 < -t_{\alpha/2,n-2}$ where $-t_{0.005,26} = -2.78$ or $t_0 > t_{0.005,26} = 2.78$

7) Using the results from Exercise 10-4

$$t_0 = \frac{-0.0070251}{0.00125965} = -5.577$$

8) Since −5.577 < −2.78 reject H_0 and conclude the regressor is useful in the model at $\alpha = 0.01$.

10-21. Refer to ANOVA of Exercise 10-5

a) 1) The parameter of interest is the regressor variable coefficient, β_1.

2) $H_0:\beta_1 = 0$

3) $H_1:\beta_1 \neq 0$

4) $\alpha = 0.05$, using t-test

5) The test statistic is $t_0 = \dfrac{\hat{\beta}_1}{se(\hat{\beta}_1)}$

6) Reject H_0 if $t_0 < -t_{\alpha/2,n-2}$ where $-t_{0.025,22} = -2.074$ or $t_0 > t_{0.025,22} = 2.074$

7) Using the results from Exercise 10-5

$$t_0 = \frac{3.32437}{0.390276} = 8.518$$

8) Since 8.518 > 2.074 reject H_0 and conclude the model is useful $\alpha = 0.05$.

b) 2) $H_0:\beta_1 = 0$

3) $H_1:\beta_1 \neq 0$

4) $\alpha = 0.05$

5) The test statistic is $f_0 = \dfrac{MS_R}{MS_E} = \dfrac{SS_R / 1}{SS_E / (n-2)}$

6) Reject H_0 if $f_0 > f_{\alpha,1,22}$ where $f_{0.01,1,22} = 4.303$

7) Using the results from Exercise 10-5

$$f_0 = \frac{636.15569 / 1}{192.89056 / 22} = 72.5563$$

8) Since 72.5563 > 4.303, reject H_0 and conclude the model is useful $\alpha = 0.05$.

The F-test is the t-test squared, but it is restricted to a two-sided test.

c) $\hat{\sigma}^2 = 8.76775$

10-22. Refer to ANOVA for Exercise 10-6

a) 1) The parameter of interest is the regressor variable coefficient, β_1.

2) $H_0: \beta_1 = 0$

3) $H_1: \beta_1 \neq 0$

4) $\alpha = 0.01$

5) The test statistic is $f_0 = \dfrac{MS_R}{MS_E} = \dfrac{SS_R / 1}{SS_E / (n-2)}$

6) Reject H_0 if $f_0 > f_{\alpha,1,22}$ where $f_{0.01,1,10} = 10.049$

7) Using the results from Exercise 10-6

$$f_0 = \frac{280583.12 / 1}{37.746089 / 10} = 74334.4$$

8) Since 74334.4 > 10.049, reject H_0 and conclude the model is useful $\alpha = 0.01$. P-value < 0.000001

b) $MSE = SS_E/(n-p) = \hat{\sigma}^2 = 3.77461$

c) 1) The parameter of interest is the regressor variable coefficient, β_1.

2) $H_0: \beta_1 = 10$

3) $H_1: \beta_1 \neq 10$

4) $\alpha = 0.01$

5) The test statistic is $t_0 = \dfrac{\hat{\beta}_1 - \beta_{1,0}}{se(\hat{\beta}_1)}$

6) Reject H_0 if $t_0 < -t_{\alpha/2,n-2}$ where $-t_{0.005,10} = -3.17$ or $t_0 > t_{0.005,10} = 3.17$

7) Using the results from Exercise 10-6

$$t_0 = \frac{9.21 - 10}{0.0338} = -23.37$$

8) Since $-23.37 < -3.17$ reject H_0 and conclude the regressor is useful at $\alpha = 0.01$. P-value = 0.

10-23. Refer to ANOVA table of Exercise 10-7

a) $H_0: \beta_1 = 0$

$H_1: \beta_1 \neq 0 \quad \alpha = 0.01$

$f_0 = 4.53158$

$f_{.01,1,18} = 8.289$

$f_0 \not> f_{\alpha,1,18}$

Therefore, do not reject H_0. P-value = 0.04734. Insufficient evidence to conclude that the model is a useful relationship.

b) $\hat{\sigma} = \sqrt{MS_E} = \sqrt{13.39232} = 3.6596$

$se(\hat{\beta}_1) = 0.0166281$

$se(\hat{\beta}_0) = 2.61396$

c) $H_0: \beta_1 = -0.05$

$H_1: \beta_1 < -0.05 \quad \alpha = 0.01$

$$t_0 = \frac{-0.0354 - (-0.05)}{0.0166281} = 0.87803$$

$t_{.01,18} = 2.552$

$t_0 \not< -t_{\alpha,18}$

Therefore, do not reject H_0. P-value = 0.804251. Insufficient evidence to conclude that β_1 is ≥ -0.05.

d) $H_0: \beta_0 = 0 \quad H_1: \beta_0 \neq 0 \quad \alpha = 0.01$

$t_0 = 12.8291$

$t_{.005,18} = 2.878$

$t_0 > t_{\alpha/2,18}$

Therefore, reject H_0. P-value = 0

10-24. Refer to ANOVA of Exercise 10-8

a) $H_0 : \beta_1 = 0$

$H_1 : \beta_1 \neq 0$

$\alpha = 0.05$

$f_0 = 44.0279$

$f_{.05,1,11} = 4.84$

$f_0 > f_{\alpha,1,11}$

Therefore, rejectH_0. P-value = 0.00004.

b) $\hat{\sigma} = \sqrt{MS_E} = \sqrt{7.324951} = 2.7065$

$se(\hat{\beta}_1) = 0.0104524$

$se(\hat{\beta}_0) = 9.84346$

c) $H_0 : \beta_0 = 0$

$H_1 : \beta_0 \neq 0$

$\alpha = 0.05$

$t_0 = -1.67718$

$t_{.025,11} = 2.201$

$|t_0| \not< -t_{\alpha/2,11}$

Therefore, do not rejectH_0. P-value = 0.12166.

10-25. Refer to ANOVA of Exercise 10-9

a) $H_0 : \beta_1 = 0$

$H_1 : \beta_1 \neq 0$

$\alpha = 0.05$

$f_0 = 44.6567$

$f_{.05,1,18} = 4.416$

$f_0 > f_{\alpha,1,18}$

Therefore, rejectH_0. P-value = 0.000003.

b) $\hat{\sigma}^2 = 1.982231$

c) $se(\hat{\beta}_1) = 0.0256613$

$se(\hat{\beta}_0) = 2.13526$

d) $H_0 : \beta_0 = 0$

$H_1 : \beta_0 \neq 0$

$\alpha = 0.05$

$t_0 = -4.59573$

$t_{.025,18} = 2.101$

$|t_0| > t_{\alpha/2,18}$

Therefore, reject H_0. P-value = 0.00022.

10-26. Refer to ANOVA of Exercise 10-11

a) $H_0 : \beta_1 = 0$

$H_1 : \beta_1 \neq 0$

$\alpha = 0.01$

$f_0 = 92.224$

$f_{.01,1,16} = 8.534$

$f_0 > f_{\alpha,1,16}$

Therefore, reject H_0.

b) P-value < 0.00001

c) $\hat{\sigma}^2 = 13.80920$

$se(\hat{\beta}_1) = 2.14169$

$se(\hat{\beta}_0) = 1.93591$

d) $H_0 : \beta_0 = 0$

$H_1 : \beta_0 \neq 0$

$\alpha = 0.01$

$t_0 = 0.243$

$t_{.005,16} = 2.921$

$t_0 \not> t_{\alpha/2,16}$

Therefore, do not rejectH_0. Conclude, Yes, the intercept should be removed.

10-27. Refer to ANOVA of Exercise 10-12

a) $H_0 : \beta_1 = 0$

$H_1 : \beta_1 \neq 0$

$\alpha = 0.01$

$f_0 = 155$

$f_{.01,1,18} = 8.289$

$f_0 > f_{\alpha,1,18}$

Therefore, rejectH_0. P-value < 0.00001.

b) $\hat{\sigma}^2 = 9811.21$

$se(\hat{\beta}_1) = 45.3468$

$se(\hat{\beta}_0) = 2.96681$

c) $H_0 : \beta_1 = -30$

$H_1 : \beta_1 \neq -30$

$\alpha = 0.01$

$$t_0 = \frac{-36.9618 - (-30)}{2.96681} = -2.3466$$

$t_{.005,18} = 2.878$

$|t_0| \not> -t_{\alpha/2,18}$

Therefore, do not rejectH_0. P-value = 0.0153(2) = 0.0306.

d) $H_0 : \beta_0 = 0$

$H_1 : \beta_0 \neq 0$

$\alpha = 0.01$

$t_0 = 57.8957$

$t_{.005,18} = 2.878$

$t_0 > t_{\alpha/2,18}$, therefore, rejectH_0. P-value < 0.00001.

e) $H_0 : \beta_0 = 2500$

$H_1 : \beta_0 \neq 2500$

$\alpha = 0.01$

$$t_0 = \frac{2625.39 - 2500}{45.3468} = 2.7651$$

$t_{.005,18} = 2.878$

$t_0 < t_{\alpha/2,18}$, therefore, do not reject H_0. P-value = 0.0064(2) = 0.0128.

10-28. $t_0 = \frac{\hat{\beta}_1}{\sqrt{\hat{\sigma}^2 / S_{xx}}}$ After the transformation $\hat{\beta}_1^* = \frac{b}{a}\hat{\beta}_1$, $S_{xx}^* = a^2 S_{xx}$, $\bar{x}^* = a\bar{x}$, $\hat{\beta}_0^* = b\hat{\beta}_0$, and

$\hat{\sigma}^* = b\hat{\sigma}$. Therefore, $t_0^* = \frac{b\hat{\beta}_1 / a}{\sqrt{(b\hat{\sigma})^2 / a^2 S_{xx}}} = t_0$.

10-29. a) $\frac{\hat{\beta}}{\sqrt{\frac{\hat{\sigma}^2}{\sum x_i^2}}}$ has a t distribution with n-1 degree of freedom.

b) From Exercise 10-17, $\hat{\beta} = 21.031461, \hat{\sigma} = 3.611768,$ and $\sum x_i^2 = 14.7073$.

The t-statistic in part a. is 22.3314 and $H_0: \beta_0 = 0$ is rejected at usual α values.

10-30. $d = \frac{|-0.01 - (-0.005)|}{2.4\sqrt{\frac{27}{3608611.96}}} = 0.76$, $S_{xx} = 3608611.96$.

Assume $\alpha = 0.05$, from Chart VI and interpolating between the curves for n = 20 and n = 30, $\beta \cong 0.05$.

Sections 10-6 and 10-7

10-31. $t_{\alpha/2,n-2} = t_{0.025,12} = 2.179$

a) 95% confidence interval on β_1.

$\hat{\beta}_1 \pm t_{\alpha/2,n-2} se(\hat{\beta}_1)$

$-2.3298017 \pm t_{.025,12}(0.2697)$

$-2.3298017 \pm 2.179(0.2697)$

$-2.917478 \le \beta_1 \le -1.7421254$

b) 95% confidence interval on β_0.

$\hat{\beta}_0 \pm t_{.025,12} se(\hat{\beta}_0)$

$48.012962 \pm 2.179(0.5959)$

$46.714496 \le \beta_0 \le 49.311428$

c) 95% confidence interval on μ when $x_0 = 2.5$.

$\hat{\mu}_{Y|x_0} = 48.012962 - 2.3298017(2.5) = 42.188458$

$\hat{\mu}_{Y|x_0} \pm t_{.025,12}\sqrt{\hat{\sigma}^2(\frac{1}{n} + \frac{(x_0-\bar{x})^2}{S_{xx}})}$

$42.188458 \pm (2.179)\sqrt{1.844(\frac{1}{14} + \frac{(2.5-1.7525)^2}{25.348571})}$

42.188458 ± 0.904643

$41.283815 \le \hat{\mu}_{Y|x_0} \le 43.093101$

d) 95% on prediction interval when $x_0 = 2.5$.

$\hat{y}_0 \pm t_{.025,12}\sqrt{\hat{\sigma}^2(1 + \frac{1}{n} + \frac{(x_0-\bar{x})^2}{S_{xx}})}$

$42.188458 \pm 2.179\sqrt{1.844(1 + \frac{1}{14} + \frac{(2.5-1.7525)^2}{25.348571})}$

42.188458 ± 3.0941504

$39.094307 \le y_0 \le 45.282608$

It is wider, because it depends on both the error associated with the fitted model as well as that with the future observation.

10-32. $t_{\alpha/2,n-2} = t_{0.005,18} = 2.878$

a) $\hat{\beta}_1 \pm \left(t_{0.005,18}\right)se(\hat{\beta}_1)$.

$0.0041612 \pm (2.878)(0.000484)$

$0.0027682 \le \beta_1 \le 0.0055542$

b) $\hat{\beta}_0 \pm \left(t_{0.005,18}\right)se(\hat{\beta}_0)$.

$0.3299892 \pm (2.878)(0.04095)$

$0.2121351 \le \beta_0 \le 0.4478433$

c) 99% confidence interval on μ when $x_0 = 85^\circ F$.

$\hat{\mu}_{Y|x_0} = 0.683689$

$\hat{\mu}_{Y|x_0} \pm t_{.005,18}\sqrt{\hat{\sigma}^2(\frac{1}{n} + \frac{(x_0-\bar{x})^2}{S_{xx}})}$

$0.683689 \pm (2.878)\sqrt{0.00796(\frac{1}{20} + \frac{(85-73.9)^2}{33991.6})}$

0.683689 ± 0.0594607

$0.6242283 \le \hat{\mu}_{Y|x_0} \le 0.7431497$

d) 99% prediction interval when $x_0 = 90^\circ F$.

$\hat{y}_0 = 0.7044949$

$\hat{y}_0 \pm t_{.005,18}\sqrt{\hat{\sigma}^2(1 + \frac{1}{n} + \frac{(x_0-\bar{x})^2}{S_{xx}})}$

$0.7044949 \pm 2.878\sqrt{0.00796(1 + \frac{1}{20} + \frac{(90-73.9)^2}{33991.6})}$

0.7044949 ± 0.2640665

$0.4404284 \le y_0 \le 0.9685614$

Note for Problems 10-33 through 10-35: These computer printouts were obtained from Statgraphics. For Minitab users, the standard errors are obtained from the Regression subroutine.

10-33.

95 percent confidence intervals for coefficient estimates

	Estimate	Standard error	Lower Limit	Upper Limit
CONSTANT	21.7883	2.69623	16.2448	27.3318
Yards	-0.00703	0.00126	-0.00961	-0.00444

a) $-0.00961 \le \beta_1 \le -0.00444$.

$16.2448 \le \beta_0 \le 27.3318$.

b) $9.143 \pm (2.056)\sqrt{5.72585(\frac{1}{28} + \frac{(1800-2110.14)^2}{3608325.5})}$

9.143 ± 1.2286697

$7.91433 \le \hat{\mu}_{Y|x_0} \le 10.37167$

c) $9.143 \pm (2.056)\sqrt{5.72585(1 + \frac{1}{28} + \frac{(1800-2110.14)^2}{3608325.5})}$

9.143 ± 5.07086

$4.07214 \le y_0 \le 14.21386$

10-34.

95 percent confidence intervals for coefficient estimates

	Estimate	Standard error	Lower Limit	Upper Limit
CONSTANT	13.3202	2.57172	7.98547	18.6549
Taxes	3.32437	0.39028	2.51479	4.13395

a) $2.51479 \le \beta_1 \le 4.13395$.
$7.98547 \le \beta_0 \le 18.6549$.

b) $38.253 \pm (2.074)\sqrt{8.76775(\frac{1}{24} + \frac{(7.5-6.40492)^2}{57.563139})}$

38.253 ± 1.5353

$36.7177 \le \hat{\mu}_{Y|x_0} \le 39.7883$

c) $38.253 \pm (2.074)\sqrt{8.76775(1 + \frac{1}{24} + \frac{(7.5-6.40492)^2}{57.563139})}$

38.253 ± 6.3302

$31.9228 \le y_0 \le 44.5832$

10-35.

95 percent confidence intervals for coefficient estimates

	Estimate	Standard error	Lower Limit	Upper Limit
CONSTANT	-6.33550	1.66765	-10.0522	-2.61876
Temperature	9.20836	0.03377	9.13309	9.28364

a) $9.10130 \le \beta_1 \le 9.31543$
b) $-11.6219 \le \beta_0 \le -1.04911$

c) $500.124 \pm (2.228)\sqrt{3.774609(\frac{1}{12} + \frac{(55-46.5)^2}{3308.9994})}$

500.124 ± 1.4037586

$498.72024 \le \hat{\mu}_{Y|x_0} \le 501.52776$

d) $500.124 \pm (2.228)\sqrt{3.774609(1 + \frac{1}{12} + \frac{(55-46.5)^2}{3308.9994})}$

500.124 ± 4.5505644

$495.57344 \le y_0 \le 04.67456$

It is wider because more error is included - both from the fitted model and from that associated with the future observation.

10-36. a) (-0.07034, -0.00045)
b) (28.0417, 39.0279)

c) $28.26 \pm (2.101)\sqrt{13.39232(\frac{1}{20} + \frac{(150-149.3)^2}{48436.256})}$

28.26 ± 1.7194236

$(26.5406, 29.9794)$

d) $28.26 \pm (2.101)\sqrt{13.39232(1 + \frac{1}{20} + \frac{(150-149.3)^2}{48436.256})}$

28.26 ± 7.87863

$(20.3814, 36.1386)$

10-37. a) (0.03689, 0.10183) b) (-47.0877, 14.0691)

c) $46.6041 \pm (3.106)\sqrt{7.324951\left(\frac{1}{13} + \frac{(910-939)^2}{67045.97}\right)}$

46.6041 ± 2.514401

$(44.0897, 49.1185)$

d) $46.6041 \pm (3.106)\sqrt{7.324951\left(1 + \frac{1}{13} + \frac{(910-939)^2}{67045.97}\right)}$

46.6041 ± 8.779266

$(37.8298, 55.3784)$

10-38. a) (0.11756, 0.22541) b) (-14.3002, -5.32598)

c) $4.76301 \pm (2.101)\sqrt{1.982231\left(\frac{1}{20} + \frac{(85-82.3)^2}{3010.2111}\right)}$

4.76301 ± 0.6772655

$(4.0857, 5.4403)$

d) $4.76301 \pm (2.101)\sqrt{1.982231\left(1 + \frac{1}{20} + \frac{(85-82.3)^2}{3010.2111}\right)}$

4.76301 ± 3.0345765

$(1.7284, 7.7976)$

10-39. a) (201.552, 266.590) b) (-4.67015, -2.34696)

c) $128.814 \pm (2.365)\sqrt{398.2804\left(\frac{1}{9} + \frac{(30-24.5)^2}{1651.4214}\right)}$

128.814 ± 16.980124

$(111.8339, 145.7941)$

10-40. a) (14.3107, 26.8239) b) (-5.18501, 6.12594)

c) $21.038 \pm (2.921)\sqrt{13.8092\left(\frac{1}{18} + \frac{(1-0.806111)^2}{3.01062}\right)}$

21.038 ± 2.8314277

$(18.2066, 23.8694)$

d) $21.038 \pm (2.921)\sqrt{13.8092\left(1 + \frac{1}{18} + \frac{(1-0.806111)^2}{3.01062}\right)}$

21.038 ± 11.217861

$(9.8201, 32.2559)$

10-41. a) (-43.1964, -30.7272) b) (2530.09, 2720.68)

c) $1886.154 \pm (2.101)\sqrt{9811.21\left(\frac{1}{20} + \frac{(20-13.3375)^2}{1114.6618}\right)}$

1886.154 ± 62.370688

$(1823.7833, 1948.5247)$

d) $1886.154 \pm (2.101)\sqrt{9811.21\left(1 + \frac{1}{20} + \frac{(20-13.3375)^2}{1114.6618}\right)}$

1886.154 ± 217.25275

$(1668.9013, 2103.4067)$

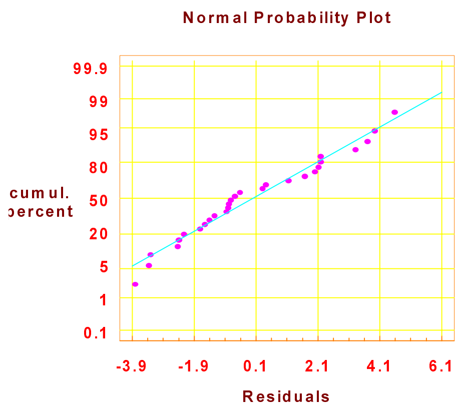
Section 10-8

10-42. Use the results of Exercise 10-4 to answer the following questions.

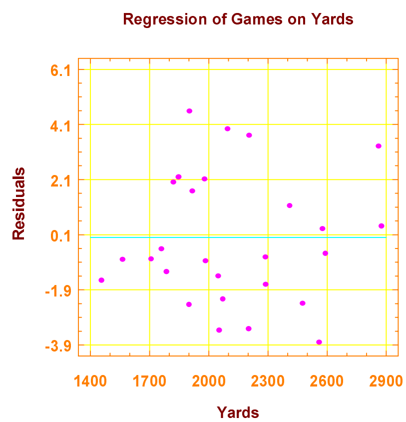
a) $R^2 = 0.544684$; The proportion of variability explained by the model.

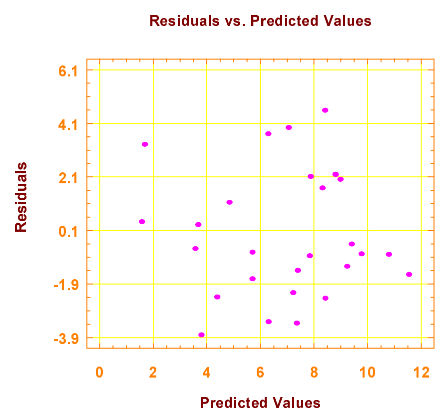
$$R^2_{Adj} = 1 - \frac{148.87197/26}{326.96429/27} = 1 - 0.473 = 0.527$$

b) Yes, normality seems to be satisfied since the data appear to fall along the straight line.



c) Since the residuals plots appear to be random, the plots do not include any serious model inadequacies.



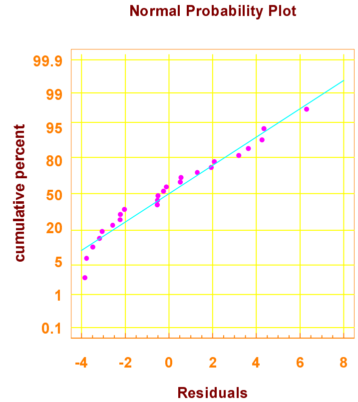


10-43. Use the Results of exercise 10-5 to answer the following questions.

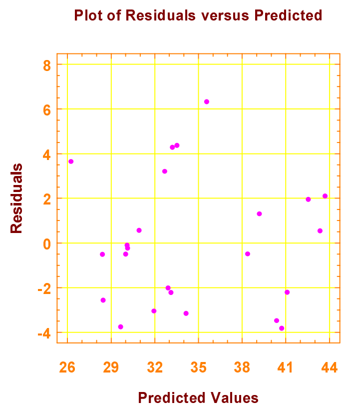
a)

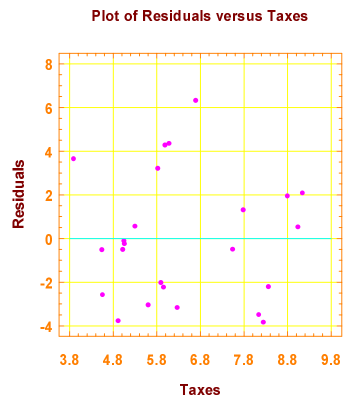
SalePrice	Taxes	Predicted	Residuals
25.9	4.9176	29.6681073	-3.76810726
29.5	5.0208	30.0111824	-0.51118237
27.9	4.5429	28.4224654	-0.52246536
25.9	4.5573	28.4703363	-2.57033630
29.9	5.0597	30.1405004	-0.24050041
29.9	3.8910	26.2553078	3.64469225
30.9	5.8980	32.9273208	-2.02732082
28.9	5.6039	31.9496232	-3.04962324
35.9	5.8282	32.6952797	3.20472030
31.5	5.3003	30.9403441	0.55965587
31.0	6.2712	34.1679762	-3.16797616
30.9	5.9592	33.1307723	-2.23077234
30.0	5.0500	30.1082540	-0.10825401
36.9	8.2464	40.7342742	-3.83427422
41.9	6.6969	35.5831610	6.31683901
40.5	7.7841	39.1974174	1.30258260
43.9	9.0384	43.3671762	0.53282376
37.5	5.9894	33.2311683	4.26883165
37.9	7.5422	38.3932520	-0.49325200
44.5	8.7951	42.5583567	1.94164328
37.9	6.0831	33.5426619	4.35733807
38.9	8.3607	41.1142499	-2.21424985
36.9	8.1400	40.3805611	-3.48056112
45.8	9.1416	43.7102513	2.08974865

b) Assumption of normality does not seem to be violated since the data appear to fall along a straight line.



c) No serious departure from assumption of constant variance. This is evident by the random pattern of the residuals.



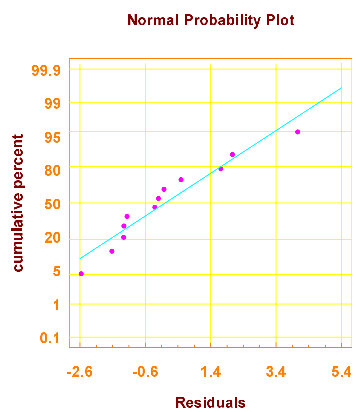


d) $R^2 \equiv 76.73\%$;

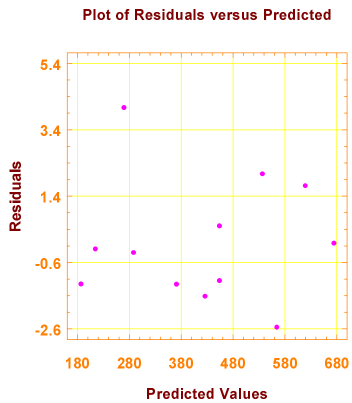
10-44. Use the results of Exercise 10-6 to answer the following questions

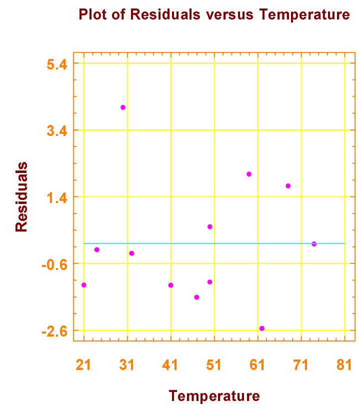
a) $R^2 = 99.986\%$; The proportion of variability explained by the model.

b) Yes, normality seems to be satisfied since the data appear to fall along the straight line.



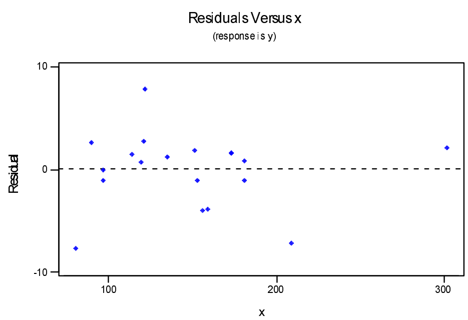
c) There might be lower variance at the middle settings of x. However, this data does not indicate a serious departure from the assumptions.

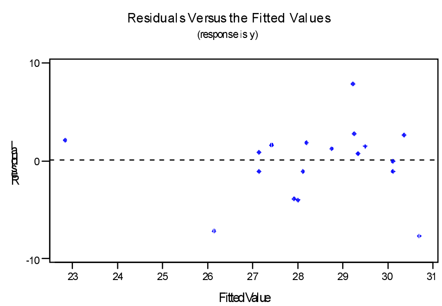




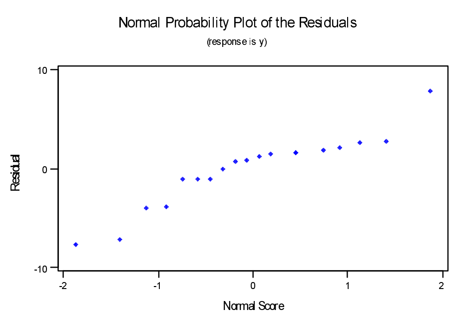
10-45. a) $R^2 = 20.1121\%$

b) These plots indicate presence of outliers, but no real problem with assumptions.

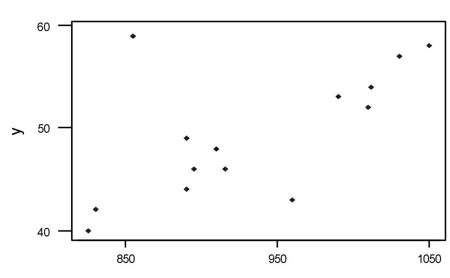




c) The normality assumption appears marginal.



10-46. a)



$\hat{y} = 0.677559 + 0.0521753x$

b) $H_0: \beta_1 = 0$ $H_1: \beta_1 \neq 0$ $\alpha = 0.05$

$f_0 = 7.9384$

$f_{.05,1,12} = 4.75$

$f_0 > f_{\alpha,1,12}$

Reject H_o.

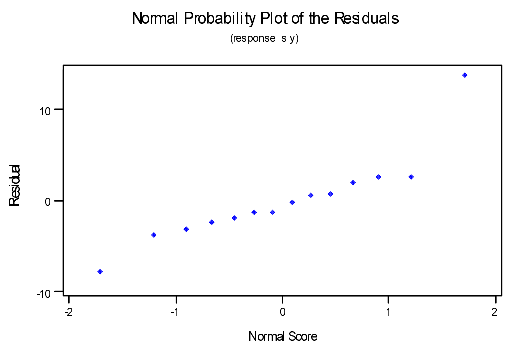
c) $\hat{\sigma}^2 = 25.23842$

d) $\hat{\sigma}_{orig}^2 = 7.324951$

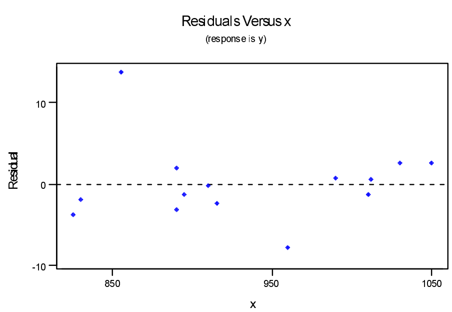
The new estimate is larger because the new point added additional variance not accounted for by the model.

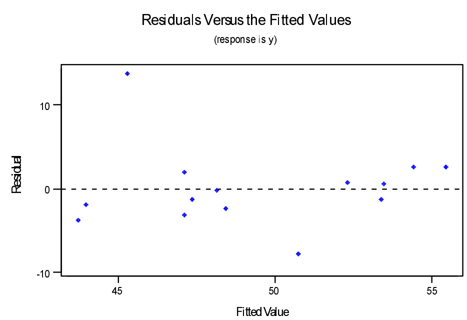
e) Yes, e_{14} is especially large compared to the other residuals.

f) The one added point is an outlier and the normality assumption is not as valid with the point included.



g) Constant variance assumption appears valid except for the added point.





10-47. a) $R^2 = 71.27\%$

b) H_0: simple linear regression model correct

H_1: simple linear regression model not correct

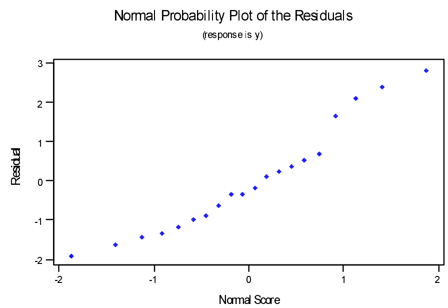
$f_0 = 0.196493$

$f_{.05,m-2,n-m} = f_{.05,8,10} = 3.072$

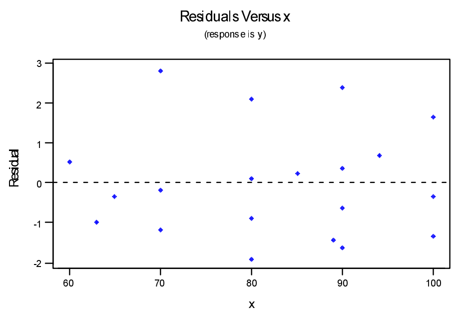
$f_0 < f_{\alpha,8,10}$

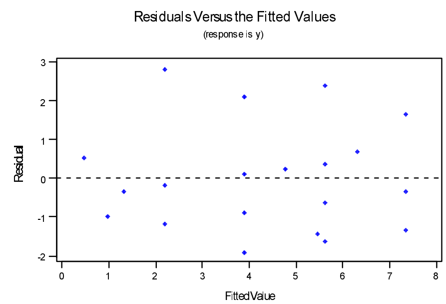
Do not reject H_0. P-value = 0.9849.

c) No major departure from normality assumptions.



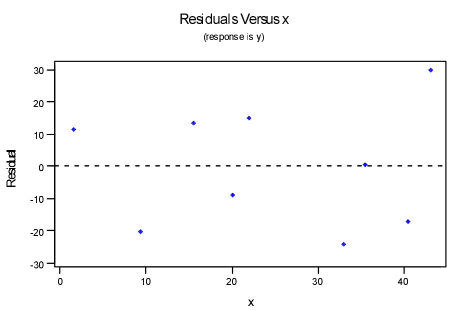
d) Assumption of constant variance appears reasonable.

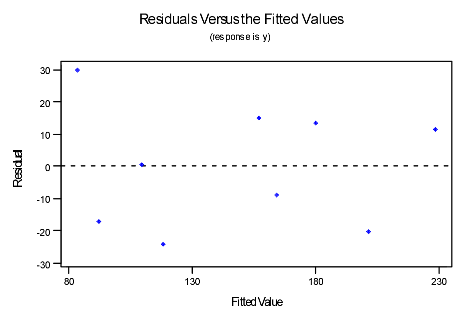




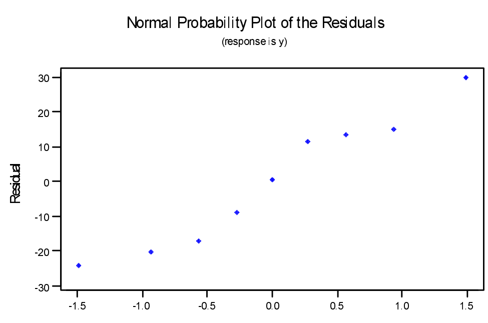
10-48. a) $R^2 = 0.879397$

b) No departure from constant variance are noted.





c) Normality assumption appears reasonable.



10-49. a) $R^2 = 85.22\%$

b) $SS_{PE} = 4.505$

c) H_0: model is correct

H_1: model is not correct

$\alpha = 0.05$

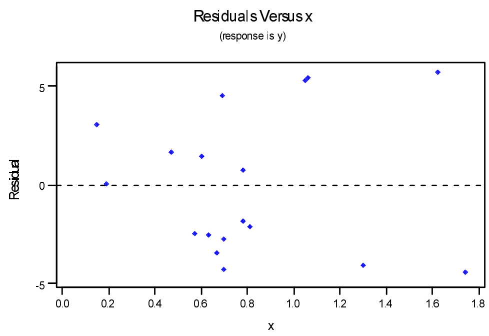
$f_0 = 8.8638$

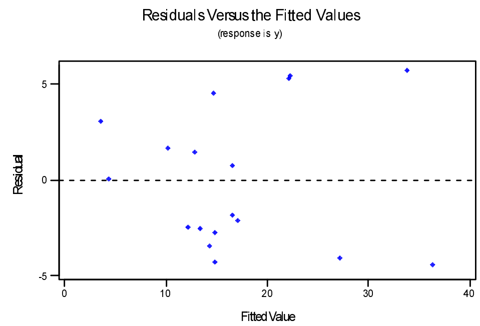
$f_{.05,14,2} = 19.424$

$f_0 < f_{\alpha,14,2}$

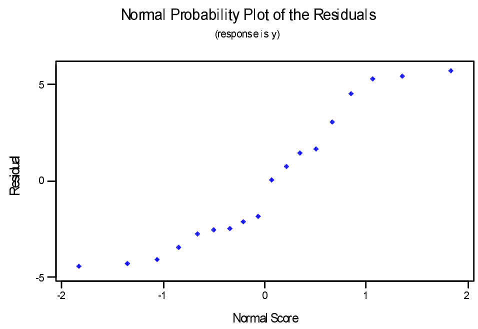
Do not reject H_0. Conclude that the simple linear regression model is reasonable. P-value = 0.134.

d) Assumptions appear reasonable, but there is a suggestion that variability increases with $\hat{y}$.





e) Normality assumption seems reasonable.



10-50. a) $R^2 = 0.896081$

b) Yes, the two points with residuals much larger in magnitude than the others.

c) $R^2_{new\ model} = 0.9573$

Larger, because the model is better able to account for the variability in the data with these two outlying data points removed.

d) $\hat{\sigma}^2_{old\ model} = 9811.21$

$\hat{\sigma}^2_{new\ model} = 4022.93$

Yes, reduced more than 50%, because the two removed points accounted for a large amount of the error.

10-51. Using $R^2 = 1 - \frac{SS_E}{S_{yy}}$, $F_0 = \frac{(n-2)(1-\frac{SS_E}{S_{yy}})}{\frac{SS_E}{S_{yy}}} = \frac{S_{yy} - SS_E}{\frac{SS_E}{n-2}} = \frac{S_{yy} - SS_E}{\hat{\sigma}^2}$

Also,

$$SS_E = \sum (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2$$
$$= \sum (y_i - \bar{y} - \hat{\beta}_1(x_i - \bar{x}))^2$$
$$= \sum (y_i - \bar{y}) + \hat{\beta}_1^2 \sum (x_i - \bar{x})^2 - 2\hat{\beta}_1 \sum (y_i - \bar{y})(x_i - \bar{x})$$
$$= \sum (y_i - \bar{y})^2 - \hat{\beta}_1^2 \sum (x_i - \bar{x})^2$$
$$S_{yy} - SS_E = \hat{\beta}_1^2 \sum (x_i - \bar{x})^2$$

Therefore, $F_0 = \frac{\hat{\beta}_1^2}{\hat{\sigma}^2 / S_{xx}} = t_0^2$

Because the square of a t random variable with n-2 degrees of freedom is an F random variable with 1 and n-2 degrees of freedom, the usually t-test that compares $|t_0|$ to $t_{\alpha/2,n-2}$ is equivalent to comparing $f_0 = t_0^2$ to $f_{\alpha,1,n-2} = t_{\alpha/2,n-2}$.

10-52. a) $f_0 = \frac{0.9(23)}{1-0.9} = 207$. Reject $H_0: \beta_1 = 0$.

b) Because $f_{.05,1,23} = 4.28$, H_0 is rejected if $\frac{23R^2}{1-R^2} > 4.28$.

That is, H_0 is rejected if

$$23R^2 > 4.28(1-R^2)$$
$$27.28R^2 > 4.28$$
$$R^2 > 0.157$$

10-53. Yes, the larger residuals are easier to identify.

10-54. For two random variables X_1 and X_2,

$$V(X_1 + X_2) = V(X_1) + V(X_2) + 2Cov(X_1, X_2)$$

Then,

$$V(Y_i - \hat{Y}_i) = V(Y_i) + V(\hat{Y}_i) - 2Cov(Y_i, \hat{Y}_i)$$
$$= \sigma^2 + V(\hat{\beta}_0 + \hat{\beta}_1 x_i) - 2\sigma^2 \left[\frac{1}{n} + \frac{(x_i - \bar{x})^2}{S_{xx}} \right]$$
$$= \sigma^2 + \sigma^2 \left[\frac{1}{n} + \frac{(x_i - \bar{x})^2}{S_{xx}} \right] - 2\sigma^2 \left[\frac{1}{n} + \frac{(x_i - \bar{x})^2}{S_{xx}} \right]$$
$$= \sigma^2 \left[1 - (\frac{1}{n} + \frac{(x_i - \bar{x})^2}{S_{xx}}) \right]$$

a) Because e_i is divided by an estimate of its standard error (when σ^2 is estimated by $\hat{\sigma}^2$), r_i has approximate unit variance.

b) No, the term in brackets in the denominator is necessary.

c) If x_i is near $\bar{x}$ and n is reasonably large, r_i is approximately equal to the standardized residual.

d) If x_i is far from $\bar{x}$, the standard error of e_i is small. Consequently, extreme points are better fit by least squares regression than points near the middle range of x. Because the studentized residual at any point has variance of approximately one, the studentized residuals can be used to compare the fit of points to the regression line over the range of x.

Section 10-10

10-55. a) $\hat{y} = -0.0280411 + 0.990987x$

b) $H_0: \beta_1 = 0$ $\quad H_1: \beta_1 \neq 0$ $\quad \alpha = 0.05$

$f_0 = 79.838$

$f_{.05,1,18} = 4.41$

$f_0 >> f_{\alpha,1,18}$

Reject H_0.

c) $r = \sqrt{0.816} = 0.903$

d) $H_0: \rho = 0$ $\quad H_1: \rho \neq 0$ $\quad \alpha = 0.05$

$$t_0 = \frac{R\sqrt{n-2}}{\sqrt{1-R^2}} = \frac{0.90334\sqrt{18}}{\sqrt{1-0.816}} = 8.9345$$

$t_{.025,18} = 2.101$

$t_0 > t_{\alpha/2,18}$

Reject H_0.

e) $H_0: \rho = 0.5$ $\quad H_1: \rho \neq 0.5$ $\quad \alpha = 0.05$

$z_0 = 3.879$

$z_{.025} = 1.96$

$z_0 > z_{\alpha/2}$

Reject H_0.

f) $\tanh(\text{arctanh } 0.90334 - \frac{z_{.025}}{\sqrt{17}}) \leq \rho \leq \tanh(\text{arctanh } 0.90334 + \frac{z_{.025}}{\sqrt{17}})$ where $z_{.025} = 1.96$.

$0.7677 \leq \rho \leq 0.9615$.

10-56. a) $\hat{y} = 69.1044 + 0.419415x$

b) $H_0: \beta_1 = 0$ $\quad H_1: \beta_1 \neq 0$ $\quad \alpha = 0.05$

$f_0 = 35.744$

$f_{.05,1,24} = 4.262$

$f_0 > f_{\alpha,1,24}$

Reject H_0.

c) $r = 0.77349$

d) $H_0: \rho = 0$ $\quad H_1: \rho \neq 0$ $\quad \alpha = 0.05$

$$t_0 = \frac{0.77349\sqrt{24}}{\sqrt{1-0.5983}} = 5.9787$$

$t_{.025,24} = 2.064$

$t_0 > t_{\alpha/2,24}$

Reject H_0.

e) $H_0: \rho = 0.6$ $\quad H_1: \rho \neq 0.6$ $\quad \alpha = 0.05$

$z_0 = (\text{arctanh } 0.77349 - \text{arctanh } 0.6)(23)^{1/2} = 1.6105$

$z_{.025} = 1.96$

$z_0 \not> z_{\alpha/2}$

Do not reject H_0.

f) $\tanh(\text{arctanh } 0.77349 - \frac{z_{.025}}{\sqrt{23}}) \leq \rho \leq \tanh(\text{arctanh } 0.77349 + \frac{z_{.025}}{\sqrt{23}})$ where $z_{.025} = 1.96$.

$0.5513 \leq \rho \leq 0.8932$.

10-57. a) r = -0.738027

b) $H_0: \rho = 0$ $\quad H_1: \rho \neq 0$ $\quad \alpha = 0.05$

$$t_0 = \frac{-0.738027\sqrt{26}}{\sqrt{1-0.5447}} = -5.577$$

$$t_{.025,26} = 2.056$$

$$|t_0| > t_{\alpha/2,26}$$

Reject H_0. P-value = (3.69E-6)(2) = 7.38E-6

c) $\tanh(\text{arctanh } -0.738 - \frac{z_{.025}}{\sqrt{25}}) \le \rho \le \tanh(\text{arctanh } -0.738 + \frac{z_{.025}}{\sqrt{25}})$

where $z_{.025} = 1.96$. $\quad -0.871 \le \rho \le -0.504$.

d) $H_0: \rho = -0.7$ $\quad H_1: \rho \neq -0.7$ $\quad \alpha = 0.05$

$$z_0 = (\text{arctanh } -0.738 - \text{arctanh } -0.7)(25)^{1/2} = -0.394$$

$$z_{.025} = 1.96$$

$$|z_0| < z_{\alpha/2}$$

Do not reject H_0. P-value = (0.3468)(2) = 0.6936

10-58 $R = \hat{\beta}_1 \left(\frac{S_{xx}}{S_{yy}} \right)^{1/2}$ and $1 - R^2 = \frac{SS_E}{S_{yy}}$.

Therefore, $T_0 = \frac{R\sqrt{n-2}}{\sqrt{1-R^2}} = \frac{\hat{\beta}_1 \left(\frac{S_{xx}}{S_{yy}} \right)^{1/2} \sqrt{n-2}}{\left(\frac{SS_E}{S_{yy}} \right)^{1/2}} = \frac{\hat{\beta}_1}{\sqrt{\frac{\hat{\sigma}^2}{S_{xx}}}}$ where $\hat{\sigma}^2 = \frac{SS_E}{n-2}$.

10-59 n = 50 r = 0.62

a) $H_0: \rho = 0$ $\quad H_1: \rho \neq 0$ $\quad \alpha = 0.01$

$$t_0 = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \frac{0.62\sqrt{48}}{\sqrt{1-(0.62)^2}} = 5.475$$

$$t_{.005,48} = 2.683$$

$$t_0 > t_{\alpha/2,48}$$

Reject H_0. P-value $\cong 0$

b) $\tanh(\text{arctanh } 0.62 - \frac{z_{.005}}{\sqrt{47}}) \le \rho \le \tanh(\text{arctanh } 0.62 + \frac{z_{.005}}{\sqrt{47}})$

where $z_{.005} = 2.575$. $\quad 0.3358 \le \rho \le 0.8007$.

c) Yes.

10-60. n = 10000, r = 0.02

a) $H_0: \rho = 0$ $\quad H_1: \rho \neq 0$ $\quad \alpha = 0.05$

$$t_0 = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \frac{0.02\sqrt{10000}}{\sqrt{1-(0.02)^2}} = 2.0004$$

$$t_{.025,9998} = 1.96$$

$$t_0 > t_{\alpha/2,9998}$$

Reject H_0. P-value = 2(0.0234) = 0.0468

b) Since the sample size is so large, the standard error is very small. Therefore, very small differences are found to be "statistically" significant. However, the practical significance is minimal since r = 0.02 is essentially zero.

10-61. a) r = 0.933203

b) $H_0: \rho = 0$ $\quad H_1: \rho \neq 0$ $\quad \alpha = 0.05$

$$t_0 = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \frac{0.933203\sqrt{15}}{\sqrt{1-(0.8709)}} = 10.06$$

$t_{.025,15} = 2.131$

$t_0 > t_{\alpha/2,15}$

Reject H_0.

c) $\hat{y} = 0.72538 + 0.498081x$

$H_0: \beta_1 = 0$ $\quad H_1: \beta_1 \neq 0$ $\quad \alpha = 0.05$

$f_0 = 101.16$

$f_{.05,1,15} = 4.545$

$f_0 >> f_{\alpha,1,15}$

Reject H_0. Conclude that the model is significant at $\alpha = 0.05$. This test and the one in part b. are identical.

d) $H_0: \beta_0 = 0$ $\quad H_1: \beta_0 \neq 0$ $\quad \alpha = 0.05$

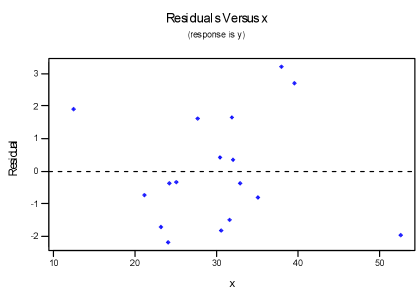
$t_0 = 0.468345$

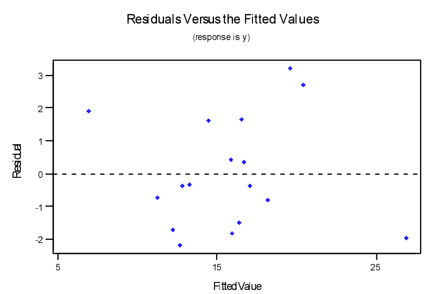
$t_{.025,15} = 2.131$

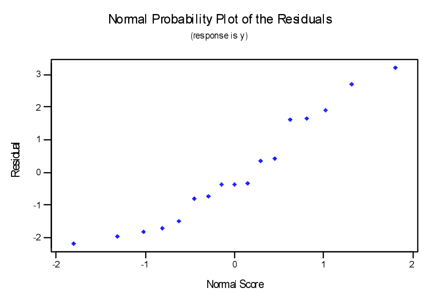
$t_0 \not> t_{\alpha/2,15}$

Do not reject H_0. We cannot conclude β_0 is different from zero.

e) No problems with model assumptions are noted.







10-62. n = 25 $\quad$ r = 0.83

a) $H_0: \rho = 0$ $\quad H_1: \rho \neq 0$ $\quad \alpha = 0.05$

$$t_0 = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \frac{0.83\sqrt{23}}{\sqrt{1-(0.83)^2}} = 7.137$$

$t_{.025,23} = 2.069$

$t_0 > t_{\alpha/2,23}$

Reject H_0. P-value = 0.

b) $\tanh(\text{arctanh } 0.83 - \frac{z_{.025}}{\sqrt{22}}) \leq \rho \leq \tanh(\text{arctanh } 0.83 + \frac{z_{.025}}{\sqrt{22}})$

where $z_{.025} = 1.96$. $0.6471 \leq \rho \leq 0.9226$.

c) $H_0: \rho = 0.8$ $\quad H_1: \rho \neq 0.8$ $\quad \alpha = 0.05$

$z_0 = (\text{arctanh } 0.83 - \text{arctanh } 0.8)(22)^{1/2} = 0.4199$

$z_{.025} = 1.96$

$z_0 \not> z_{\alpha/2}$

Do not reject H_0 . P-value = (0.3373)(2) = 0.6746.

Supplemental Exercises

10-63. a) $\sum_{i=1}^{n}(y_i - \hat{y}_i) = \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} \hat{y}_i$ and $\sum y_i = n\hat{\beta}_0 + \hat{\beta}_1 \sum x_i$ from normal equation

Then,

$$(n\hat{\beta}_0 + \hat{\beta}_1 \sum_{i=1}^{n} x_i) - \sum \hat{y}_i$$

$$= n\hat{\beta}_0 + \hat{\beta}_1 \sum_{i=1}^{n} x_i - \sum_{i=1}^{n}(\hat{\beta}_0 + \hat{\beta}_1 x_i)$$

$$= n\hat{\beta}_0 + \hat{\beta}_1 \sum_{i=1}^{n} x_i - n\hat{\beta}_0 - \hat{\beta}_1 \sum_{i=1}^{n} x_i = 0$$

b) $\sum_{i=1}^{n}(y_i - \hat{y}_i)x_i = \sum_{i=1}^{n} y_i x_i - \sum_{i=1}^{n} \hat{y}_i x_i$

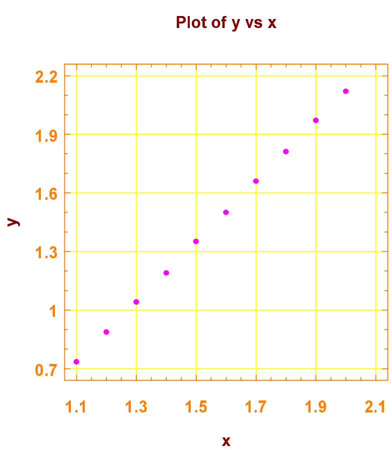
and $\sum_{i=1}^{n} y_i x_i = \hat{\beta}_0 \sum_{i=1}^{n} x_i + \hat{\beta}_1 \sum_{i=1}^{n} x_i^2$ from normal equations. Then,

$$\hat{\beta}_0 \sum_{i=1}^{n} x_i + \hat{\beta}_1 \sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n}(\hat{\beta}_0 + \hat{\beta}_1 x_i)x_i =$$

$$\hat{\beta}_0 \sum_{i=1}^{n} x_i + \hat{\beta}_1 \sum_{i=1}^{n} x_i^2 - \hat{\beta}_0 \sum_{i=1}^{n} x_i - \hat{\beta}_1 \sum_{i=1}^{n} x_i^2 = 0$$

c) $\frac{1}{n}\sum_{i=1}^{n} \hat{y}_i$

10-64. a)



Yes, a straight line relationship seems plausible.

b)

```
                       Model fitting results for: y
Independent variable           coefficient  std. error     t-value   sig.level
CONSTANT                         -0.966824    0.004845   -199.5413      0.0000
x                                 1.543758    0.003074    502.2588      0.0000
--------------------------------------------------------------------------------
R-SQ. (ADJ.) = 1.0000  SE=       0.002792  MAE=       0.002063  DurbWat=  2.843
Previously:    0.0000            0.000000             0.000000            0.000
10 observations fitted, forecast(s) computed for 0 missing val. of dep. var.
```

$\hat{y} = -0.966824 + 1.54376x$

c)

```
                    Analysis of Variance for the Full Regression
Source                 Sum of Squares     DF    Mean Square      F-Ratio   P-value
Model                         1.96613      1        1.96613      252264.     .0000
Error                    0.0000623515      8  0.00000779394
--------------------------------------------------------------------------------
Total (Corr.)                 1.96619      9
R-squared = 0.999968                             Stnd. error of est. = 2.79176E-3
R-squared (Adj. for d.f.) = 0.999964           Durbin-Watson statistic = 2.84309
```

2) $H_0 : \beta_1 = 0$

3) $H_1 : \beta_1 \neq 0$

4) $\alpha = 0.05$

5) The test statistic is $f_0 = \dfrac{SS_R / k}{SS_E / (n-p)}$

6) Reject H_0 if $f_0 > f_{\alpha,1,8}$ where $f_{0.05,1,8} = 5.32$

7) Using the results from the ANOVA table

$$f_0 = \frac{1.96613/1}{0.0000623515/8} = 255263.9$$

8) Since 2552639 > 5.32 reject H_0 and conclude that the regression model is significant at $\alpha = 0.05$.
P-value < 0.000001

d)

```
             95 percent confidence intervals for coefficient estimates
--------------------------------------------------------------------------------
                         Estimate  Standard error     Lower Limit     Upper Limit
CONSTANT                 -0.96682         0.00485        -0.97800        -0.95565
x                         1.54376         0.00307         1.53667         1.55085
--------------------------------------------------------------------------------
```

$-0.97800 \leq \beta_0 \leq -0.95565$

e) 2) $H_0 : \beta_0 = 0$

3) $H_1 : \beta_0 \neq 0$

4) $\alpha = 0.05$

5) The test statistic is $t_0 = \dfrac{\hat{\beta}_0}{se(\hat{\beta}_0)}$

6) Reject H_0 if $t_0 < -t_{\alpha/2,n-2}$ where $-t_{0.025,8} = -2.306$ or $t_0 > t_{0.025,8} = 2.306$

7) Using the results from the table above

$$t_0 = \frac{-0.96682}{0.00485} = -199.34$$

8) Since $-199.34 < -2.306$ reject H_0 and conclude the intercept is significant at $\alpha = 0.05$.

10-65. a) $\hat{y} = 93.34 + 15.64x$

b) $H_0 : \beta_1 = 0$ $H_1 : \beta_1 \neq 0$ $\alpha = 0.05$

$f_0 = 12.872$

$f_{.05,1,14} = 4.60$

$f_0 > f_{\alpha,1,14}$

Reject H_0. Conclude that $\beta_1 \neq 0$ at $\alpha = 0.05$.

c) $(7.961 \le \beta_1 \le 23.322)$

d) $(74.758 \le \beta_0 \le 111.923)$

e) $\hat{y} = 93.34 + 15.64(2.5) = 132.44$

$$132.44 \pm 2.145\sqrt{136.27\left[\frac{1}{16} + \frac{(2.5 - 2.325)^2}{977.03}\right]}$$

132.44 ± 6.26

$126.18 \le \hat{\mu}_{Y|x_0} \le 138.70$

f) H_0: simple linear regression model is correct

H_1: simple linear regression model is not correct

$\alpha = 0.05$

$f_0 = 2.002$

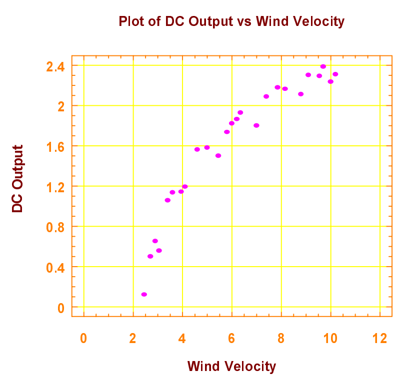
$f_{\alpha,m-2,n-m} = f_{.05,8,6} = 4.1469$

$f_0 \not> f_{\alpha,8,6}$

Do not reject H_0. Conclude evidence is insufficient to claim that the model is not correct at $\alpha = 0.05$

P-value = 0.2066

10-66. a)



Scatter diagram shows definite curvature. So, a higher order model may be appropriate or a transformation of variables.

b)

```
                        Model fitting results for: DC Output
--------------------------------------------------------------------------------
Independent variable                coefficient  std. error     t-value   sig.level
--------------------------------------------------------------------------------
CONSTANT                               0.130875     0.125989      1.0388      0.3097
Wind Vel                               0.241149     0.019049     12.6593      0.0000
--------------------------------------------------------------------------------
R-SQ. (ADJ.) = 0.8690  SE=        0.236052  MAE=        0.188745  DurbWat=  1.214
Previously:    0.0000             0.000000              0.000000             0.000
25 observations fitted, forecast(s) computed for 0 missing val. of dep. var.
```

$\hat{y} = 0.1309 + 0.2411x$

c)

Analysis of Variance for the Full Regression

Source	Sum of Squares	DF	Mean Square	F-Ratio	P-value
Model	8.92961	1	8.92961	160.257	.0000
Error	1.28157	23	0.0557206		
Total (Corr.)	10.2112	24			

R-squared = 0.874493 Stnd. error of est. = 0.236052
R-squared (Adj. for d.f.) = 0.869036 Durbin-Watson statistic = 1.21409

2) $H_0:\beta_1 = 0$

3) $H_1:\beta_1 \neq 0$

4) $\alpha = 0.05$

5) The test statistic is $f_0 = \dfrac{SS_R / k}{SS_E / (n-p)}$

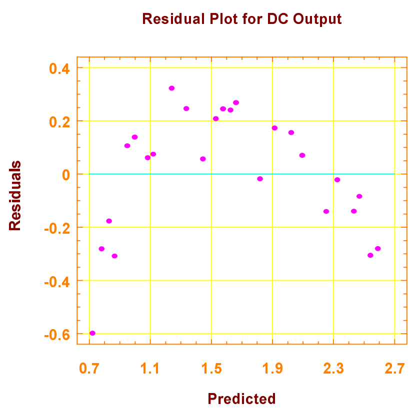
6) Reject H_0 if $f_0 > f_{\alpha,1,23}$ where $f_{0.05,1,23} = 4.28$

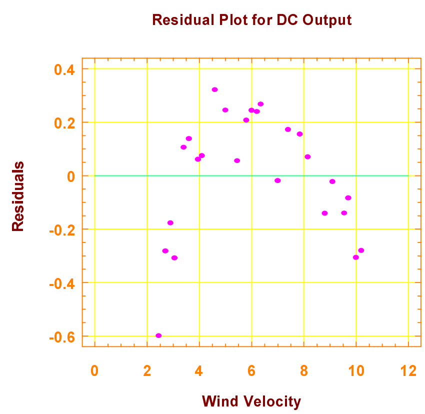
7) Using the results from the ANOVA table

$$f_0 = \frac{8.92961/1}{1.28157/23} = 160.257$$

8) Since 160.257 > 4.28 reject H_0 and conclude that the regression model is significant at $\alpha = 0.05$.

d)





Conclude plots indicate model inadequacy since the residual plots exhibit nonrandom patterns.

e) Examining the residual plots in part d), a transformation on the x-variable, y-variable, or both would be appropriate. A simple linear regression of y on the transformed variable 1/x may be satisfactory.

f) The following analysis employs the transformed variable, 1/x

Model fitting results for: DC Output

Independent variable	coefficient	std. error	t-value	sig.level
CONSTANT	2.97886	0.044902	66.3410	0.0000
1/WindVel	-6.934547	0.206434	-33.5922	0.0000

R-SQ. (ADJ.) = 0.9792 SE= 0.094171 MAE= 0.073975 DurbWat= 1.567
Previously: 0.0000 0.000000 0.000000 0.000
25 observations fitted, forecast(s) computed for 0 missing val. of dep. var.

$$\hat{y} = 2.9789 - 6.9345x^*$$

Analysis of Variance for the Full Regression

Source	Sum of Squares	DF	Mean Square	F-Ratio	P-value
Model	10.0072	1	10.0072	1128.43	.0000
Error	0.203970	23	0.00886825		
Total (Corr.)	10.2112	24			

R-squared = 0.980025 Stnd. error of est. = 0.0941714
R-squared (Adj. for d.f.) = 0.979156 Durbin-Watson statistic = 1.56651

2) $H_0: \beta_1 = 0$

3) $H_1: \beta_1 \neq 0$

4) $\alpha = 0.05$

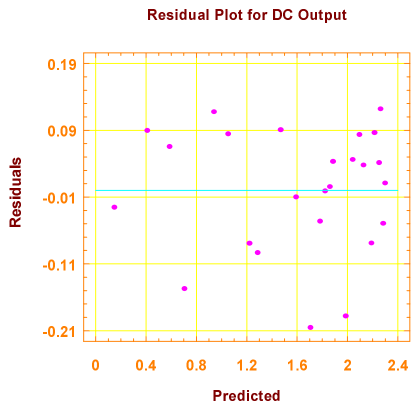
5) The test statistic is $f_0 = \dfrac{SS_R / k}{SS_E / (n-p)}$

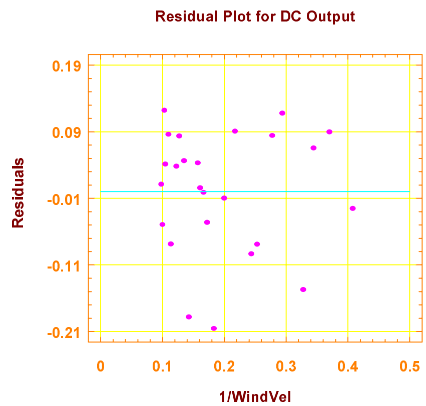
6) Reject H_0 if $f_0 > f_{\alpha,1,23}$ where $f_{0.05,1,23} = 4.28$

7) Using the results from the ANOVA table

$$f_0 = \frac{10.0072 / 1}{0.203970 / 23} = 1128.43$$

8) Since 1128.43 > 4.28 reject H_0 and conclude that the regression model is significant at $\alpha = 0.05$.

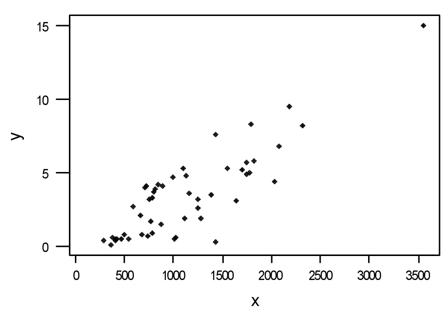




Using 1/WindVel

Conclude from the random appearance of the residuals in the plots and significance of regression that the model is adequate. The transformation, 1/(Wind Velocity), appears to be satisfactory as a regressor of Output.

10-67. a)



b) $\hat{y} = -0.8819 + 0.00385x$

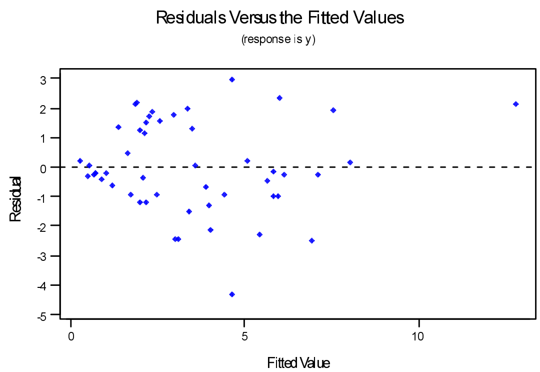
c) $H_0: \beta_1 = 0$ $H_1: \beta_1 \neq 0$ $\alpha = 0.05$

$f_0 = 122.03$

$f_0 > f_{\alpha,1,48}$

Reject H_0. Conclude that regression model is significant at $\alpha = 0.05$

d) No, it seems the variance is not constant.



e) $\hat{y}^* = 0.5967 + 0.00097x$. Yes, the transformation stabilizes the variance.

10-68. $\hat{y}^* = 1.2232 + 0.5075x$ where $y^* = \frac{1}{y}$. No, model does not seem reasonable. The residual plots indicate a possible outlier.

10-69. $\hat{y} = 0.7916x$
Even though y should be zero when x is zero, because the regressor variable does not normally assume values near zero, a model with an intercept fits this data better. Without an intercept, the MS_E is larger and the residuals plots are not satisfactory.

10-70. $\hat{y} = 4.5755 + 2.2047x$, r= 0.992, $R^2 = 98.40$
The model appears to be an excellent fit. Significance of regressor is strong and R^2 is large. Both regression coefficients are significant. No, the existence of a strong correlation does not imply a cause and effect relationship.

Mind-Expanding Exercises

10-71. a) $\hat{\beta}_1 = \frac{S_{xY}}{S_{xx}}$, $\hat{\beta}_0 = \overline{Y} - \hat{\beta}_1\overline{x}$

$Cov(\hat{\beta}_0, \hat{\beta}_1) = Cov(\overline{Y}, \hat{\beta}_1) - \overline{x}Cov(\hat{\beta}_1, \hat{\beta}_1)$

$Cov(\overline{Y}, \hat{\beta}_1) = \frac{Cov(\overline{Y}, S_{xY})}{S_{xx}} = \frac{Cov(\sum Y_i, \sum Y_i(x_i - \overline{x}))}{nS_{xx}} = \frac{\sum(x_i - \overline{x})\sigma^2}{nS_{xx}} = 0$. Therefore, $Cov(\hat{\beta}_0, \hat{\beta}_1) = \frac{-\overline{x}\sigma^2}{S_{xx}}$

$Cov(\hat{\beta}_1, \hat{\beta}_1) = V(\hat{\beta}_1) = \frac{\sigma^2}{S_{xx}}$

b) The requested result is shown in part a.

10-72. a) $MS_E = \frac{\sum (Y_i - \hat\beta_0 - \hat\beta_1 x_i)^2}{n-2} = \frac{\sum e_i^2}{n-2}$

$$E(e_i) = E(Y_i) - E(\hat\beta_0) - E(\hat\beta_1)x_i = 0$$

From Exercise 10-54, $V(e_i) = \sigma^2[1 - (\frac{1}{n} + \frac{(x_i - \bar{x})^2}{S_{xx}})]$ Therefore,

$$E(MS_E) = \frac{\sum E(e_i^2)}{n-2} = \frac{\sum V(e_i)}{n-2}$$

$$= \frac{\sum \sigma^2[1 - (\frac{1}{n} + \frac{(x_i - \bar{x})^2}{S_{xx}})]}{n-2}$$

$$= \frac{\sigma^2[n-1-1]}{n-2} = \sigma^2$$

b) Using the fact that $SS_R = MS_R$, we obtain

$$E(MS_R) = E(\hat\beta_1^2 S_{xx}) = S_{xx}\left\{V(\hat\beta_1) + [E(\hat\beta_1)]^2\right\}$$

$$= S_{xx}\left(\frac{\sigma^2}{S_{xx}} + \beta_1^2\right) = \sigma^2 + \beta_1^2 S_{xx}$$

10-73. $\hat\beta_1 = \frac{S_{x_1Y}}{S_{x_1x_1}}$

$$E(\hat\beta_1) = \frac{E\left[\sum_{i=1}^{n} Y_i(x_{1i} - \bar{x}_1)\right]}{S_{x_1x_1}} = \frac{\sum_{i=1}^{n} (\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i})(x_{1i} - \bar{x}_1)}{S_{x_1x_1}}$$

$$= \frac{\beta_1 S_{x_1x_1} + \beta_2 \sum_{i=1}^{n} x_{2i}(x_{1i} - \bar{x}_1)}{S_{x_1x_1}} = \beta_1 + \frac{\beta_2 S_{x_1x_2}}{S_{x_1x_1}}$$

No, $\hat\beta_1$ is no longer unbiased.

10-74. $V(\hat\beta_1) = \frac{\sigma^2}{S_{xx}}$. To minimize $V(\hat\beta_1)$, S_{xx} should be maximized. Because $S_{xx} = \sum_{i=1}^{n} (x_i - \bar{x})^2$, S_{xx} is maximized by choosing approximately half of the observations at each end of the range of x.

From a practical perspective, this allocation assumes the linear model between Y and x holds throughout the range of x and observing Y at only two x values prohibits verifying the linearity assumption. It is often preferable to obtain some observations at intermediate values of x.

10-75. One might minimize a weighted some of squares $\sum_{i=1}^{n} w_i(y_i - \beta_0 - \beta_1 x_i)^2$ in which a Y_i with small variance (w_i large) receives greater weight in the sum of squares.

$$\frac{\partial}{\beta_0}\sum_{i=1}^{n} w_i(y_i - \beta_0 - \beta_1 x_i)^2 = -2\sum_{i=1}^{n} w_i(y_i - \beta_0 - \beta_1 x_i)$$
$$\frac{\partial}{\beta_1}\sum_{i=1}^{n} w_i(y_i - \beta_0 - \beta_1 x_i)^2 = -2\sum_{i=1}^{n} w_i(y_i - \beta_0 - \beta_1 x_i)x_i$$

Setting these derivatives to zero yields

$$\hat{\beta}_0 \sum w_i + \hat{\beta}_1 \sum w_i x_i = \sum w_i y_i$$

$$\hat{\beta}_0 \sum w_i x_i + \hat{\beta}_1 \sum w_i x_i^2 = \sum w_i x_i y_i$$

as requested.

10-76. $$\hat{y} = \bar{y} + r\frac{s_y}{s_x}(x - \bar{x})$$
$$= \bar{y} + \frac{S_{xy}\sqrt{\sum(y_i - \bar{y})^2}(x - \bar{x})}{\sqrt{S_{xx}S_{yy}}\sqrt{\sum(x_i - \bar{x})^2}}$$
$$= \bar{y} + \frac{S_{xy}}{S_{xx}}(x - \bar{x})$$
$$= \bar{y} + \hat{\beta}_1 x - \hat{\beta}_1 x = \hat{\beta}_0 + \hat{\beta}_1 x$$

10-77. a) $$\frac{\partial}{\beta_1}\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)^2 = -2\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)x_i$$

Upon setting the derivative to zero, we obtain

$$\beta_0 \sum x_i + \beta_1 \sum x_i^2 = \sum x_i y_i$$

Therefore,

$$\hat{\beta}_1 = \frac{\sum x_i y_i - \beta_0 \sum x_i}{\sum x_i^2} = \frac{\sum x_i(y_i - \beta_0)}{\sum x_i^2}$$

b) $$V(\hat{\beta}_1) = V\left(\frac{\sum x_i \ (Y_i - \beta_0)}{\sum x_i^2}\right) = \frac{\sum x_i^2 \sigma^2}{[\sum x_i^2]^2} = \frac{\sigma^2}{\sum x_i^2}$$

c) $$\hat{\beta}_1 \pm t_{\alpha/2,n-1}\sqrt{\frac{\hat{\sigma}^2}{\sum x_i^2}}$$

This confidence interval is shorter because $\sum x_i^2 \geq \sum(x_i - \bar{x})^2$. Also, the t value based on n-1 degrees of freedom is slightly smaller than the corresponding t value based on n-2 degrees of freedom.

CHAPTER 11

Sections 11-2 and 11-3

11-1. a) $X'X = \begin{bmatrix} 10 & 223 & 553 \\ 223 & 5200.9 & 12352 \\ 553 & 12352 & 31729 \end{bmatrix}$

$X'y = \begin{bmatrix} 1916.0 \\ 43550.8 \\ 104736.8 \end{bmatrix}$

b) $\hat{\beta} = \begin{pmatrix} 171.054 \\ 3.713 \\ -1.126 \end{pmatrix}$

c) at x = (1, 18, 43), $\hat{y} = 189.481$

11-2. a) $\hat{\beta} = (X'X)^{-1}X'y$

$\hat{\beta} = \begin{pmatrix} -1.9122 \\ 0.0931 \\ 0.2532 \end{pmatrix}$

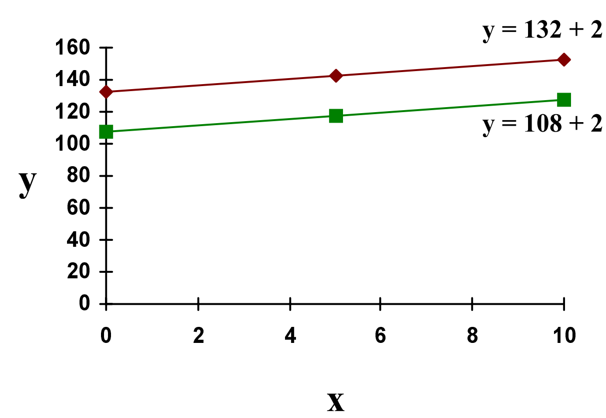
b) $\hat{y} = -1.9122 + 0.0931x_1 + 0.2532x_2$

$\hat{y} = -1.9122 + 0.0931(200) + 0.2532(50) = 29.3678$

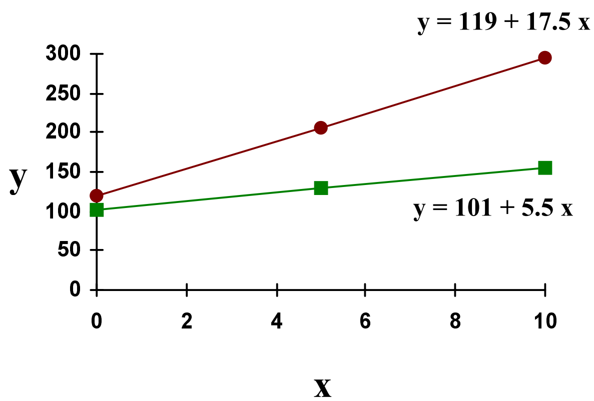
11-3. a)

	Model 1	Model 2
$x_2 = 2$	$\hat{y} = 100 + 2x_1 + 8$	$\hat{y} = 95 + 1.5x_1 + 3(2) + 4x_1$
	$\hat{y} = 108 + 2x_1$	$\hat{y} = 101 + 5.5x_1$
$x_2 = 8$	$\hat{y} = 100 + 2x_1 + 4(8)$	$\hat{y} = 95 + 1.5x_1 + 3(8) + 16x_1$
	$\hat{y} = 132 + 2x_1$	$\hat{y} = 119 + 17.5x_1$

MODEL 1



MODEL 2



The interaction term in model 2 affects the slope of the regression equation. That is, it modifies the amount of change per unit of x_1 on $\hat{y}$.

b) $x_2 = 5$ $\quad$ $\hat{y} = 100 + 2x_1 + 4(5)$

$\hat{y} = 120 + 2x_1$

Then, 2 is the expected change on $\hat{y}$ per unit of x_1.

NO, it does not depend on the value of x_2, because there is no relationship or interaction between these two variables in model 1.

c)

	$x_2 = 5$	$x_2 = 2$	$x_2 = 8$
	$\hat{y} = 95 + 1.5x_1 + 3(5) + 2x_1(5)$ $\hat{y} = 110 + 11.5x_1$	$\hat{y} = 101 + 5.5x_1$	$\hat{y} = 119 + 17.5x_1$
Change per unit of X_1	11.5	5.5	17.5

Yes, result does depend on the value of x_2, because x_2 interacts with x_1.

11-4.

```
Predictor        Coef        StDev          T        P
Constant       -7.634        7.848      -0.97    0.340
x2          0.0039765    0.0007136       5.57    0.000
X7            0.24777      0.08896       2.79    0.010
X8          -0.003893     0.001296      -3.00    0.006
S = 1.797       R-Sq = 76.3%     R-Sq(adj) = 73.3%
Analysis of Variance
Source        DF          SS          MS         F        P
Regression     3     249.454      83.151     25.75    0.000
Error         24      77.511       3.230
Total         27     326.964
```

a) $\hat{y} = -7.6345 + 0.0040x_2 + 0.2478x_7 - 0.0039x_8$

b) $\hat{y} = -7.6345 + 0.0040(2000) + 0.2478(60) - 0.0039(1800)$

$\hat{y} = 8.21 \approx 8$ games.

11-5.

Predictor	Coef	StDev	T	P
Constant	33.449	1.576	21.22	0.000
xl	-0.054349	0.006329	-8.59	0.000
x6	1.0782	0.6997	1.54	0.138

S = 2.834 R-Sq = 82.9% R-Sq(adj) = 81.3%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	2	856.24	428.12	53.32	0.000
Error	22	176.66	8.03		
Total	24	1032.90			

a) $\hat{y} = 33.4491 - 0.05435x_1 + 1.07822x_2$

b) $\hat{y} = 33.4491 - 0.05435(300) + 1.07822(4) = 21.457$ mpg.

11-6.

Independent variable	coefficient	std. error	t-value	sig.level
CONSTANT	-102.713236	207.858851	-0.4941	0.6363
x1	0.605371	0.368897	1.6410	0.1448
x2	8.923644	5.300522	1.6835	0.1361
x3	1.437457	2.391621	0.6010	0.5668
x4	0.013609	0.733821	0.0185	0.9857

R-SQ. (ADJ.) = 0.5989 SE= 15.579333 MAE= 9.776513 DurbWat= 1.772
Previously: 0.0000 0.000000 0.000000 0.000
12 observations fitted, forecast(s) computed for 0 missing val. of dep. var.

a) $\hat{y} = -102.7132 + 0.6054x_1 + 8.9236x_2 + 1.4375x_3 + 0.0136x_4$

b) $\hat{y} = -102.7132 + 0.6054(75) + 8.9236(24) + 1.4375(90) + 0.0136(98) = 287.566$

11-7.

Independent variable	coefficient	std. error	t-value	sig.level
CONSTANT	350.994271	74.753074	4.6954	0.0183
x1	-1.271994	1.16914	-1.0880	0.3562
x2	-0.153904	0.08953	-1.7190	0.1841

R-SQ. (ADJ.) = 0.7696 SE= 25.497858 MAE= 16.319125 DurbWat= 2.565
Previously: 0.0000 0.000000 0.000000 0.000
6 observations fitted, forecast(s) computed for 0 missing val. of dep. var.

a) $\hat{y} = 350.9943 - 1.272x_1 - 0.1539x_2$

b) $\hat{y} = 350.9943 - 1.272(25) - 0.1539(1000) = 165.29$

c)

Independent variable	coefficient	std. error	t-value	sig.level
CONSTANT	125.865548	197.957166	0.6358	0.5899
x1	7.758641	7.514795	1.0324	0.4104
x2	0.094304	0.220657	0.4274	0.7107
x1*x2	0.009186	0.007564	-1.2145	0.3485

R-SQ. (ADJ.) = 0.8011 SE= 23.691404 MAE= 12.527757 DurbWat= 3.003
Previously: 0.7696 25.497858 16.319125 2.565
6 observations fitted, forecast(s) computed for 0 missing val. of dep. var.

$\hat{y}' = 125.8655 - 7.7586x_1 - 0.0943x_2 - 0.0092x_1x_2$

d) $\hat{y}' = 125.8655 - 7.7586(25) - 0.0943(1000) - 0.0092(25)(1000)$

$\hat{y}' = 184.13$ The predicted value is larger

11-8.

The regression equation is
y = 7.46 - 0.030 x2 + 0.521 x3 - 0.102 x4 - 2.16 x5

Predictor	Coef	StDev	T	P
Constant	7.458	7.226	1.03	0.320
x2	-0.0297	0.2633	-0.11	0.912
x3	0.5205	0.1359	3.83	0.002
x4	-0.10180	0.05339	-1.91	0.077
x5	-2.161	2.395	-0.90	0.382

S = 0.8827 R-Sq = 67.2% R-Sq(adj) = 57.8%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	4	22.3119	5.5780	7.16	0.002
Error	14	10.9091	0.7792		
Total	18	33.2211			

a) $\hat{y} = 7.4578 - 0.0297x_2 + 0.5205x_3 - 0.1018x_4 - 2.1606x_5$

b) $\hat{y} = 7.4578 - 0.0297(20) + 0.5205(30) - 0.1018(90) - 2.1606(2.0)$ $\hat{y} = 10.037$

11-9.

Independent variable	coefficient	std. error	t-value	sig.level
CONSTANT	47.173999	49.581476	0.9514	0.3555
x1	-9.735202	3.691625	-2.6371	0.0179
x2	0.428287	0.223933	1.9126	0.0739
x3	18.237455	1.311802	13.9026	0.0000

```
R-SQ. (ADJ.) = 0.9925  SE=      3.479627  MAE=       2.511105  DurbWat=  1.778
Previously:    0.0000           0.000000             0.000000            0.000
20 observations fitted, forecast(s) computed for 0 missing val. of dep. var.
```

a) $\hat{y} = 47.174 - 9.7352x_1 + 0.4283x_2 + 18.2375x_3$

b) $\hat{y} = 47.174 - 9.7352(14.5) + 0.4283(220) + 18.2375(5) = 91.43$

11-10.

Predictor	Coef	StDev	T	P
Constant	-0.07837	0.07360	-1.06	0.297
TEMP	0.00004392	0.00003627	1.21	0.237
SOAKTIME	0.0024544	0.0002082	11.79	0.000
SOAKPCT	0.01835	0.02014	0.91	0.371
DFTIME	0.007786	0.001349	5.77	0.000
DIFFPCT	-0.003134	0.008053	-0.39	0.700

S = 0.002272   R-Sq = 96.9%   R-Sq(adj) = 96.3%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	5	0.00419226	0.00083845	162.43	0.000
Error	26	0.00013421	0.00000516		
Total	31	0.00432647			

a) $\hat{y} = -0.0784 + 0.00004x_1 + 0.0025x_2 + 0.0183x_3 + 0.0078x_4 - 0.0031x_5$

where $x_1 = \text{TEMP}$ $x_2 = \text{SOAKTIME}$ $x_3 = \text{SOAKPCT}$ $x_4 = \text{DFTIME}$ $x_5 = \text{DIFFPCT}$

b) $\hat{y} = -0.0784 + 0.00004(1650) + 0.0025(1) + 0.0183(1.10) + 0.0078(1) - 0.0031(0.8)$

$\hat{y} = 0.01555$

11-11.

Predictor	Coef	StDev	T	P
Constant	-8.01	16.18	-0.50	0.630
Pts	0.49400	0.04806	10.28	0.000
GF	0.001847	0.006424	0.29	0.779
GA	0.00231	0.02091	0.11	0.914
PPG	0.03830	0.05146	0.74	0.472
PPcT	-0.2068	0.2611	-0.79	0.445
SHG	-0.01284	0.02562	-0.50	0.626
PPGA	0.03074	0.03802	0.81	0.436
PKPcT	0.0407	0.1483	0.27	0.789
SHGA	-0.2083	0.1110	-1.88	0.087

S = 1.523   R-Sq = 98.8%   R-Sq(adj) = 97.8%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	9	2125.43	236.16	101.79	0.000
Error	11	25.52	2.32		
Total	20	2150.95			

$$\hat{y} = -8.0119 + 0.494x_1 + 0.0018x_2 + 0.0023x_3 + 0.0383x_4 - 0.2868x_5 - 0.0128x_6 + 0.030x_7 + 0.0407x_8 - 0.2083x_9$$

where

$x_1 = \text{Pts}$ $x_2 = \text{GF}$ $x_3 = \text{GA}$ $x_4 = \text{PPG}$ $x_5 = \text{PPct}$ $x_6 = \text{SHG}$ $x_7 = \text{PPGA}$

$x_8 = \text{PKPcT}$ $x_9 = \text{SHGA}$

11-12. a) $f(\beta_0', \beta_1, \beta_2) = \sum[y_i - \beta_0' - \beta_1(x_{i1} - \bar{x}_1) - \beta_2(x_{i2} - \bar{x}_2)]^2$

$$\frac{\partial f}{\partial \beta_0'} = -2\sum[y_i - \beta_0' - \beta_1(x_{i1} - \bar{x}_1) - \beta_2(x_{i2} - \bar{x}_2)]$$

$$\frac{\partial f}{\partial \beta_1} = -2\sum[y_i - \beta_0' - \beta_1(x_{i1} - \bar{x}_1) - \beta_2(x_{i2} - \bar{x}_2)](x_{i1} - \bar{x}_1)$$

$$\frac{\partial f}{\partial \beta_2} = -2\sum[y_i - \beta_0' - \beta_1(x_{i1} - \bar{x}_1) - \beta_2(x_{i2} - \bar{x}_2)](x_{i2} - \bar{x}_2)$$

Setting the derivatives equal to zero yields

$$n\beta_0' = \sum y_i$$

$$n\beta_0' + \beta_1\sum(x_{i1} - \bar{x}_1)^2 + \beta_2\sum(x_{i1} - \bar{x}_1)(x_{i2} - \bar{x}_2) = \sum y_i(x_{i1} - \bar{x}_1)$$

$$n\beta_0' + \beta_1\sum(x_{i1} - \bar{x}_1)(x_{i2} - \bar{x}_2) + \beta_2\sum(x_{i2} - \bar{x}_2)^2 = \sum y_i(x_{i2} - \bar{x}_2)$$

b) From the first normal equation, $\hat{\beta}_0' = \bar{y}$.

c) Substituting $y_i - \bar{y}$ for y_i in the first normal equation yields $\hat{\beta}_0' = 0$.

<u>Sections 11-4 and 11-5</u>

11-13. n = 10, k = 2, p = 3, $\alpha = 0.05$

$H_0: \beta_1 = \beta_2 = \ldots = \beta_k = 0$

$H_1: \beta_j \neq 0$ for at least one j

$$S_{yy} = 371595.6 - \frac{(1916)^2}{10} = 4490$$

$$X'y = \begin{bmatrix} \sum y_i \\ \sum x_{i1}y_i \\ \sum x_{i2}y_i \end{bmatrix} = \begin{bmatrix} 1030 \\ 21310 \\ 44174 \end{bmatrix}$$

$$\hat{\beta}'X'y = \begin{bmatrix} 171.054 & 3.713 & -1.126 \end{bmatrix} \begin{bmatrix} 1916 \\ 43550.8 \\ 104736.8 \end{bmatrix} = 371535.9$$

$$SS_R = 371535.9 - \frac{1916^2}{10} = 4430.38$$

$$SS_E = S_{yy} - SS_R = 4490 - 4430.38 = 59.62$$

$$f_0 = \frac{\frac{SS_R}{k}}{\frac{SS_E}{n-p}} = \frac{4430.38/2}{59.62/7} = 260.09$$

$$f_{0.05,2,7} = 4.74$$

$$f_0 > f_{0.05,2,7}$$

Reject H_0 and conclude that the regression model is significant at $\alpha = 0.05$.

b) $\hat{\sigma}^2 = MS_E = \frac{SS_E}{n-p} = 8.517$

c) $se(\hat{\beta}_1) = \sqrt{\hat{\sigma}^2 c_{11}} = \sqrt{8.517(0.00439)} = 0.195$

$se(\hat{\beta}_2) = \sqrt{\hat{\sigma}^2 c_{22}} = \sqrt{8.517(0.00087)} = 0.0861$

d) $H_0: \beta_1 = 0$ $\quad$ $\beta_2 = 0$

$H_1: \beta_1 \neq 0$ $\quad$ $\beta_2 \neq 0$

$$t_0 = \frac{\hat{\beta}_1}{se(\hat{\beta}_1)} = \frac{3.713}{0.1934} = 19.20 \qquad t_0 = \frac{\hat{\beta}_2}{se(\hat{\beta}_2)} = \frac{-1.126}{0.0861} = -13.08$$

$t_{\alpha/2,7} = t_{.025,7} = 2.365$

Reject H_0 $\quad$ Reject H_0

Both significant

11-14 $\quad S_{yy} = 742.00$

a) $H_0: \beta_1 = \beta_2 = 0$

$H_1: \beta_j \neq 0$ for at least one j

$\alpha = 0.01$

$$SS_R = \hat{\beta}'X'y - \frac{(\sum_{i=1}^{n} y_i)^2}{n}$$

$$= \begin{pmatrix} -1.9122 & 0.0931 & 0.2532 \end{pmatrix} \begin{pmatrix} 220 \\ 36768 \\ 9965 \end{pmatrix} - \frac{220^2}{10}$$

$$= 5525.5548 - 4840$$

$$= 685.55$$

$$SS_E = S_{yy} - SS_R$$

$$= 742 - 685.55$$

$$= 56.45$$

$$f_0 = \frac{\frac{SS_R}{k}}{\frac{SS_E}{n-p}} = \frac{685.55/2}{56.45/7} = 42.51$$

$f_{0.01,2,7} = 9.55$

$f_0 > f_{\alpha,2,7}$

Reject H_0 and conclude that the regression model is significant at $\alpha = 0.01$. P-value = 0.000121

b) $\hat{\sigma}^2 = MS_E = \frac{SS_E}{n-p} = \frac{56.45}{7} = 8.0643$

c) $se(\hat{\beta}_1) = \sqrt{8.0643(7.9799E-5)} = 0.0254$

d) $H_0: \beta_1 = 0$

$H_1: \beta_1 \neq 0$

$$t_0 = \frac{\hat{\beta}_1}{se(\hat{\beta}_1)} = \frac{0.0931}{0.0254} = 3.67$$

$t_{.005,7} = 3.499$

$|t_0| > t_{\alpha/2,7}$

Reject H_0 and conclude that $\hat{\beta}_1$ is significant in the model at $\alpha = 0.01$

P-value = $2(1 - P(t < t_0))$ = 2(1 - 0.996018) = 0.007964

11-15. a) $H_0: \beta_2 = \beta_7 = \beta_8 = 0$

H_1 for at least one $\beta_j \neq 0$

$\alpha = 0.05$

$f_0 = 25.7465$

$f_{.05,3,24} = 3.0103$

$f_0 > f_{\alpha,3,24}$

Reject H_0 and conclude regression model is significant at $\alpha = 0.05$
P-value < 0.000001

b) $\hat{\sigma}^2 = MS_E = 3.2296$

c) $se(\hat{\beta}_2) = \sqrt{\hat{\sigma}^2 c_{jj}} = 0.000714$

d) $H_0: \beta_2 = 0$

$H_1: \beta_2 \neq 0$

$\alpha = 0.05$

$t_0 = 5.5723$

$t_{.025,28-4} = t_{.025,24} = 2.064$

$|t_0| > t_{\alpha/2,24}$, Reject H_0

$H_0: \beta_7 = 0$ | $H_0: \beta_8 = 0$

$H_1: \beta_7 \neq 0$ | $H_1: \beta_8 \neq 0$

$\alpha = 0.05$ | $\alpha = 0.05$

$t_0 = 2.7058$ | $t_0 = -3.0047$

Reject H_0 | Reject H_0

Conclude the variables are significant at $\alpha = 0.05$

11-16. a) $SS_R(\beta_8|\beta_7,\beta_2,\beta_0) = 29.1576$

b) $H_0: \beta_8 = 0$

$H_1: \beta_8 \neq 0$

$\alpha = 0.05$

$$f_0 = \frac{SS_R(\beta_8|\beta_7,\beta_2,\beta_0)/1}{MS_E} = \frac{29.15765}{3.2296} = 9.03$$

$f_{.05,1,24} = 4.26$

$f_0 > f_{\alpha,1,24}$

Reject H_0 and conclude X_8 regressor is significant at $\alpha = 0.05$
P-value = 0.0061

11-17. a) $H_0: \beta_1 = \beta_6 = 0$

H_1: at least one $\beta \neq 0$

$f_0 = 53.3162$

$f_{\alpha,2,22} = f_{.05,2,22} = 3.44$

$f_0 > f_{\alpha,2,22}$

Reject H_0 and conclude regression model is significant at $\alpha = 0.05$

b) $H_0: \beta_1 = 0$

$H_1: \beta_1 \neq 0$

$t_0 = -8.59$

$t_{.025,25-3} = t_{.025,22} = 2.074$

$|t_0| > t_{\alpha/2,22}$, Reject H_0 and conclude x_1 is significant at $\alpha = 0.05$

$H_0: \beta_6 = 0$

$H_1: \beta_6 \neq 0$

$\alpha = 0.05$

$t_0 = 1.5411$

$|t_0| \not> t_{\alpha/2,22}$, Do not reject H_0, conclude that evidence is not significant to state x_6 is significant at α = 0.05.

NO, only x_1 contributes significantly to the regression.

11-18. a) $\beta_1: t_0 = 4.82$ P-value = 2(4.08 E-5) = 8.16 E-5

$\beta_2: t_0 = 8.21$ P-value = 291.91 E-8) = 3.82 E-8

$\beta_3: t_0 = 0.98$ P-value = 2 (0.1689) = 0.3378

b) $H_0: \beta_3 = 0$

$H_1: \beta_3 \neq 0$

$\alpha = 0.05$

$t_0 = 0.98$

$t_{\alpha/2,23} = 2.074$

$|t_0| \not> t_{\alpha/2,23}$

Do not reject H_0.

11-19. a) $H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$

H_1 at least one $\beta_j \neq 0$

$\alpha = 0.01$

$f_0 = 5.10602$

$f_{.01,4,7} = 7.85$

$f_0 \not> f_{\alpha,4,7}$

Do not reject H_0

P-value = 0.0303

b) $\hat{\sigma}^2 = 242.716$

c) $\alpha = 0.01$

$H_0: \beta_1 = 0$	$\beta_2 = 0$	$\beta_3 = 0$	$\beta_4 = 0$
$H_1: \beta_1 \neq 0$	$\beta_2 \neq 0$	$\beta_3 \neq 0$	$\beta_4 \neq 0$
$t_0 = 1.64$	$t_0 = 1.68$	$t_0 = 0.60$	$t_0 = 0.02$

$t_{\alpha/2,n-p} = t_{.005,7} = 3.499$

All $|t_0| \not> t_{\alpha/2,7}$

Do not reject H_0 for any regressor.

11-20. $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2$, Assume no interaction model.

a) $H_0: \beta_1 = \beta_2 = 0$

H_1 at least one $\beta_j \neq 0$

$f_0 = 9.35$

$f_{.05,2,3} = 9.55$

$f_0 \not> f_{\alpha,2,3}$

Do not reject H_0

P-value = 0.0514

b) $H_0: \beta_1 = 0$ | $H_0: \beta_2 = 0$

$H_1: \beta_1 \neq 0$ | $H_1: \beta_2 \neq 0$

$t_0 = -1.09$ | $t_0 = -1.72$

$t_{\alpha/2,3} = t_{.025,3} = 3.182$ | $t_{\alpha/2,3} = t_{.025,3} = 3.182$

$|t_0| \ngtr t_{\alpha/2}$, | $|t_0| \ngtr t_{\alpha/2,3}$

Do not reject H_0 for either regressor

c) $SS_R(\beta_2|\beta_1,\beta_0) = 1921.21$

$H_0: \beta_2 = 0$

$H_1: \beta_2 \neq 0$

$\alpha = 0.05$

$f_0 = 2.96$

$f_{\alpha,1,3} = f_{.05,1,3} = 10.13$

$f_0 \ngtr f_{\alpha,1,3}$

Do not reject H_0

11-21. a) $H_0: \beta_1 = \beta_2 = \beta_{12} = 0$

H_1 at least one $\beta_j \neq 0$

$\alpha = 0.05$

$f_0 = 7.714$

$f_{\alpha,3,2} = f_{.05,3,2} = 19.16$

$f_0 \ngtr f_{\alpha,3,2}$

Do not reject H_0

b) $H_0: \beta_{12} = 0$

$H_1: \beta_{12} \neq 0$

$\alpha = 0.05$

$f_0 = 1.47$

$f_{.05,1,2} = 18.51$

$f_0 \ngtr f_{\alpha,1,2}$

Do not reject H_0

c) $\hat{\sigma}^2 = 561.283$

$\hat{\sigma}^2$ (no interaction term) = 650.141

$MS_E(\hat{\sigma}^2)$ was reduced in the interaction term model due to the addition of this term.

11-22. a) $H_0: \beta_j = 0$ for all j

$H_1: \beta_j \neq 0$ for at least one j

$f_0 = 7.16$

$f_{.05,4,14} = 3.11$

$f_0 > f_{\alpha,4,14}$

Reject H_0 and conclude regression is significant at $\alpha = 0.05$. P-value = 0.0023

b) $\hat{\sigma} = 0.7792$

c) $\alpha = 0.05$ $t_{\alpha/2,n-p} = t_{.025,14} = 2.145$

$H_0: \beta_2 = 0$	$\beta_3 = 0$	$\beta_4 = 0$	$\beta_5 = 0$
$H_1: \beta_2 \neq 0$	$\beta_3 \neq 0$	$\beta_4 \neq 0$	$\beta_5 \neq 0$
$t_0 = -0.113$	$t_0 = 3.83$	$t_0 = -1.91$	$t_0 = -0.9$
$\lvert t_0\rvert \ngtr t_{\alpha/2,14}$	$\lvert t_0\rvert > t_{\alpha/2,14}$	$\lvert t_0\rvert \ngtr t_{\alpha/2,14}$	$\lvert t_0\rvert \ngtr t_{\alpha/2,14}$
Do not reject H_0	Reject H_0	Do not reject H_0	Do not reject H_0

NO.

11-23. a) $H_0: \beta_j = 0$ for all j

$H_1: \beta_j \neq 0$ for at least one j

$f_0 = 840.55$

$f_{.05,3,16} = 3.24$

$f_0 > f_{\alpha,3,16}$

Reject H_0 and conclude regression is significant at $\alpha = 0.05$

b) $\hat{\sigma}^2 = 12.1078$

c) $\alpha = 0.05$ $t_{\alpha/2,n-p} = t_{.025,16} = 2.12$

$H_0: \beta_1 = 0$	$\beta_2 = 0$	$\beta_3 = 0$
$H_1: \beta_1 \neq 0$	$\beta_2 \neq 0$	$\beta_3 \neq 0$
$t_0 = -2.637$	$t_0 = 1.91$	$t_0 = 13.9$
$\|t_0\| > t_{\alpha/2,16}$	$\|t_0\| \not> t_{\alpha/2,16}$	$\|t_0\| > t_{\alpha/2,16}$
Reject H_0	Do not reject H_0	Reject H_0

11-24. a) $H_0: \beta_j = 0$ for all j

$H_1: \beta_j \neq 0$ for at least one j

$f_0 = 162.43$

$f_{.05,5,26} = 2.59$

$f_0 > f_{\alpha,5,26}$

Reject H_0 and conclude regression is significant at $\alpha = 0.05$

P-value < 0.000001

b) $\hat{\sigma}^2 = 5.162E-6$

c) $\alpha = 0.05$ $t_{\alpha/2,n-p} = t_{.025,26} = 2.056$

$H_0: \beta_{temp} = 0$	$\beta_{soaktime} = 0$	$\beta_{soakpct} = 0$	$\beta_{DFtime} = 0$	$\beta_{Diffpct} = 0$
$H_1: \beta_{temp} \neq 0$	$\beta_{soaktime} \neq 0$	$\beta_{soakpct} \neq 0$	$\beta_{DFtime} \neq 0$	$\beta_{Diffpct} \neq 0$
$t_0 = 1.21$	$t_0 = 11.79$	$t_0 = 0.91$	$t_0 = 5.77$	$t_0 = -0.39$
Do not Reject H_0	Do not Reject H_0	Do not reject H_0	Reject H_0	Reject H_0

d) $\hat{y} = 0.010918 + 0.002691x_1 + 0.009356x_2$

e) $H_0: \beta_j = 0$ for all j

$H_1: \beta_j \neq 0$ for at least one j

$f_0 = 332.225$

$f_{.05,2,29} = 3.33$

$f_0 > f_{\alpha,2,29}$

Reject H_0 and conclude regression is significant at $\alpha = 0.05$

$\alpha = 0.05$ $t_{\alpha/2,n-p} = t_{.025,29} = 2.045$

$H_0: \beta_1 = 0$	$\beta_2 = 0$
$H_1: \beta_1 \neq 0$	$\beta_2 \neq 0$
$t_0 = 18.97$	$t_0 = 6.62$
$\|t_0\| > t_{\alpha/2,29}$	$\|t_0\| > t_{\alpha/2,29}$

Reject H_0 for either regressor variable and conclude that both variables are significant at $\alpha = 0.05$

f) $\hat{\sigma}_{part\ d.} = 6.24E-6$. Part b. is smaller, suggesting a better model.

11-25. a) $H_0: \beta_j = 0$ for all j

$H_1: \beta_j \neq 0$ for at least one j

$f_0 = 101.79$

$f_{.05,9,11} = 2.90$

$f_0 > f_{\alpha,9,11}$

Reject H_0 and conclude regression is significant at $\alpha = 0.05$

b) $H_0: \beta_j = 0$

$H_1: \beta_j \neq 0$

$t_{.025,11} = 2.201$

PTS :	$t_0 = 10.28$	Reject H_0
GF :	$t_0 = 0.29$	Do not reject H_0
GA :	$t_0 = 0.11$	Do not reject H_0
PPG :	$t_0 = 0.74$	Do not reject H_0
PPcT :	$t_0 = -0.79$	Do not reject H_0
SHG :	$t_0 = -0.50$	Do not reject H_0
PPGA :	$t_0 = 0.81$	Do not reject H_0
PKPcT :	$t_0 = 0.27$	Do not reject H_0
SHGA :	$t_0 = -1.88$	Do not reject H_0

NO, only "PTS" is significant at $\alpha = 0.05$

c)

$\hat{y} = -5.531 + 0.497x_{PTS} - 0.004x_{PPG}$

$f_0 = 510.12$

$f_{.05,2,18} = 3.55$

Reject H_0

$H_0: \beta_{PTS} = 0$ $\beta_{PPG} = 0$

$H_1: \beta_{PTS} \neq 0$ $\beta_{PPG} \neq 0$

$t_0 = 30.60$ $t_0 = -0.165$

Reject H_0 Do not reject H_0

Sections 11-6 and 11-7

11-26. a) $\hat{\beta}_1 \pm t_{\alpha/2,n-p}\sqrt{\hat{\sigma}^2 c_{11}}$

$0.0931 \pm t_{.025,7}\ se(\hat{\beta}_1)$

$0.0931 \pm (2.365)(0.0254)$

0.0931 ± 0.060071

$0.0330 \leq \beta_1 \leq 0.1532$

b) $x_1 = 200$

$x_2 = 50$

$\hat{y}_0 = 29.37$

$X_0'(X'X)^{-1}X_0 = 0.211088$

$29.37 \pm (2.365)\sqrt{8.0643(0.211088)}$

29.37 ± 3.08565

$26.28 \leq \mu_{Y|x_0} \leq 32.46$

c) $\alpha = 0.05$

$x_1 = 200$

$x_2 = 50$

$\hat{y}_0 = 29.37$

$X_0'(X'X)^{-1}X_0 = 0.211088$

$\hat{y} \pm t_{\alpha/2,n-p}\sqrt{\hat{\sigma}^2(1+X_0'(X'X)^{-1}X_0)}$

$29.37 \pm (2.365)\sqrt{8.0643(1.211088)}$

29.37 ± 7.391

$21.98 \le y_0 \le 36.76$

11-27. a) $-0.00657 \le \beta_8 \le -0.00122$

b) $\sqrt{\hat{\sigma}^2 X_0'(X'X)^{-1}X_0} = 0.497648 = se(\hat{\mu}_{Y|x_0})$

c) $\hat{\mu}_{Y|x_0} = -7.63449 + 0.00398(2000) + 0.24777(60) - 0.00389(1800) = 8.19$

$\hat{\mu}_{Y|x_0} \pm t_{.025,24}\ se(\hat{\mu}_{Y|x_0})$

$8.19 \pm (2.064)(0.497648)$

8.19 ± 1.03

$7.16 \le \mu_{Y|x_0} \le 9.22$

11-28. a) $-0.07219 \le \beta_1 \le -0.03651$

$-0.89430 \le \beta_6 \le 3.05073$

b) $\hat{\mu}_{Y|x_0} = 21.4572$ $\quad se(\hat{\mu}_{Y|x_0}) = 1.08785$

$\hat{\mu}_{Y|x_0} \pm t_{.005,22}\ se(\hat{\mu}_{Y|x_0})$

$21.4572 \pm (2.819)(1.08785)$

$18.391 \le \mu_{Y|x_0} \le 24.524$

c) $\hat{y} = 32.0223 - 0.064x_1 + 0.01269x_2 + 1.06181x_6 + 0.00069x_{10}$

$$\left.\begin{array}{l} -0.12973 \le \beta_1 \le 0.00174 \\ -0.13236 \le \beta_2 \le 0.15774 \\ -1.76898 \le \beta_6 \le 3.89259 \\ -0.00575 \le \beta_{10} \le 0.00713 \end{array}\right\} \quad \alpha = 0.01$$

d) Intervals of large model are wider (part c. are wider). YES, adding x_2 and x_{10} did not help the significance of the model, the standard errors increased.

11-29. a) $-0.26718 \le \beta_1 \le 1.47793$

$-3.61373 \le \beta_2 \le 21.4610$

$-4.21947 \le \beta_3 \le 7.09438$

$-1.72211 \le \beta_4 \le 1.74932$

b) $\hat{\mu}_{Y|x_0} = 287.562$ $\quad se(\hat{\mu}_{Y|x_0}) = 10.054$ $\quad t_{.025,7} = 2.365$

$\hat{\mu}_{Y|x_0} \pm t_{\alpha/2,n-p}\ se(\hat{\mu}_{Y|x_0})$

$287.562 \pm (2.365)(10.054)$

$263.78 \le \mu_{Y|x_0} \le 311.34$

c) $\hat{y}_0 \pm t_{\alpha/2,n-p}\sqrt{\hat{\sigma}^2(1+X_0'(X'X)^{-1}X_0}$

$287.562 \pm 2.365(18.5418)$

$243.711 \le y_0 \le 331.413$

11-30. a) $-8.101 \le \beta_1 \le 5.557$

$-0.677 \le \beta_2 \le 0.369$

b) $-66.825 \le \beta_1 \le 82.342$

$-2.096 \le \beta_2 \le 2.284$

$-0.084 \le \beta_{12} \le 0.066$

These part b. intervals are much wider.
Yes, the addition of this term increased the standard error of the regression coefficient estimators.

11-31. a) $-0.595 \le \beta_2 \le 0.535$

$0.229 \le \beta_3 \le 0.812$

$-0.216 \le \beta_4 \le 0.013$

$-7.298 \le \beta_5 \le 2.977$

b) $\hat{\mu}_{Y|x_0} = 8.99568$ $\quad se(\hat{\mu}_{Y|x_0}) = 0.472445$ $\quad t_{.025,14} = 2.145$

$\hat{\mu}_{Y|x_0} \pm t_{\alpha/2,n-p} \ se(\hat{\mu}_{Y|x_0})$

$8.99568 \pm (2.145)(0.472445)$

$7.982 \le \mu_{Y|x_0} \le 10.009$

c) $\hat{y}_0 = 8.99568$ $\quad se(\hat{y}_0) = 1.00121$

$8.99568 \pm 2.145(1.00121)$

$6.8481 \le y_0 \le 11.143$

11-32. a) $-20.520 \le \beta_1 \le 1.049$

$-0.226 \le \beta_2 \le 1.082$

$14.405 \le \beta_3 \le 22.070$

b) $\hat{y}_0 = 91.424$ $\quad se(\hat{y}_0) = 4.68262$ $\quad t_{.005,16} = 2.921$

$91.424 \pm 2.921(4.68262)$

$77.746 \le y_0 \le 105.102$

c) $\hat{\mu}_{Y|x_0} = 91.424$ $\quad se(\hat{\mu}_{Y|x_0}) = 3.13355$

$91.424 \pm (2.921)(3.13355)$

$82.271 \le \mu_{Y|x_0} \le 100.577$

11-33. a) $-0.00003 \le \beta_{Temp} \le 0.00012$

$0.00203 \le \beta_{soaktime} \le 0.00288$

$-0.02306 \le \beta_{soakpct} \le 0.05976$

$0.00501 \le \beta_{DFtime} \le 0.01056$

$-0.01969 \le \beta_{DiffPct} \le 0.01342$

b) $\hat{\mu}_{Y|x_0} = 0.0220086$ $\quad se(\hat{\mu}_{Y|x_0}) = 6.714E-4$ $\quad t_{.025,26} = 2.056$

$0.0220086 \pm (2.056)(6.714E-4)$

$2.06E-2 \le \mu_{Y|x_0} \le 2.34E-2$

11-34. a) $\hat{\mu}_{Y|x_0} = 0.0214$ $\quad se(\hat{\mu}_{Y|x_0}) = 5.21E-4$ $\quad t_{.025,29} = 2.045$

$0.0214 \pm (2.045)(5.21E-4)$

$2.03E-2 \le \mu_{Y|x_0} \le 2.25E-2$

b) : width = 2.2 E-3
: width = 2.8 E-3
The interaction model has a shorter confidence interval. Yes, this suggests the interaction model is preferable.

11-35. a) $0.3882 \leq \beta_{Pts} \leq 0.5998$

b) $\hat{y} = -5.767703 + 0.496501x_{Pts}$

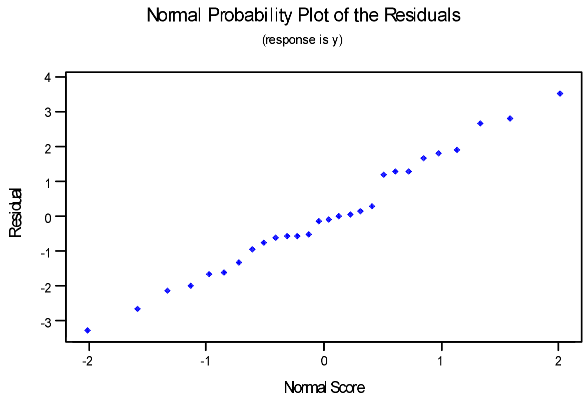
c) $0.4648 \leq \beta_{Pts} \leq 0.5282$

d) The simple linear regression model has the shorter interval. YES, the simple model in this case is preferable.

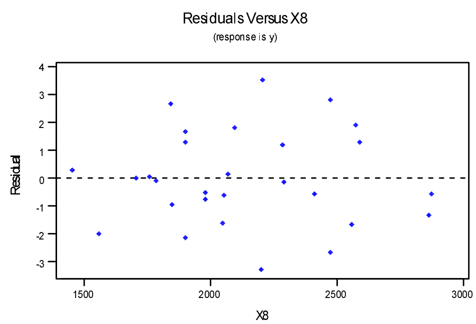
Section 11-8

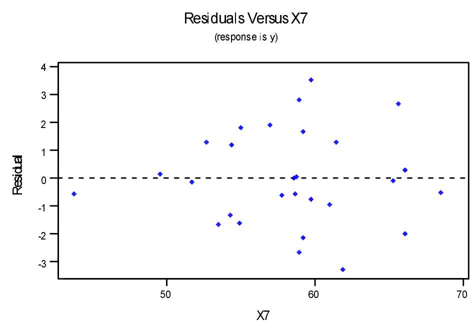
11-36. a) $r^2 = 0.762938$ or 76.29 %

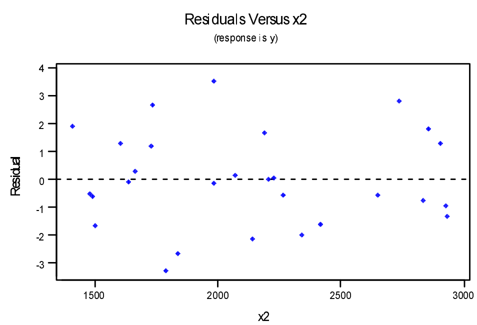
b) Assumption of normality appears adequate.

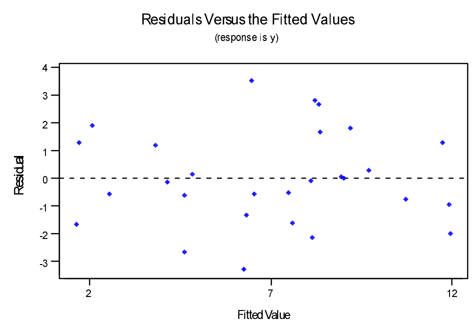


c) Model appears adequate. Some suggestion of nonconstant variance in the plot of x_7.





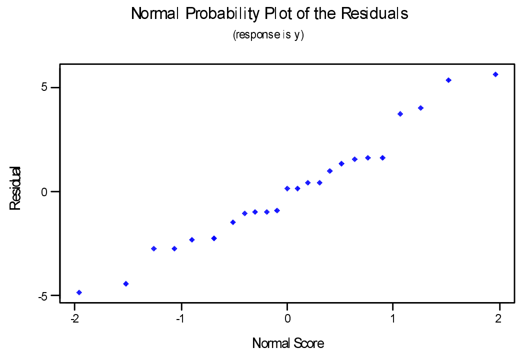




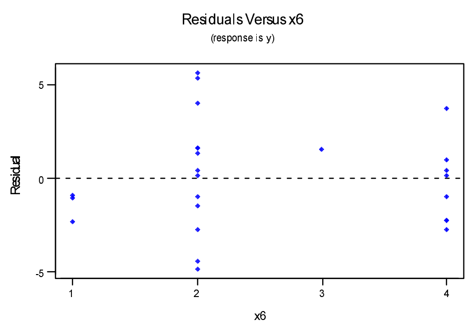
d) Yes, Observations 1, 8, and 21.

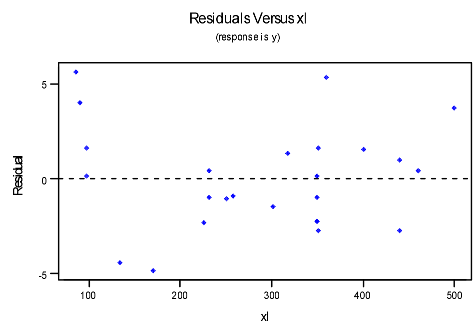
11-37. a) $r^2 = 0.82897$

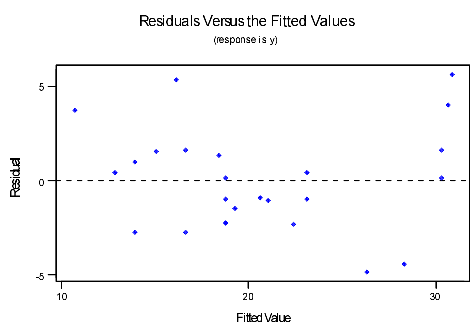
b) Normality assumption appears valid.



c) Assumption of constant variance appears reasonable.



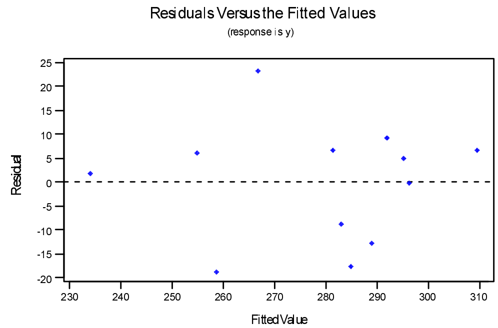




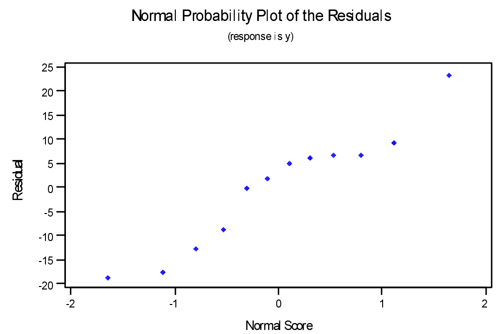
d) Yes, observations 7, 10, and 18

11-38. a) $r^2 = 0.74475$

b) The residual plots look reasonable. This is some increase in variability at the middle of the predicted values.



c) Normality assumption is reasonable.

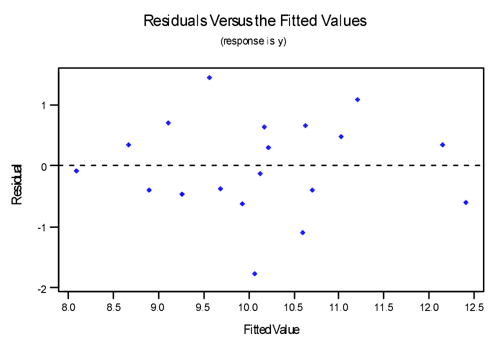


11-39. a) $r^2 = 0.86179$

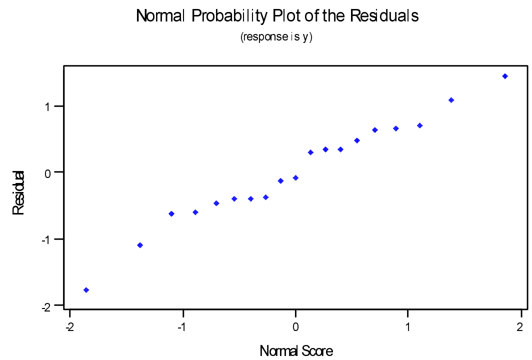
b) $r^2 = 0.920453$

r^2 increases with addition of interaction term. No, adding additional regressor will always increase r^2

11-40. a) Some indication of variability increasing with predicted value.

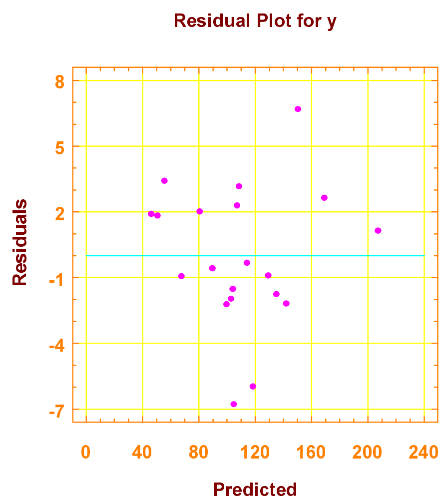


b) No, the normality assumption appears reasonable.



c) Yes, observation number 2.

11-41. a) There is some indication of nonconstant variance since the residuals appear to “fan out” with increasing values of y.



b)

Source	Sum of Squares	DF	Mean Square	F-Ratio	P-value
Model	30531.5	3	10177.2	840.546	.0000
Error	193.725	16	12.1078		
Total (Corr.)	30725.2	19			

```
R-squared = 0.993695                              Stnd. error of est. = 3.47963
R-squared (Adj. for d.f.) = 0.992513           Durbin-Watson statistic = 1.77758
```

$R^2 = 0.9937$ or 99.37 %;

$R^2_{Adj} = 0.9925$ or 99.25%;

c)

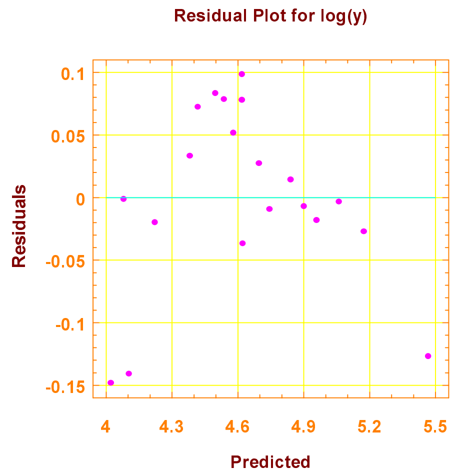
Model fitting results for: log(y)

Independent variable	coefficient	std. error	t-value	sig.level
CONSTANT	6.22489	1.124522	5.5356	0.0000
x1	-0.16647	0.083727	-1.9882	0.0642
x2	-0.000228	0.005079	-0.0448	0.9648
x3	0.157312	0.029752	5.2875	0.0001

```
R-SQ. (ADJ.) = 0.9574  SE=        0.078919  MAE=        0.053775  DurbWat=  2.031
Previously:    0.0000             0.000000              0.000000             0.000
20 observations fitted, forecast(s) computed for 0 missing val. of dep. var.
```

$\hat{y}^* = 6.22489 - 0.16647x_1 - 0.000228x_2 + 0.157312x_3$

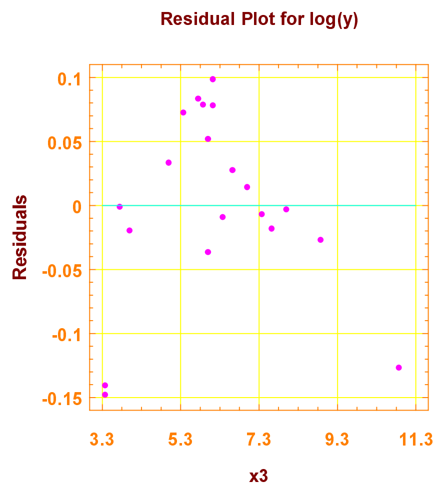
d)



Plot exhibits curvature

There is curvature in the plot. The plot does not give much more information as to which model is preferable.

e)

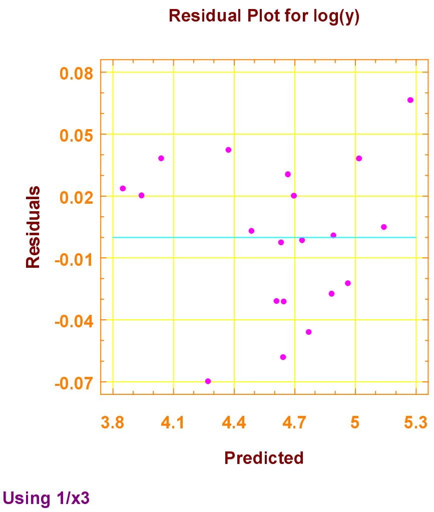


Plot exhibits curvature

Variance does not appear constant. Curvature is evident.

f)

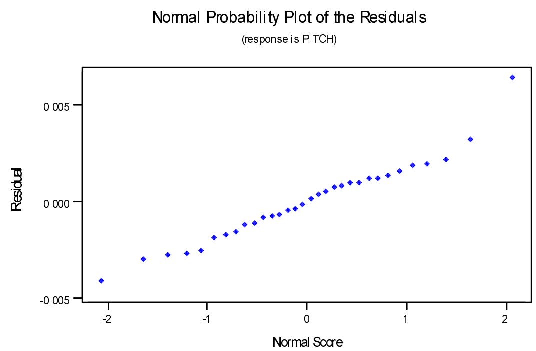
```
                    Model fitting results for: log(y)
Independent variable             coefficient  std. error     t-value   sig.level
CONSTANT                            6.222045    0.547157     11.3716      0.0000
x1                                 -0.198597    0.034022     -5.8374      0.0000
x2                                  0.009724    0.001864      5.2180      0.0001
1/x3                               -4.436229    0.351293    -12.6283      0.0000
--------------------------------------------------------------------------------
R-SQ. (ADJ.) = 0.9893  SE=       0.039499  MAE=       0.028896  DurbWat=  1.869
Previously:    0.9574            0.078919             0.053775             2.031
20 observations fitted, forecast(s) computed for 0 missing val. of dep. var.

                 Analysis of Variance for the Full Regression
Source                 Sum of Squares     DF    Mean Square      F-Ratio   P-value
Model                         2.75054      3       0.916847      587.649     .0000
Error                       0.0249631     16     0.00156020
Total (Corr.)                 2.77550     19
R-squared = 0.991006                             Stnd. error of est. = 0.0394993
R-squared (Adj. for d.f.) = 0.98932            Durbin-Watson statistic = 1.86891
```



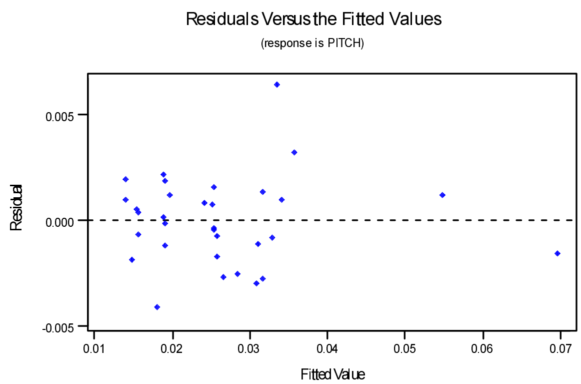
The residual plot indicates better conformance to assumptions.

Curvature is removed when using $1/x_3$ as the regressor instead of x_3 and the log of the data.

11-42. a) 0.969
b) Normality is acceptable



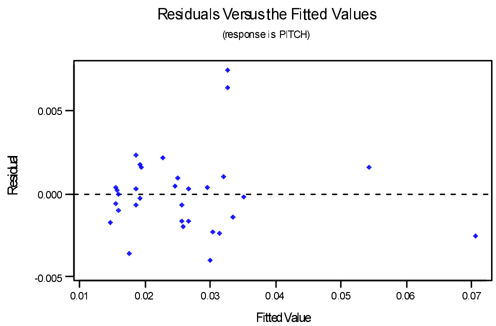
c) Plot is acceptable.

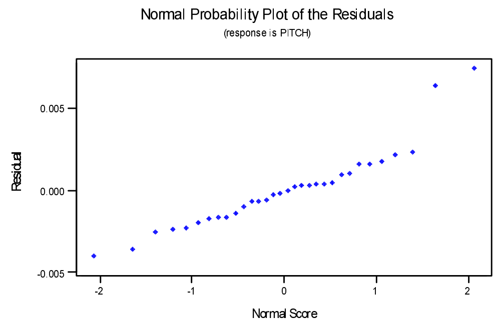


d) None of the points appear to be influential.

11-43. a) $R^2 = 0.9582$; Yes, the R^2 using these two regressors is nearly as large as the R2 from the model with five regressors.

b) Normality is acceptable, but there is some indication of outliers.



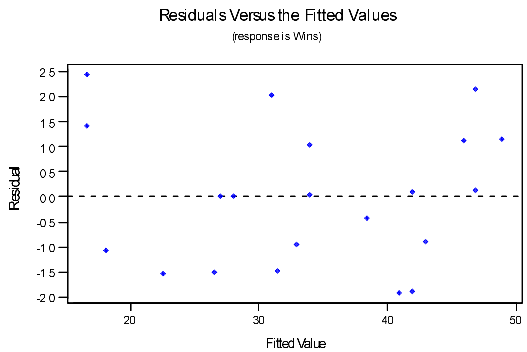


c) The last observation is very influential

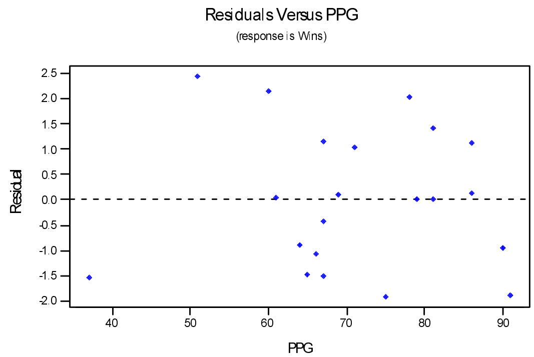
11-44. a) $\hat{y} = -5.767703 + 0.496501x_1$

b) $r^2 = 0.982637$ or 98.26 %

c) Model appears adequate.



d) No, the residuals do not seem to be related to PPG. If PPG were significant, we would detect some nonrandom pattern in this plot.



11-45. a) $p = k + 1 = 2 + 1 = 3$

Average size = $p/n = 3/25 = 0.12$

b) Leverage point criteria:

$h_{ii} > 2(p/n)$

$h_{ii} > 2(0.12)$

$h_{ii} > 0.24$

$h_{17,17} = 0.2593$

$h_{18,18} = 0.2929$

Points 17 and 18 are leverage points.

Sections 11-9 and 11-10

11-46. a) $\hat{y} = -26219.15 + 189.205x - 0.33x^2$

b) $H_0: \beta_j = 0$ for all j

$H_1: \beta_j \neq 0$ for at least one j

$\alpha = 0.05$

$f_0 = 17.2045$

$f_{.05,2,5} = 5.79$

$f_0 > f_{.05,2,5}$

Reject H_0 and conclude that model is significant at $\alpha = 0.05$

c) $H_0: \beta_{11} = 0$

$H_1: \beta_{11} \neq 0$

$\alpha = 0.05$

$t_0 = -2.45$

$t_{\alpha,n-p} = t_{.025,8-3} = t_{.025,5} = 2.571$

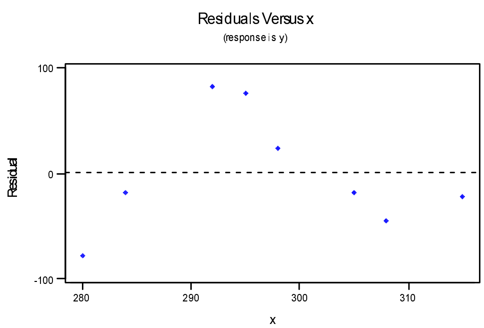
$|t_0| \ngtr 2.571$

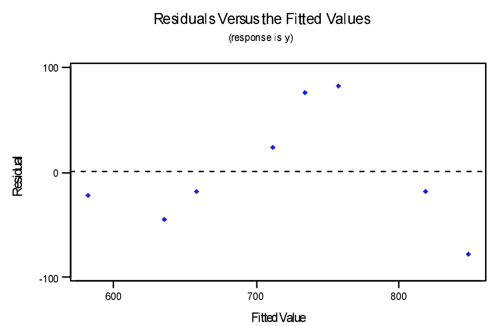
Do not reject H_0 and conclude insufficient evidence to support value of quadratic term in model at $\alpha = 0.05$

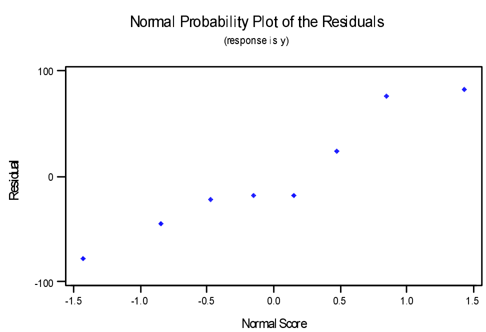
d) One residual is outlier

Normality assumption appears acceptable

Residuals against predicted and x are somewhat unusual plots, but the impact of the outlier should be considered.





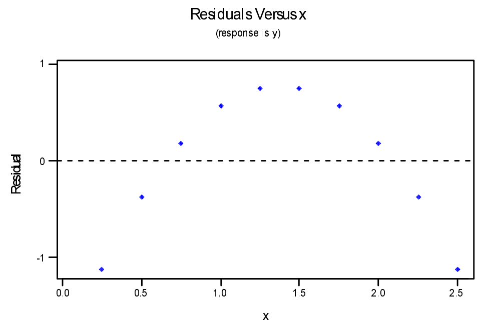


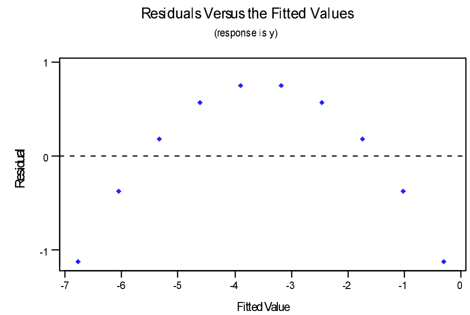
11-47. a) $\hat{y} = -1.633 + 1.232x - 1.495x^2$

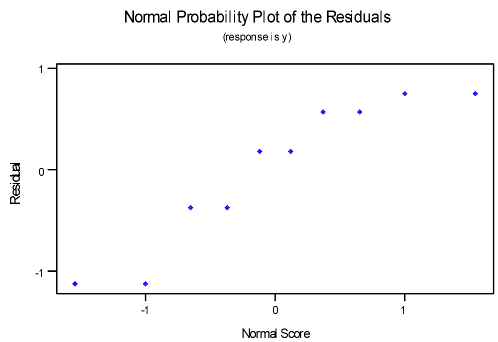
b) $f_0 = 1858613$, reject H_0

c) $t_0 = -601.64$, reject H_0

d) Model is acceptable, observation number 10 has large leverage.







11-48.

$$\hat{y} = \beta_0^* + \beta_1^* x' + \beta_{11}^* (x')^2$$

$$\hat{y} = 759.395 - 7.607x' - 0.331(x')^2$$

$$\hat{y} = 759.395 - 7.607(x - 297.125) - 0.331(x - 297.125)^2$$

$$\hat{y} = -26202.14 + 189.09x - 0.331x^2$$

11-49. $\hat{y} = 759.395 - 90.783x' - 47.166(x')^2$, where $x' = \dfrac{x - \bar{x}}{S_x}$

a) At x = 285 $x' = \dfrac{285 - 297.125}{11.9336} = -1.016$

$$\hat{y} = 759.395 - 90.783(-1.106) - 47.166(-1.106)^2 = 802.943 \text{ psi}$$

b) $\hat{y} = 759.395 - 90.783\left(\dfrac{x - 297.125}{11.9336}\right) - 47.166\left(\dfrac{x - 297.125}{11.9336}\right)^2$

$$\hat{y} = 759.395 - 7.607(x - 297.125) - 0.331(x - 297.125)^2$$

$$\hat{y} = -26204.14 + 189.09x - 0.331x^2$$

c) They are the same.

d) $\hat{y}' = 0.385 - 0.847x' - 0.440(x')^2$

where $y' = \dfrac{y - \bar{y}}{S_y}$ and $x' = \dfrac{x - \bar{x}}{S_x}$

The "proportion" of S_{yy} allocated to SS_E is the same for both standardized and unstandardized models.

Therefore, R^2 is the same for both models.

$y' = \beta_0^* + \beta_1^* x' + \beta_{11}^* (x')^2$ where $y' = \dfrac{y - \bar{y}}{S_y}$ and $x' = \dfrac{x - \bar{x}}{S_x}$

$$y' = \beta_0^* + \beta_1^* x' + \beta_{11}^* (x')^2$$

11-50. a) $\hat{y} = -4.46 + 1.38x + 1.47x^2$

b) $H_0: \beta_j = 0$ for all j

$H_1: \beta_j \neq 0$ for at least one j

$\alpha = 0.05$

$f_0 = 1044.99$

$f_{.05,2,9} = 4.26$

$f_0 > f_{\alpha,2,9}$

Reject H_0 and conclude regression model is significant at $\alpha = 0.05$

c) $H_0: \beta_{11} = 0$

$H_1: \beta_{11} \neq 0$ $\alpha = 0.05$

$t_0 = 2.97$

$t_{.025,9} = 2.262$

$|t_0| > t_{\alpha/2,9}$

Reject H_0 and conclude that β_{11} is significant at $\alpha = 0.05$

d) Observation number 9 is an extreme outlier.

e) $\hat{y} = -87.36 + 48.01x - 7.04x^2 + 0.51x^3$

$H_0: \beta_{33} = 0$

$H_1: \beta_{33} \neq 0$ $\alpha = 0.05$

$t_0 = 0.91$

$t_{.025,8} = 2.306$

$|t_0| \not> t_{\alpha/2,8}$

Do not reject H_0 and conclude that cubic term is not significant at $\alpha = 0.05$

11-51. a) $\hat{y} = -1.769 + 0.421x_1 + 0.222x_2 - 0.128x_3 - 0.02x_1x_2 + 0.009x_1x_3 +$
$0.003x_2x_3 - 0.019x_1^2 - 0.007x_2^2 + 0.001x_3^2$

b) H_0 all $\beta_j = 0$

H_1 not all $\beta_j = 0$

$f_0 = 19.628$

$f_{.05,9,16} = 2.54$

$f_0 > f_{\alpha,9,16}$

Reject H_0 and conclude that the model is significant at $\alpha = 0.05$

c) Model is acceptable.

d) $H_0: \beta_{11} = \beta_{22} = \beta_{33} = \beta_{12} = \beta_{13} = \beta_{23} = 0$

H_1 at least one $\beta_{ij} \neq 0$

$$f_0 = \frac{SS_R(\beta_{11}\beta_{22}\beta_{33}\beta_{12}\beta_{13}\beta_{23}|\beta_1\beta_2\beta_3\beta_0)/r}{MS_E} = \frac{\frac{0.0359}{6}}{0.003712} = 1.612$$

$f_{.05,6,16} = 2.74$

$f_0 \not> f_{.05,6,16}$

Do not reject H_0

$$SS_R(\beta_{11}\beta_{22}\beta_{33}\beta_{12}\beta_{13}\beta_{23}|\beta_1\beta_2\beta_3\beta_0) = SS_R(\beta_{11}\beta_{22}\beta_{33}\beta_{12}\beta_{13}\beta_{23}\beta_1\beta_2\beta_3|\beta_0) - SS_R(\beta_1\beta_2\beta_3|\beta_0)$$
$$= 0.65567068 - 0.619763$$
$$= 0.0359$$

Reduced Model: $y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3$

11-52. a) Use indicator variable for transmission type

$$x_{11} = \begin{cases} 0 & \text{for manual} \\ 1 & \text{for automatic} \end{cases}$$

b) $\hat{y} = 32.77 - 0.06x_1 + 0.05x_2 + 0.12x_{11}$

c) $H_0: \beta_{11} = 0$

$H_1: \beta_{11} \neq 0 \quad \alpha = 0.05$

$t_0 = 0.05$

$t_{.025,21} = 2.08$

$|t_0| \not> t_{\alpha/2,21}$

Do not reject H_0. No evidence that transmission type affects mileage.

11-53. $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$

$\hat{y} = 11.503 + 0.153x_1 - 6.094x_2 - 0.031x_1x_2$

where $x_2 = \begin{cases} 0 & \text{for tool type 302} \\ 1 & \text{for tool type 416} \end{cases}$

Test of different slopes:

$H_0: \beta_{12} = 0$

$H_1: \beta_{12} \neq 0 \quad \alpha = 0.05$

$t_0 = -1.79$

$t_{.025,16} = 2.12$

$|t_0| \not> t_{\alpha/2,16}$

Do not reject H_0. Conclude that data is insufficient to claim that (2) regression models are needed.
Test of different intercepts and slopes using extra sums of squares:

$H_0: \beta_2 = \beta_{12} = 0$

H_1 at least one is not zero

$$\begin{aligned} SS(\beta_2, \beta_{12} | \beta_0) &= SS(\beta_1, \beta_2, \beta_{12} | \beta_0) - SS(\beta_1 | \beta_0) \\ &= 1013.35995 - 130.60910 \\ &= 882.7508 \end{aligned}$$

$$f_0 = \frac{SS(\beta_2, \beta_{12} | \beta_0)/2}{MS_E} = \frac{882.7508/2}{0.4059} = 1087.40$$

Reject H_0.

Section 11-11

11-54. a) Min. MS_E x_2, x_4, x_7, x_8, x_9

$MS_E = 2.975$

$c_p = 3.126$

$\hat{y} = -10.584 + 0.004x_2 + 0.048x_4 + 0.284x_7 - 0.003x_8 - 0.002x_9$

Min. c_p: x_2, x_7, x_8

$c_p = 2.647$

$MS_E = 3.230$

$\hat{y} = -7.634 + 0.004x_2 + 0.284x_7 - 0.004x_8$

b), c), & d) Same as the minimum c_p model.

e) The minimum c_p model seems the best because stepwise, forward and backward selection all converged to the same model. Also, the model contains fewer variables than the minimum MS_E model.

11-55. a) Min. MS_E $x_1, x_3, x_4, x_5, x_7, x_8, x_{10}$

$MS_E = 6.58$ $c_p = 5.88$

Min. c_p: x_5, x_8, x_{10}

$c_p = 5.02$ $MS_E = 7.97$

b) $\hat{y} = 34.434 - 0.048x_1$

$MS_E = 8.81$ $c_p = 5.55$

c) Same as part b.

d) $\hat{y} = 0.341 + 2.862x_5 + 0.246x_8 - 0.010x_{10}$

$MS_E = 7.97$ $c_p = 5.02$

The same as minimum c_p model

e) Minimum c_p and backward elimination are the same model. Stepwise and forward selection are the same model. Because it is much smaller, the minimum c_p model seems preferable.

11-56. a) Min. MS_E: x_1, x_2

Min. c_p: x_1, x_2

$c_p = 1.37$ $MS_E = 198.66$

$\hat{y} = 0.529 + 0.497x_1 + 10.267x_2$

b) $\hat{y} = -90.161 + 15.161x_2$

$MS_E = 236.58$ $c_p = 1.75$

c) Same as part b.

d) Same as part b.

e) Stepwise, forward and backward models all are the same. Min c_p and MS_E models add in x_1 for a small gain in MS_E and c_p.

11-57. a) $\hat{y} = 4.656 + 0.511x_3 - 0.124x_4$

b) Same as part a.

c) Same as part a.

d) All models are the same.

11-58. a) $\hat{y} = -3.517 + 0.486x_1 - 0.156x_9$

b) Same as part a.

c) Same as part a.

d) All models are the same.

11-59 a) $\hat{y} = -0.304 + 0.083x_1 - 0.031x_3 + 0.004x_2^2$

$c_p = 4.04$ $MS_E = 0.004$

b) $\hat{y} = -0.256 + 0.078x_1 + 0.022x_2 - 0.042x_3 + 0.0008x_3^2$

$c_p = 4.66$ $MS_E = 0.004$

c) The forward selection model is more parsimonious with a lower c_p and equivalent MS_E. Therefore, we prefer model in a.

11-60. Min. c_p: x_3, x_4

$\hat{y} = 4.656 + 0.511x_3 - 0.124x_4$

Adjusted $R^2 = 0.603$. Min. $c_p = 1.65$

Max. adjusted R^2 x_1, x_3, x_4

$\hat{y} = 2.419 + 0.553x_1 + 0.479x_3 - 0.123x_4$

Adjusted $R^2 = 0.607$, $c_p = 2.61$

No, Max. adjusted R^2 has x_1 in addition to x_3 & x_4 in the model. However, the adjusted R^2 values are very close.

11-61. a) Min. c_p:

$\hat{y} = -3.517 + 0.486x_1 - 0.156x_9$

$c_p = -1.67$

b) Min MS_E x_1, x_7, x_9, $MS_E = 1.67$, $c_p = -0.77$

$\hat{y} = -5.964 + 0.495x_1 + 0.025x_7 - 0.163x_9$

c) Max. adjusted R^2 x_1, x_7, x_9, Adj. $R^2 = 0.98448$ YES, same as Min. MS_E model.

11-62. n = 30, k = 9, p =9 + 1 = 10 in full model.

a) $\hat{\sigma}^2 = MS_E = 100$ $R^2 = 0.92$

$$R^2 = \frac{SS_R}{S_{yy}} = 1 - \frac{SS_E}{S_{yy}}$$

$$SS_E = MS_E(n-p) = 100(30-10) = 2000$$

$$0.92 = 1 - \frac{2000}{S_{yy}}$$

$$25000 = S_{yy}$$

$$SS_R = S_{yy} - SS_E = 25000 - 2000 = 23000$$

$$MS_R = \frac{SS_R}{k} = \frac{23000}{9} = 2555.56$$

$$f_0 = \frac{MS_R}{MS_E} = \frac{2555.56}{100} = 25.56$$

$$f_{.05,9,20} = 2.39$$

$$f_0 > f_{\alpha,9,20}$$

Reject H_0 and conclude at least one β_j is significant at $\alpha = 0.05$.

b) k = 4 p = 5 $SS_E = 2200$

$$MS_E = \frac{SS_E}{n-p} = \frac{220}{30-5} = 88$$

YES, MS_E is reduced with new model (k = 4).

c) $c_p = \dfrac{SS_E(p)}{\hat{\sigma}^2} - n + 2p$ $c_p = \dfrac{2200}{100} - 30 + 2(5) = 2$

Yes, c_p is reduced from the full model.

11-63. n = 25 k = 7 p = 8 $MS_{E(\text{full})} = 10$

a) p =4 $SS_E = 300$

$$MS_E = \frac{SS_E}{n-p} = \frac{300}{25-4} = 14.29$$

$$c_p = \frac{SS_E}{MS_{E(\text{full})}} - n + 2p = \frac{300}{10} - 25 + 2(4) = 5 + 8 = 13$$

YES, $c_p > p$

b) p = 5 $SS_E = 275$

$$MS_E = \frac{SS_E}{n-p} = \frac{275}{30-5} = 11 \quad c_p = \frac{275}{10} - 25 + 2(5) = 12.5$$

Yes, both MS_E and c_p are reduced.

Supplemental Exercises

11-64. a) $\hat{y} = 3829.26 - 0.215x_3 + 21.213x_4 + 1.657x_5$

b) $H_0: \beta_3 = \beta_4 = \beta_5 = 0$

$H_1: \beta_j \neq 0$ for at least one j

$\alpha = 0.01$

$f_0 = 1724.42$

$f_{.01,3,36} = 4.38$

Reject H_0 and conclude regression is significant.

P-value < 0.00001

c) All at $\alpha = 0.01$ $t_{.005,36} = 2.72$

$H_0: \beta_3 = 0$	$H_0: \beta_4 = 0$	$H_0: \beta_5 = 0$
$H_1: \beta_3 \neq 0$	$H_1: \beta_4 \neq 0$	$H_1: \beta_5 \neq 0$
$t_0 = -1.97$	$t_0 = 23.44$	$t_0 = 3.01$
$\lvert t_0 \rvert \not> t_{\alpha/2,36}$	$\lvert t_0 \rvert > t_{\alpha/2,36}$	$\lvert t_0 \rvert > t_{\alpha/2,36}$
Do not reject H_0	Reject H_0	Reject H_0

d) $R^2 = 0.993$ adj. $R^2 = 0.9925$

e) Normality assumption appears reasonable.

f) Plot is satisfactory.

g) Slight indication that variance increases as x_3 increases.

h) $\hat{y} = 3829.26 - 0.215(1670) + 21.213(170) + 1.657(1589) = 9709.39$

11-65. a) $H_0: \beta_3^* = \beta_4 = \beta_5 = 0$

$H_1: \beta_j \neq 0$ for at least one j

$\alpha = 0.01$

$f_0 = 1323.62$

$f_{.01,3,36} = 4.38$

$f_0 >> f_{\alpha,3,36}$

Reject H_0 and conclude regression is significant.

P-value < 0.00001

b) $\alpha = 0.01$ $t_{.005,36} = 2.72$

$H_0: \beta_3^* = 0$	$H_0: \beta_4 = 0$	$H_0: \beta_5 = 0$
$H_1: \beta_3^* \neq 0$	$H_1: \beta_4 \neq 0$	$H_1: \beta_5 \neq 0$
$t_0 = -1.32$	$t_0 = 19.97$	$t_0 = 2.48$
$\lvert t_0 \rvert \not> t_{\alpha/2,36}$	$\lvert t_0 \rvert > t_{\alpha/2,36}$	$\lvert t_0 \rvert \not> t_{\alpha/2,36}$
Do not reject H_0	Reject H_0	Do not reject H_0

c) Curvature is evident in the residuals plots from this model.

11-66. a) $\hat{y} = 2.855 - 0.290x_1 + 0.206x_2 + 0.454x_3 - 0.594x_4 + 0.005x_5$

$H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$

$H_1: \beta_j \neq 0$ for at least one j

$\alpha = 0.01$

$f_0 = 4.81$

$f_{.01,5,19} = 4.17$

$f_0 > f_{\alpha,5,19}$

Reject H_0. P-value = 0.0052.

b) $\alpha = 0.05$ $\quad t_{.025,19} = 2.093$

$H_0:\beta_1 = 0$	$H_0:\beta_2 = 0$	$H_0:\beta_3 = 0$	$H_0:\beta_4 = 0$	$H_0:\beta_5 = 0$
$H_1:\beta_1 \neq 0$	$H_1:\beta_2 \neq 0$	$H_1:\beta_3 \neq 0$	$H_1:\beta_4 \neq 0$	$H_1:\beta_5 \neq 0$
$t_0 = -2.47$	$t_0 = 2.74$	$t_0 = 2.42$	$t_0 = -2.80$	$t_0 = 0.26$
$\|t_0\| > t_{\alpha/2,19}$	$\|t_0\| > t_{\alpha/2,19}$	$\|t_0\| > t_{\alpha/2,19}$	$\|t_0\| > t_{\alpha/2,19}$	$\|t_0\| \ngtr t_{\alpha/2,19}$
Reject H_0	Reject H_0	Reject H_0	Reject H_0	Do not reject H_0

c) $\hat{y} = 3.148 - 0.290x_1 + 0.199x_2 + 0.455x_3 - 0.609x_4$

$H_0:\beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$

$H_1:\beta_j \neq 0$ for at least one j

$\alpha = 0.05$

$f_0 = 6.284$

$f_{.05,4,20} = 2.87$

$f_0 > f_{\alpha,4,20}$

Reject H_0.

$\alpha = 0.05$ $\quad t_{.025,20} = 2.086$

$H_0:\beta_1 = 0$	$H_0:\beta_2 = 0$	$H_0:\beta_3 = 0$	$H_0:\beta_4 = 0$
$H_1:\beta_1 \neq 0$	$H_1:\beta_2 \neq 0$	$H_1:\beta_3 \neq 0$	$H_1:\beta_4 \neq 0$
$t_0 = -2.53$	$t_0 = 2.89$	$t_0 = 2.49$	$t_0 = -3.05$
$\|t_0\| > t_{\alpha/2,20}$	$\|t_0\| > t_{\alpha/2,20}$	$\|t_0\| > t_{\alpha/2,20}$	$\|t_0\| > t_{\alpha/2,20}$
Reject H_0	Reject H_0	Reject H_0	Reject H_0

d) The addition of the 5^{th} regressor causes loss in one degree of freedom in the denominator and the reduction in SS_E is not enough to compensate for this loss.

e) Observation 2 is usually large.

f) R^2 for model in a. : 0.5584 R^2 for model in c. : 0.5569, R^2 for model without obs. #2 : 0.8044

R^2 has increased because observation 2 was not fit well by either of the previous models.

g) $H_0:\beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$

$H_1:\beta_j \neq 0$ for at least one j

$\alpha = 0.05$

$f_0 = 19.53$

$f_{.05,4,19} = 2.90$

$f_0 > f_{\alpha,4,19}$

Reject H_0.

$\alpha = 0.05$ $\quad t_{.025,19} = 2.093$

$H_0:\beta_1 = 0$	$H_0:\beta_2 = 0$	$H_0:\beta_3 = 0$	$H_0:\beta_4 = 0$
$H_1:\beta_1 \neq 0$	$H_1:\beta_2 \neq 0$	$H_1:\beta_3 \neq 0$	$H_1:\beta_4 \neq 0$
$t_0 = -3.96$	$t_0 = 6.43$	$t_0 = 3.64$	$t_0 = -3.39$
$\|t_0\| > t_{\alpha/2,19}$	$\|t_0\| > t_{\alpha/2,19}$	$\|t_0\| > t_{\alpha/2,19}$	$\|t_0\| > t_{\alpha/2,19}$
Reject H_0	Reject H_0	Reject H_0	Reject H_0

h) There is some indication of curvature.

11-67. a) $\hat{y} = -0.908 + 5.482x_1^* + 1.126x_2^* - 3.920x_3^* - 1.143x_4^*$

b) $H_0:\beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$

$H_1:\beta_j \neq 0$ for at least one j

$\alpha = 0.05$

$f_0 = 109.02$

$f_{.05,4,19} = 2.90$

$f_0 >> f_{\alpha,4,19}$

Reject H_0 and conclude regression is significant at $\alpha = 0.05$.

$\alpha = 0.05$ $t_{.025,19} = 2.093$

$H_0: \beta_1 = 0$	$H_0: \beta_2 = 0$	$H_0: \beta_3 = 0$	$H_0: \beta_4 = 0$
$H_1: \beta_1 \neq 0$	$H_1: \beta_2 \neq 0$	$H_1: \beta_3 \neq 0$	$H_1: \beta_4 \neq 0$
$t_0 = 11.27$	$t_0 = 14.59$	$t_0 = -6.98$	$t_0 = -8.11$
$\lvert t_0 \rvert > t_{\alpha/2,19}$	$\lvert t_0 \rvert > t_{\alpha/2,19}$	$\lvert t_0 \rvert > t_{\alpha/2,19}$	$\lvert t_0 \rvert > t_{\alpha/2,19}$
Reject H_0	Reject H_0	Reject H_0	Reject H_0

c) The residual plots are more satisfactory than in Exercise 11-66.

11-68. a) $\hat{y} = -1709.405 + 2.023x - 0.0006x^2$

b) $H_0: \beta_1 = \beta_{11} = 0$

$H_1: \beta_j \neq 0$

$\alpha = 0.05$

$f_0 = 300.11$

$f_{.05,2,7} = 4.74$

$f_0 >> f_{\alpha,2,7}$

Reject H_0.

c) $H_0: \beta_{11} = 0$

$H_1: \beta_{11} \neq 0$

$\alpha = 0.05$

$$F_0 = \frac{SS_R(\beta_{11}|\beta_1) / r}{MS_E} = \frac{2.4276 / 1}{0.04413}$$

$f_0 = 55.01$

$f_{.05,1,7} = 5.59$

$f_0 >> f_{\alpha,1,7}$

Reject H_0.

d) Some indication of nonconstant variance.

e) Normality assumption is reasonable.

11-69. a) $\hat{y} = -3982.1 + 1.0964x_1 + 0.1843x_3 + 3.7456x_4 + 0.8343x_5 - 16.2781x_6$

$MS_E(p) = 694.93$ $c_p = 5.62$

b) $\hat{y} = -4280.2 + 1.442x_1 + 0.209x_3 + 0.6467x_5 - 17.5103x_6$

$MS_E(p) = 714.20$ $c_p = 5.57$

c) Same as model b.

d) Models from parts b. and c. are identical. Model in part a. is the same with x_4 added in.

MS_E model in part a. = 694.93 $c_p = 5.62$

MS_E model in part b.&c. = 714.20 $c_p = 5.57$

11-70 a) $\hat{y}^* = -0.908 + 5.482x_1^* + 1.126x_2^* - 3.920x_3^* - 1.143x_4^*$

$MS_E(p) = 0.0797$ Min $c_p = 5.0$

b) Same as part a.

c) Same as part a.

d) All models are the same.

11-71. a) $VIF(\hat{\beta}_3^*) = 51.86$

$VIF(\hat{\beta}_4) = 9.11$

$VIF(\hat{\beta}_5) = 28.99$

Yes, VIFs for x_3^* and x_5 exceed 10.

b) Model from Exercise 11-65: $\hat{y} = 19.69 - 1.27x_3^* + 0.005x_4 + 0.0004x_5$

11-72. a) $\hat{y} = 300.0 + 0.85x_1 + 10.4x_2$

$$\hat{y} = 300 + 0.85(100) + 10.4(2) = 405.8$$

b) $S_{yy} = 1230.5 \qquad SS_E = 120.3$

$$SS_R = S_{yy} - SS_E = 1230.5 - 120.3 = 1110.2$$

$$MS_R = \frac{SS_R}{k} = \frac{1110.2}{2} = 555.1$$

$$MS_E = \frac{SS_E}{n-p} = \frac{120.3}{15-3} = 10.025$$

$$f_0 = \frac{MS_R}{MS_E} = \frac{555.1}{10.025} = 55.37$$

$H_0: \beta_1 = \beta_2 = 0$

$H_1: \beta_j \neq 0$ for at least one j

$\alpha = 0.05$

$f_0 = 55.37$

$f_{.05,2,12} = 3.89$

$f_0 > f_{\alpha,2,12}$

Reject H_0 and conclude regression model is significant at $\alpha = 0.05$

c) $R^2 = \frac{SS_R}{S_{yy}} = \frac{1110.2}{1230.5} = 0.9022$ or 90.22%

d) k = 3 p = 4 $SS_E' = 117.20$

$$MS_E' = \frac{SS_E'}{n-p} = \frac{117.2}{11} = 10.65$$

NO, MS_E increased with the addition of x_3 because the reduction in SS_E was not enough to compensate for the loss in one degree of freedom in the observation. This is why MS_E can be used as a model selection criteria.

e) $SS_R = S_{yy} - SS_E = 1230.5 - 117.20 = 1113.30$

$$\begin{aligned} SS_R(\beta_3|\beta_2,\beta_1,\beta_0) &= SS_R(\beta_3\beta_2\beta_1|\beta_0) - SS_R(\beta_2,\beta_1|\beta_0) \\ &= 1113.30 - 1110.20 \\ &= 3.1 \end{aligned}$$

$H_0: \beta_3 = 0$

$H_1: \beta_3 \neq 0$

$\alpha = 0.05$

$$f_0 = \frac{SS_R(\beta_3|\beta_2,\beta_1,\beta_0)/r}{SS_E'/n-p} = \frac{3.1/1}{117.2/11} = 0.291$$

$f_{.05,1,11} = 4.84$

$f_0 \ngtr f_{\alpha,1,11}$

Do not reject H_0.

11-73. a) $R^2 = \frac{SS_R}{S_{yy}}$

$SS_R = R^2(S_{yy}) = 0.94(0.50) = 0.47$

$SS_E = S_{yy} - SS_R = 0.5 - 0.47 = 0.03$

$H_0: \beta_1 = \beta_2 = ... = \beta_6 = 0$

$H_1: \beta_j \neq 0$ for at least one j.

$\alpha = 0.05$

$f_0 = \frac{SS_R / k}{SS_E / n-p} = \frac{0.47/6}{0.03/7} = 18.28$

$f_{.05,6,7} = 3.87$

$f_0 > f_{\alpha,6,7}$

Reject H_0.

b) $k = 5$ $n = 14$ $p = 6$ $R^2 = 0.92$

$SS_R' = R^2(S_{yy}) = 0.92(0.50) = 0.46$

$SS_E' = S_{yy} - SS_R' = 0.5 - 0.46 = 0.04$

$$\begin{aligned} SS_R(\beta_x|\beta_1,...,\beta_5,\beta_0) &= SS_R(\text{full}) - SS_R(\text{reduced}) \\ &= 0.47 - 0.46 \\ &= 0.01 \end{aligned}$$

$f_0 = \frac{SS_R(\beta_x|\beta_1,...,\beta_5)/r}{SS_E'/n-p} = \frac{0.01/1}{0.04/8} = 2$

$f_{.05,1,8} = 5.32$

$f_0 \not> f_{\alpha,1,8}$

Do not reject H_0 and conclude that the evidence is insufficient to claim that the removed variable is significant at $\alpha = 0.05$

c) $MS_E(\text{reduced}) = \frac{SS_E}{n-p} = \frac{0.04}{8} = 0.0053$

$MS_E(\text{full}) = \frac{0.03}{7} = 0.004$

NO, the MS_E is larger for the reduced model. The increase in MS_E is sufficient to compensate for the addition of one degree of freedom in the denominator.

Mind-Expanding Exercises

11-74. Because $R^2 = \frac{SS_R}{S_{yy}}$ and $1 - R^2 = \frac{SS_E}{S_{yy}}$, $F_0 = \frac{SS_R / k}{SS_E / (n-k-1)}$ which is the usual F-test for significance of regression. Then, $F_0 = \frac{0.90/4}{(1-0.9)/(20-4-1)} = 33.75$ and the critical value $f_{.05,4,15} = 3.06$. Because 33.75 > 3.06, the regression is significant.

11-75. Using n = 20, k = 4, $f_{.05,4,15} = 3.06$. Reject H_0 if

$\frac{R^2/4}{(1-R^2)/15} \geq 3.06$

$\frac{R^2}{(1-R^2)} \geq 0.816$

Then, $R^2 \geq 0.449$ results in a significant regression.

11-76. Because $\hat{\beta} = (X'X)^{-1}X'Y$, $e = Y - X\hat{\beta} = Y - X(X'X)^{-1}X'Y = (I - H)Y$

11-77. From Exercise 11-76, e_i is ith element of (I-H)Y. That is,

$$e_i = -h_{i,1}Y_1 - h_{i,2}Y_2 - ... - h_{i,i-1}Y_{i-1} + (1 - h_{i,i})Y_i - h_{i,i+1}Y_{i+1} - ... - h_{i,n}Y_n$$

and

$$V(e_i) = (h_{i,1}^2 + h_{i,2}^2 + ... + h_{i,i-1}^2 + (1 - h_{i,i}^2) + h_{i,i+1}^2 + ... + h_{i,n}^2)\sigma^2$$

The expression in parentheses is recognized to be the ith diagonal element of (I-H)(I-H') = I-H by matrix multiplication. Consequently, $V(e_i) = (1 - h_{i,i})\sigma^2$. Assume that i < j. Now,

$$e_i = -h_{i,1}Y_1 - h_{i,2}Y_2 - \ldots - h_{i,i-1}Y_{i-1} + (1 - h_{i,i})Y_i - h_{i,i+1}Y_{i+1} - \ldots - h_{i,n}Y_n$$

$$e_j = -h_{j,1}Y_1 - h_{j,2}Y_2 - \ldots - h_{j,j-1}Y_{j-1} + (1 - h_{j,j})Y_j - h_{j,j+1}Y_{j+1} - \ldots - h_{j,n}Y_n$$

Because the Y_i's are independent,

$$Cov(e_i, e_j) = (h_{i,1}h_{j,1} + h_{i,2}h_{j,2} + \ldots + h_{i,i-1}h_{j,i-1} + (1 - h_{i,i})h_{j,i}$$
$$+ h_{i,i+1}h_{j,i+1} + \ldots + h_{i,j}(1 - h_{j,j}) + h_{i,j+1}h_{j,j+1} + \ldots + h_{i,n}h_{j,n})\sigma^2$$

The expression in parentheses is recognized as the ij th element of (I-H)(I-H') = I-H. Therefore,

$Cov(e_i, e_j) = -h_{ij}\sigma^2$.

11-78. $\hat{\beta} = (X'X)^{-1}X'Y = (X'X)^{-1}X(X\beta + \varepsilon) = \beta + (X'X)^{-1}X'\varepsilon = \beta + R\varepsilon$

11-79. a) Min L = $(y - X\beta)'(y - X\beta)$

subject to :$\mathbf{T\beta = c}$

This is equivalent to Min Z = $(\mathbf{y} - \mathbf{X\beta})'(\mathbf{y} - \mathbf{X\beta}) + 2\lambda'(\mathbf{T\beta} - \mathbf{c})$

where $\lambda' = [\lambda_1, \lambda_2, ..., \lambda_p]$ is a vector of La Grange multipliers.

$$\frac{\partial \mathbf{Z}}{\partial \beta} = -2\mathbf{X'y} + 2(\mathbf{X'X})\beta + 2\mathbf{T'}\lambda$$

$$\frac{\partial \mathbf{Z}}{\partial \lambda} = \mathbf{T\beta} - \mathbf{c}. \quad \text{Set} \quad \frac{\partial \mathbf{Z}}{\partial \beta} = 0, \quad \frac{\partial \mathbf{Z}}{\partial \lambda} = 0.$$

Then we get

$$(\mathbf{X'X})\hat{\beta}_\mathbf{c} + \mathbf{T'}\lambda = \mathbf{X'y}$$

$$\mathbf{T}\hat{\beta}_\mathbf{c} = \mathbf{c}$$

where $\hat{\beta}_\mathbf{c}$ is the constrained estimator.

From the first of these equations,

$$\hat{\beta}_\mathbf{c} = (\mathbf{X'X})^{-1}(\mathbf{X'y} - \mathbf{T'}\lambda) = \hat{\beta} - (\mathbf{X'X})^{-1}\mathbf{T'}\lambda$$

From the second, $\mathbf{T}\hat{\beta} - \mathbf{T}(\mathbf{X'X})^{-1}\mathbf{T'}\lambda = \mathbf{c}$ or

$$\lambda = [\mathrm{T}(\mathrm{X'X})^{-1}\mathrm{T'}]^{-1}(\mathbf{T}\hat{\beta} - \mathbf{c})$$

Then

$$\hat{\beta}_\mathbf{c} = \hat{\beta} - (\mathbf{X'X})^{-1}\mathbf{T'}[\mathbf{T}(\mathbf{X'X})^{-1}\mathbf{T'}]^{-1}(\mathbf{T}\hat{\beta} - \mathbf{c})$$
$$= \hat{\beta} + (\mathbf{X'X})^{-1}\mathbf{T'}[\mathbf{T}(\mathbf{X'X})^{-1}\mathbf{T'}]^{-1}(\mathbf{c} - \mathbf{T}\hat{\beta})$$

b) This solution would be appropriate in situations where you have tested the hypothesis that $\mathbf{T\beta = c}$ and concluded that this hypothesis cannot be rejected.

11-80. For the piecewise linear function to be continuous at $x = x^*$, the point-slope formula for a line can be used to show that

$$y = \begin{cases} \beta_0 + \beta_1(x - x^*) & x \le x^* \\ \beta_0 + \beta_2(x - x^*) & x > x^* \end{cases}$$

where $\beta_0, \beta_1, \beta_2$ are arbitrary constants.

Let $z = \begin{cases} 0, & x \le x^* \\ 1, & x > x^* \end{cases}$.

Then, y can be written as $y = \beta_0 + \beta_1(x - x^*) + (\beta_2 - \beta_1)(x - x^*)z$.

Let

$x_1 = x - x^*$

$x_2 = (x - x^*)z$

$\beta_0^* = \beta_0$

$\beta_1^* = \beta_1$

$\beta_2^* = \beta_2 - \beta_1$

Then, $y = \beta_0^* + \beta_1^* x_1 + \beta_2^* x_2$.

11-81. If there is a discontinuity at $x = x^*$, then a model that can be used is

$$y = \begin{cases} \beta_0 + \beta_1 x & x \le x^* \\ \alpha_0 + \alpha_1 x & x > x^* \end{cases}$$

Let $z = \begin{cases} 0, & x \le x^* \\ 1, & x > x^* \end{cases}$

Then, y can be written as $y = \beta_0 + \beta_1 x + [(\alpha_0 - \beta_0) + (\alpha_1 - \beta_1)x]z = \beta_0^* + \beta_1^* x_1 + \beta_2^* z + \beta_3^* x_2$

where

$\beta_0^* = \beta_0$

$\beta_1^* = \beta_1$

$\beta_2^* = \alpha_0 - \beta_0$

$\beta_3^* = \alpha_1 - \beta_1$

$x_1 = x$

$x_2 = xz$

11-82. One could estimate x* as a parameter in the model. A simple approach is to refit the model in Exercise 11-80 with different choices for x* and to select the value for x* that minimizes the residual sum of squares.

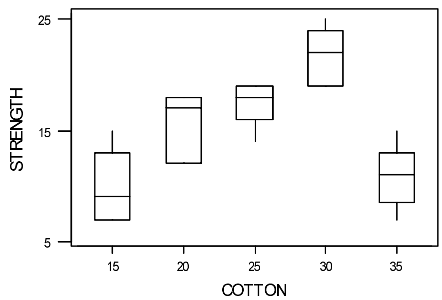
CHAPTER 12

Section 12-2

12-1. a)

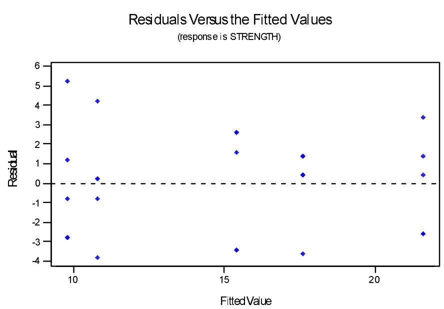
Source	DF	SS	MS	F	P
COTTON	4	475.76	118.94	14.76	0.000
Error	20	161.20	8.06		
Total	24	636.96			

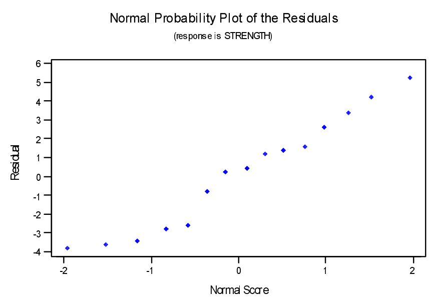
Reject H_0 and conclude that cotton percentage affects breaking strength.

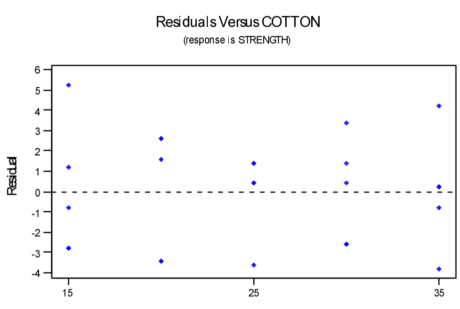


b) Tensile strength seems to increase to 30% cotton and declines at 35% cotton.

c)



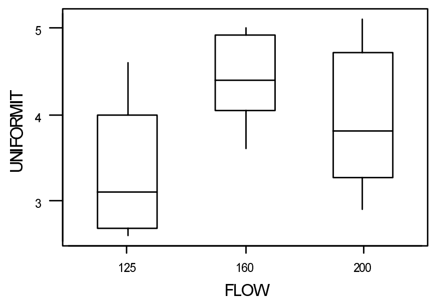




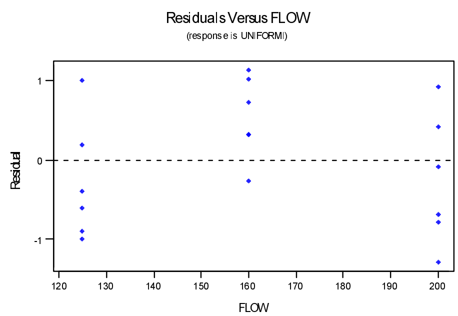
12-2. a)

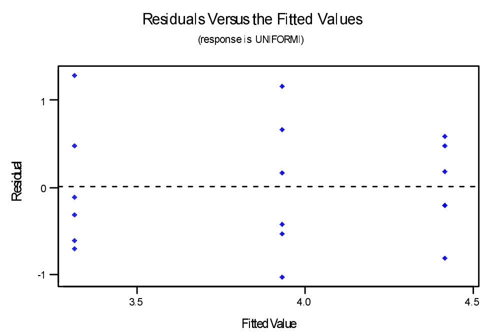
Source	DF	SS	MS	F	P
FLOW	2	3.6478	1.8239	3.59	0.053
Error	15	7.6300	0.5087		
Total	17	11.2778			

Do not reject H_0. Flow rate does not affect etch uniformity.



b) Residuals are acceptable.

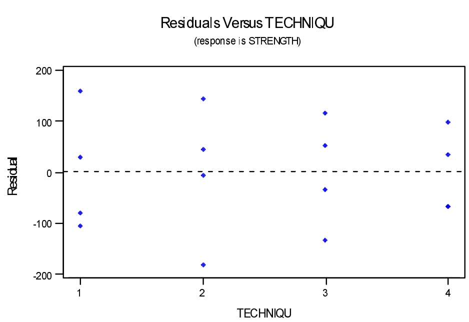


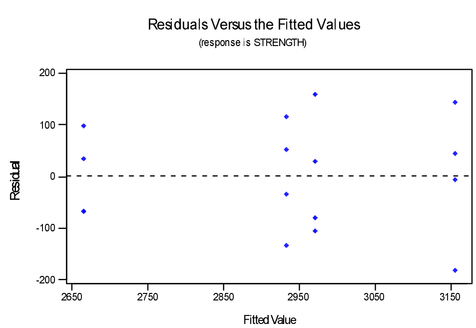


12-3. a)

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 strength of the concrete.
b) P-value = 0
c) Residuals are acceptable





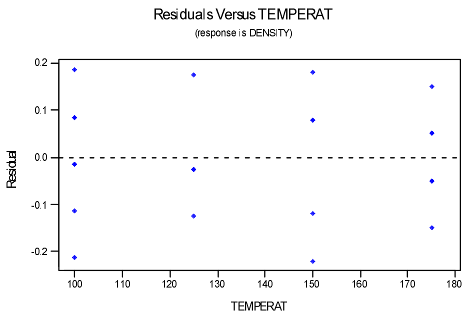
12-4. a)

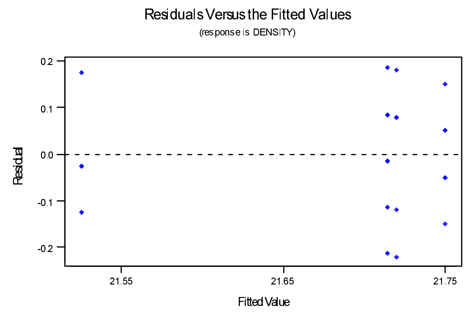
Source	DF	SS	MS	F	P
TEMPERAT	3	0.1391	0.0464	2.62	0.083
Error	18	0.3191	0.0177		
Total	21	0.4582			

Do not reject H_0

b) P-value = 0.083

c) Residuals are acceptable



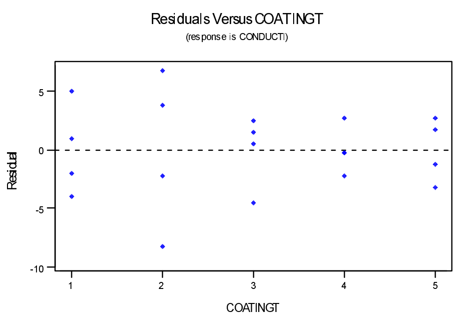


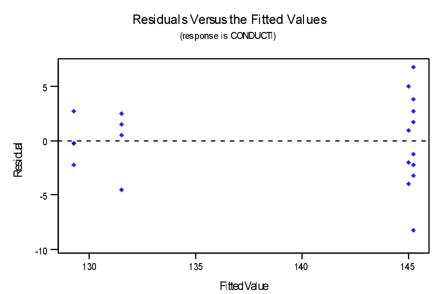
12-5. a)

Source	DF	SS	MS	F	P
COATINGT	4	1060.5	265.1	16.35	0.000
Error	15	243.3	16.2		
Total	19	1303.7			

Reject H_0

b) There is some indication of increasing variability.





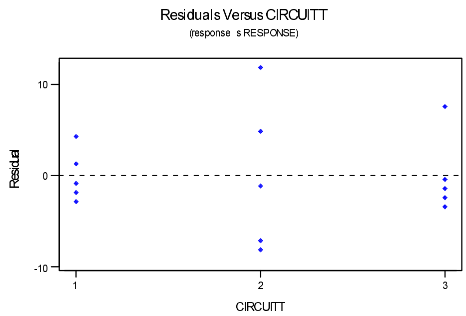
c) (141.96, 148.04)

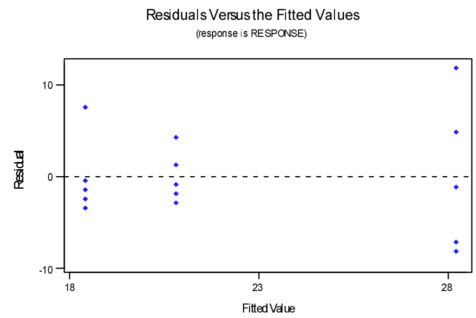
12-6. a)

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.





c) (14.47, 22.33)

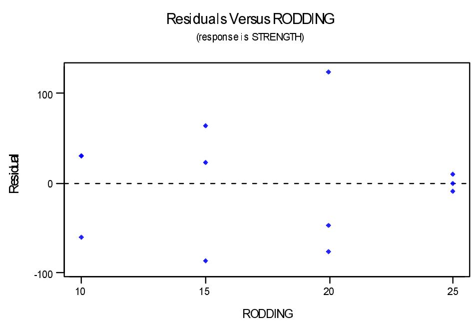
12-7. a)

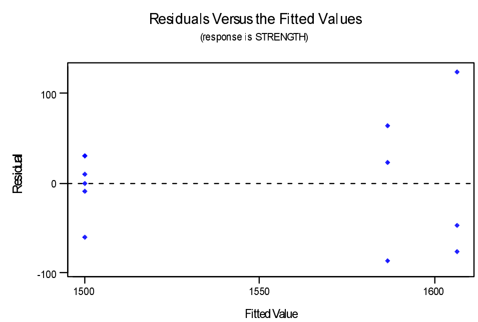
Source	DF	SS	MS	F	P
RODDING	3	28633	9544	1.87	0.214
Error	8	40933	5117		
Total	11	69567			

Do not reject H_0

b) P-value = 0.214

c)





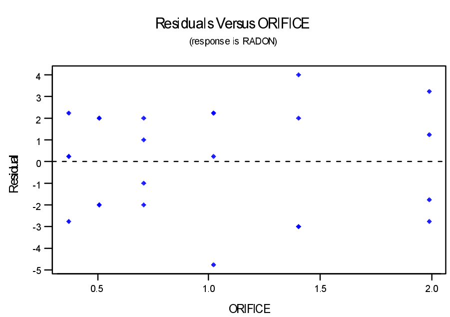
12-8. a)

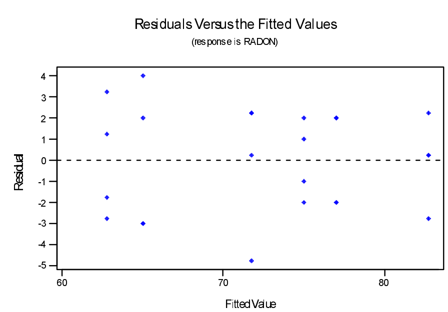
Source	DF	SS	MS	F	P
ORIFICE	5	1133.37	226.67	30.85	0.000
Error	18	132.25	7.35		
Total	23	1265.63			

Reject H_0

b) P-value = 0

c)





d) (62.99, 67.01)

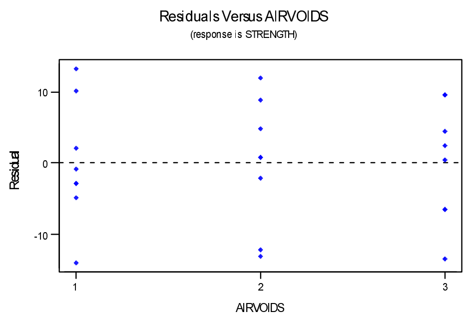
12-9. a)

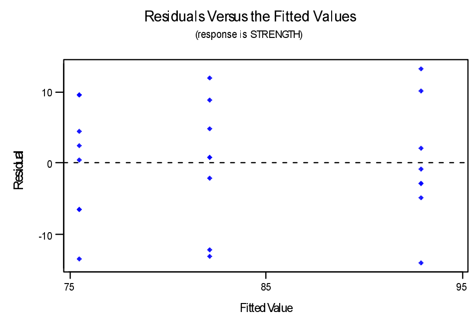
Source	DF	SS	MS	F	P
AIRVOIDS	2	1230.2	615.1	8.30	0.002
Error	21	1555.8	74.1		
Total	23	2786.0			

Reject H_0

b) P-value = 0.002

c)





d) (71.02, 79.98)

e) (8.423, 26.327)

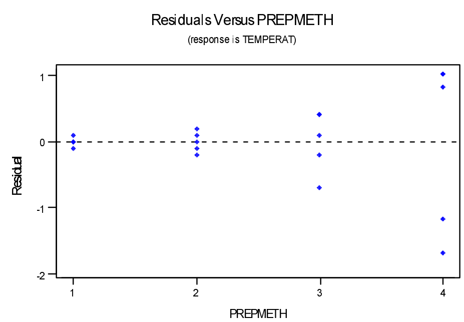
12-10. a)

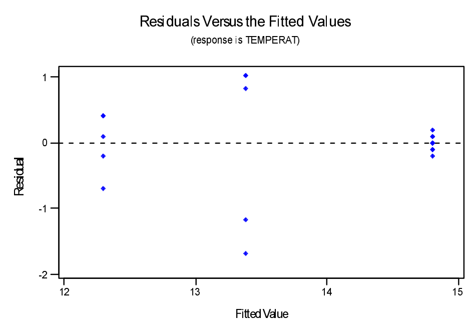
Source	DF	SS	MS	F	P
PREPMETH	3	22.124	7.375	14.85	0.000
Error	16	7.948	0.497		
Total	19	30.072			

Reject H_0

b) P-value = 0

c)





d) (14.71, 14.89)

Section 12-2

12-11. MS_E = 8.06, standard error of the mean = $\sqrt{\frac{MS_E}{n}}$ =1.27

9.8 10.8 15.4 17.6 21.6

12-12. 3.32 3.93 4.42

12-13. 2666.25 2933.75 2971.00 3156.25

12-14. 129.25 131.5 145.0 145.25 145.25

12-15. 17.4 20.8 28.2

12-16. 62.75 65.0 71.75 75.0 77.0 82.75

12-17. 75.5 82.125 92.875

12-18. 12.3 13.28 14.8 14.8

12-19.

```
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 25, 25 and 30, 25 and 35; and 30 and 35.

12-20.

```
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
     200     -1.4942     -0.3942
              0.2608      1.3608
```

No significant differences among the levels.

12-21.

1) $\mu_{125} - 0.5\mu_{160} + 0.5\mu_{200}$

2) $0\mu_{125} - \mu_{160} + \mu_{200}$

$\bar{x}_{125} = 3.32 \quad \bar{x}_{160} = 3.93 \quad \bar{x}_{200} = 4.42$

$$SS_{contrast(1)} = \frac{(-5.13)^2}{6(1+.25+.25)} = 2.92$$

$SS_{contrast(2)} = 0.720$

$MS_E = 0.509 \quad f_{0.05,1,15} = 4.54$

contrast(1): $2.92 / 0.509 = 5.74$, Re ject

contrast(2): $0.720 / 0.509 = 1.41$ Do not Re ject

12-22.

```
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 4, 2 and 3, 2 and 4, and 3 and 4.

12-23.

```
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 between levels 1 and 3, 1 and 4, 2 and 3, 2 and 4, 3 and 5, and 4 and 5.

12-24.

```
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      -1.226
              13.426      20.826
```

No significant differences among the levels.

12-25.

contrast	μ_1	μ_2	μ_3	μ_4	μ_5	SS	f_0	Decision
1	-0.25	-0.25	-0.25	-0.25	1	180	11.10	Reject
2	-0.5	-0.5	0.5	0.5	0	870.25	53.66	Reject
3	-1	1	0	0	0	0.125	0.008	Do not Reject
4	0	0	-1	1	0	10.125	0.624	Do not Reject
average	145	145.25	131.5	129.25	145.25	MS_E=16.22		

12-26.

contrast	μ_1	μ_2	μ_3	SS	f_0	Decision
1	0	1	-1	240.1	7.37	Do not reject
2	1	-0.5	-0.5	20.83	0.64	Do not reject
average	20.8	28.2	18.4	MS_E=32.57		

12-27.

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

Significant differences between all levels except 0.71 and 1.02; and 1.40 and 1.99.

12-28.
```
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      -2.326
              26.326      15.576
```
Significant differences between levels 1 and 2; and 2 and 3.

12-29.
```
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
```
Significance differences among all levels except levels 1 and 2.

12-30.

contrast	μ_1	μ_2	μ_3	μ_4	SS	f_0	Decision
1	0.5	0.5	-0.5	-0.5	19.21	38.67	Reject
2	1	-1	0	0	0	0	Do not reject
3	0	0	1	-1	2.916	5.87	Reject
average	14.8	14.8	12.3	13.38	MS_E=0.4968		

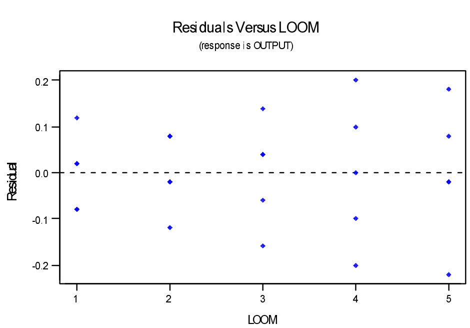
Section 12-4

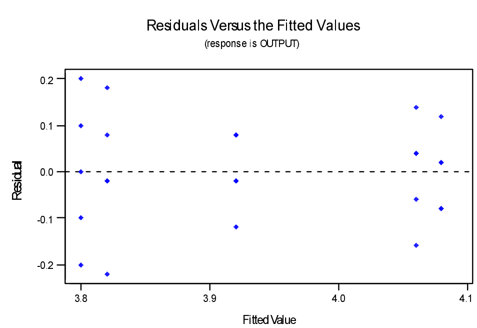
12-31. a)
```
Source    DF       SS        MS        F       P
LOOM       4   0.3416    0.0854     5.77   0.003
Error     20   0.2960    0.0148
Total     24   0.6376
```
Reject H_0, there are significant differences among the looms.

b) 0.0141

c) 0.0148

d) Residuals are acceptable





12-32. a)

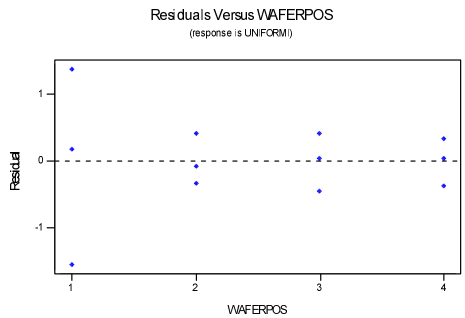
```
Analysis of Variance for UNIFORMI
Source     DF        SS        MS        F       P
WAFERPOS    3    16.220     5.407     8.29   0.008
Error       8     5.217     0.652
Total      11    21.437
```

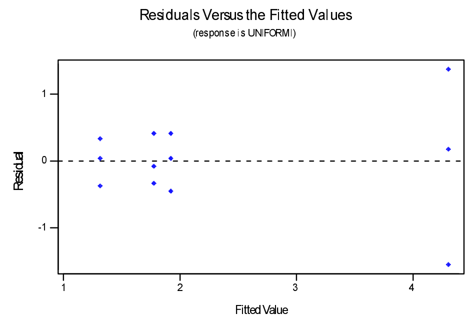
Reject H_0, there is a significant difference among wafer position.

b) 1.5848

c) 0.652

d) Greater variability at wafer position 1.





12-33. a)

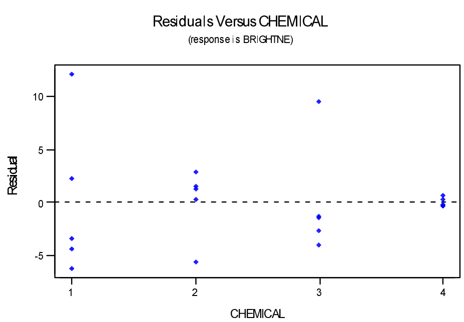
```
Source     DF        SS        MS        F       P
CHEMICAL    3      54.0      18.0     0.75   0.538
Error      16     384.0      24.0
Total      19     438.0
```

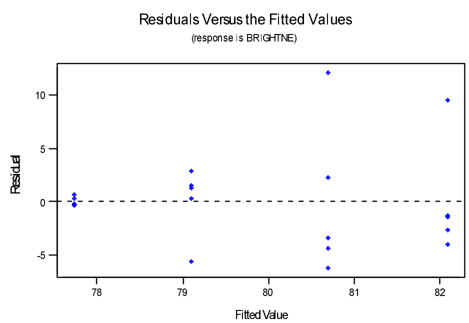
Do not reject H_0, there is no difference among the chemical types.

b) −1.201, set equal to 0

c) 23.9994

d) Variability is smaller in chemical 4.





12-34. a) $\hat{\sigma}^2_{total} = \hat{\sigma}^2_{position} + \hat{\sigma}^2 = 2.237$

b) $\dfrac{\hat{\sigma}^2_{position}}{\hat{\sigma}^2_{total}} = 0.708$

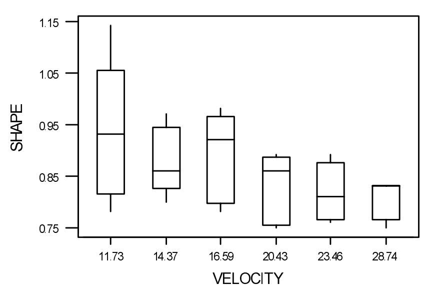
c) It could be reduced to 0.6522 - this is the reduction of approximately 70%.

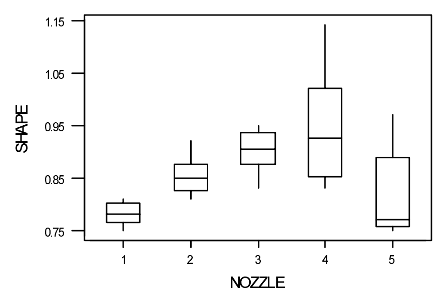
Section 12-5

12-35. a)

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			

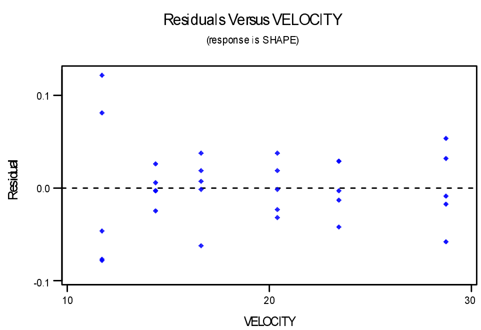
Reject H_0, nozzle type affects shape measurement.

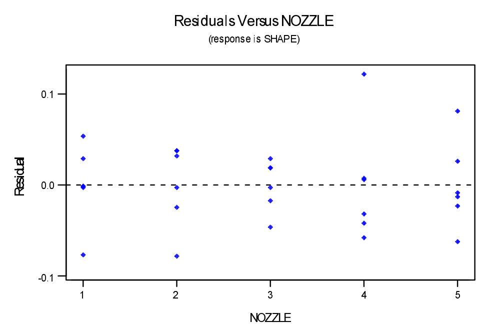


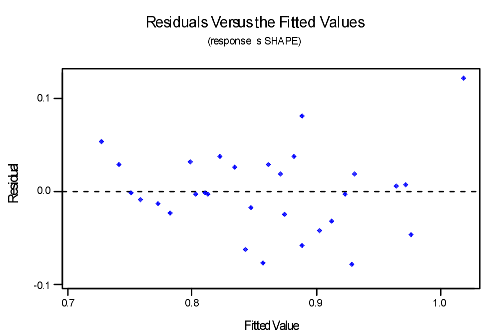


b) <u>0.78 0.81 0.85</u>
<u>0.85 0.90 0.94</u>

c)







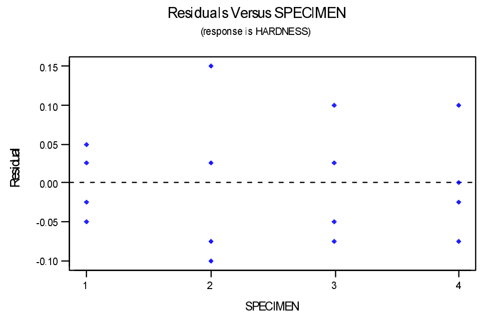
12-36. a)

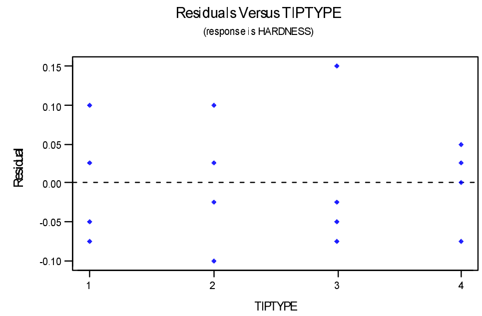
Source	DF	SS	MS	F	P
TIPTYPE	3	0.38500	0.12833	14.44	0.001
SPECIMEN	3	0.82500	0.27500	30.94	0.000
Error	9	0.08000	0.00889		
Total	15	1.29000			

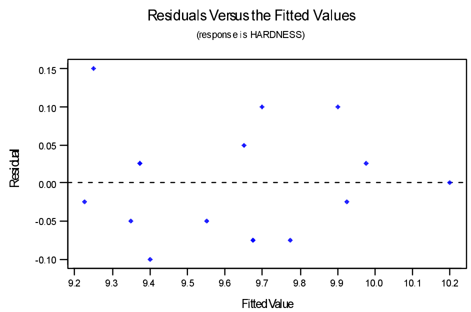
Reject H_0, there are significant differences in hardness measurements between the tips.

b) <u>9.575 9.60 9.45</u> 9.875

c) Residuals are acceptable.





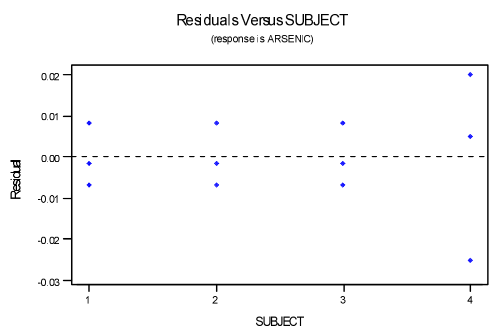


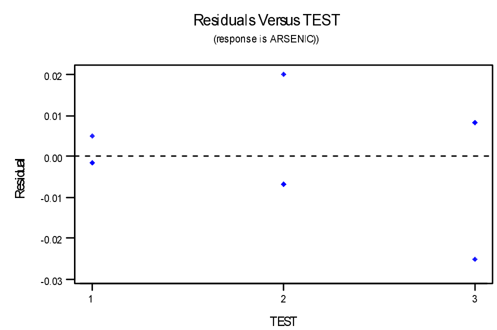
12-37. a)

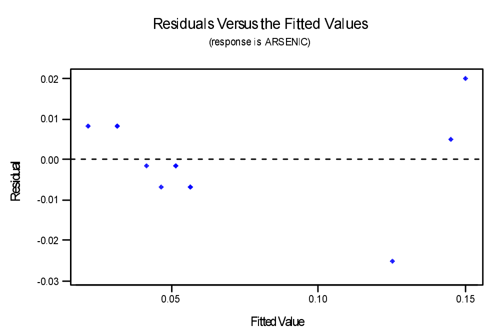
Source	DF	SS	MS	F	P
TEST	2	0.0014000	0.0007000	3.00	0.125
SUBJECT	3	0.0212250	0.0070750	30.32	0.001
Error	6	0.0014000	0.0002333		
Total	11	0.0240250			

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.







12-38. a)

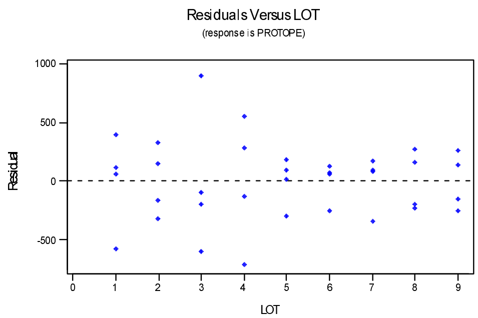
Source	DF	SS	MS	F	P
STORAGE	3	1972652	657551	4.33	0.014
LOT	8	1980499	247562	1.63	0.169
Error	24	3647150	151965		
Total	35	7600300			

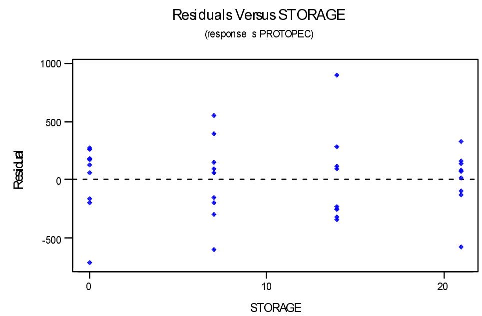
Reject H_0, the storage times are significantly different.

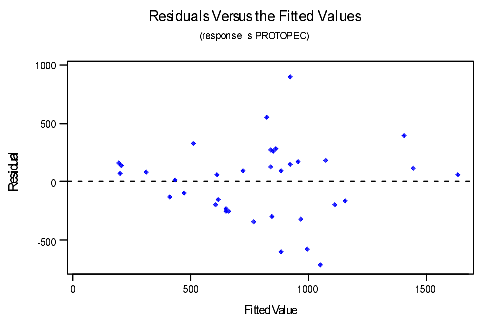
b) P-value = 0.014

c) 415.5 825.82 868.33 1057.44

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





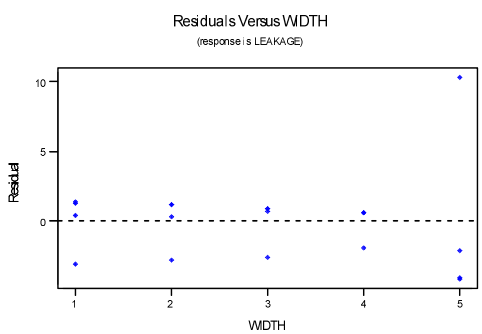


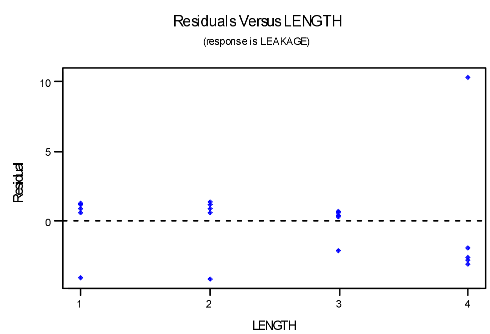
12-39. a)

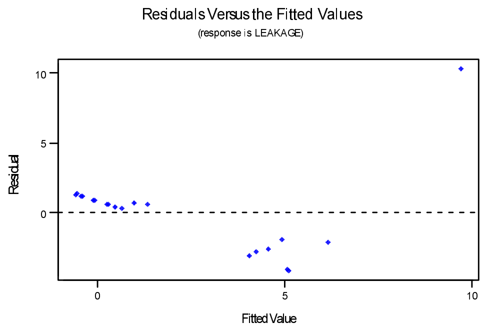
Source	DF	SS	MS	F	P
LENGTH	3	72.66	24.22	1.61	0.240
WIDTH	4	90.52	22.63	1.50	0.263
Error	12	180.83	15.07		
Total	19	344.01			

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

b) One unusual observation in width 5, length 4.







12-40.

```
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. mean leakage voltage does depend on channel length.

Section 12-6

12-41. $\overline{\mu} = 55$, $\tau_1 = -5$, $\tau_2 = 5$, $\tau_3 = -5$, $\tau_4 = 5$.

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

Various choices:

n	Φ^2	Φ	a(n-1)	Power=1-β
4	4	2	12	0.80
5	5	2.24	16	0.95

Therefore, n = 5 is needed.

12-42. $\overline{\mu} = 188$, $\tau_1 = -13$, $\tau_2 = 2$, $\tau_3 = -28$, $\tau_4 = 12$, $\tau_5 = 27$. $\Phi^2 = 366n$

n	Φ^2	Φ	a(n-1)	Power=1-β
3	10.98	3.31	12	0.96
2	7.32	2.71	16	0.05

Therefore, n = 3 is needed.

12-43. a = 4, a(n-1) = 16, and $\tau = 2.69$. Thus $\beta = 0.4$, and power = 0.6

12-44. a = 3, a(n-1) = 15, and $\tau = 2.24$. Thus $\beta = 0.5$, and power = 0.5

12-45.

n	τ^2	τ	a(n-1)	power=1-β
3	4	2	6	<0.6
4	5	2.24	9	<0.6
5	6	2.45	12	0.55
6	7	2.66	15	0.6

Therefore, n = 6 is needed.

Supplemental Exercises

12-46. a)

Analysis of Variance for RESISTAN

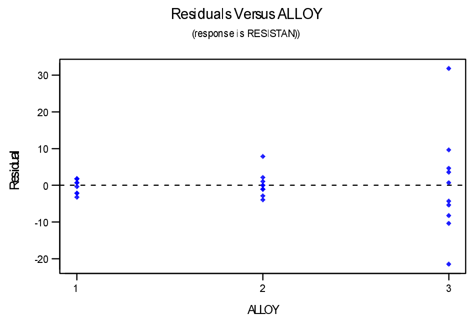
Source	DF	SS	MS	F	P
ALLOY	2	10941.8	5470.9	76.09	0.000
Error	27	1941.4	71.9		
Total	29	12883.2			

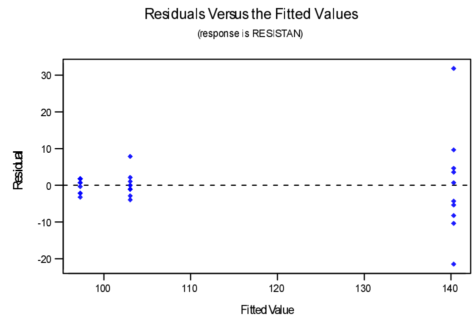
Reject H_0,

b) 97.3 103.1 140.4

c) (135.15, 145.65)

d) Variability of the residuals increases with the response





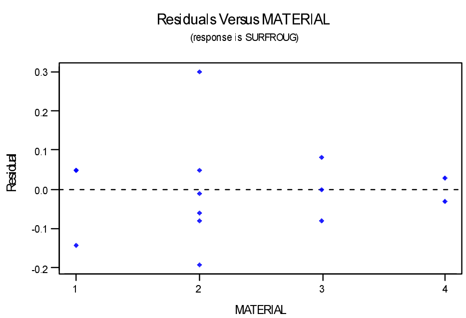
12-47. a)

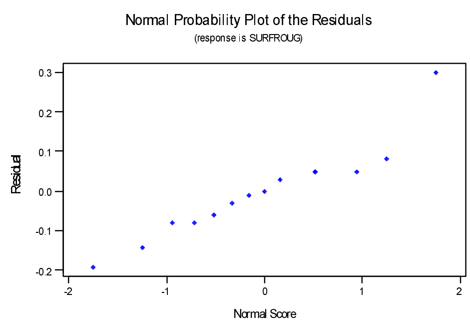
Analysis of Variance for SURFROUG

Source	DF	SS	MS	F	P
MATERIAL	3	0.2604	0.0868	5.29	0.017
Error	11	0.1804	0.0164		
Total	14	0.4408			

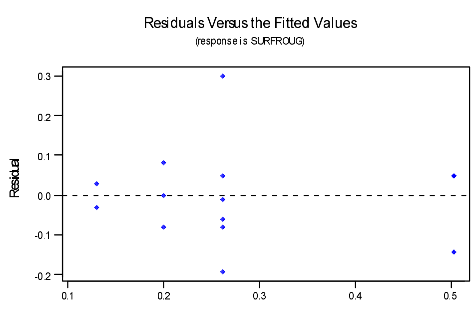
Reject H_0

b) One observation is an outlier.





c)



d) (0.128, 0.6166)

12-48. a=4, $n_H = 3.2$, $MS_E = 0.0164$,

Comparisons	Differences	R_i
1 vs 2	0.3725	0.2399
1 vs 3	0.3025	0.2341
1 vs 4	0.2408	0.2227
2 vs 4	0.1317	0.2399

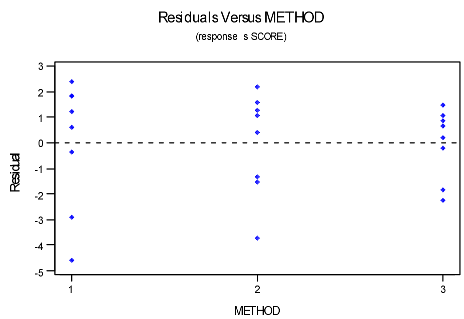
12-49. a)

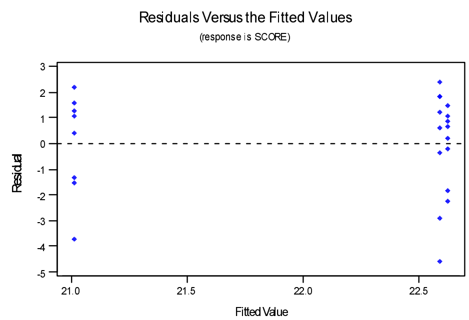
```
Analysis of Variance for SCORE
Source     DF        SS        MS        F        P
METHOD      2     13.55      6.78     1.68    0.211
Error      21     84.77      4.04
Total      23     98.32
```

Do not reject H_0

b) P-value = 0.211

c)





d) 4.037, Method: 0.342

12-50. a)

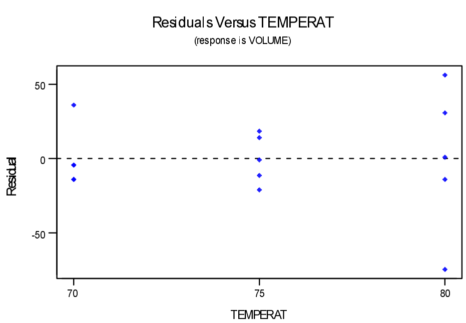
```
Analysis of Variance for VOLUME
Source     DF        SS        MS        F        P
TEMPERAT    2     16480      8240     7.84    0.007
Error      12     12610      1051
Total      14     29090
```

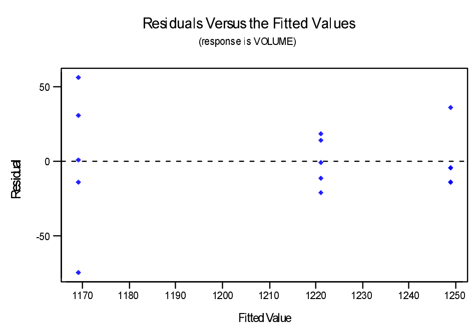
Reject H_0.

b) P-value = 0.007

c) 1169 1221 1249

d)





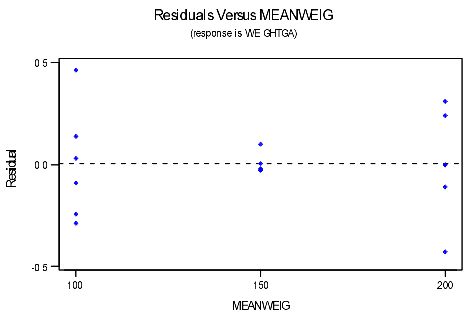
12-51. a)

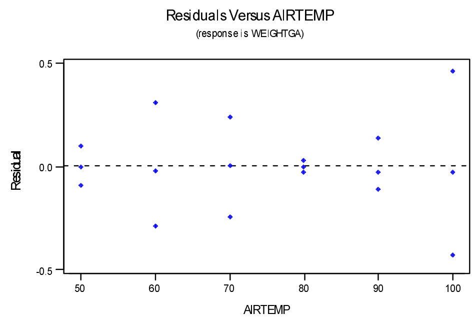
Source	DF	SS	MS	F	P
MEANWEIG	2	0.2227	0.1113	1.48	0.273
AIRTEMP	5	10.1852	2.0370	27.13	0.000
Error	10	0.7509	0.0751		
Total	17	11.1588			

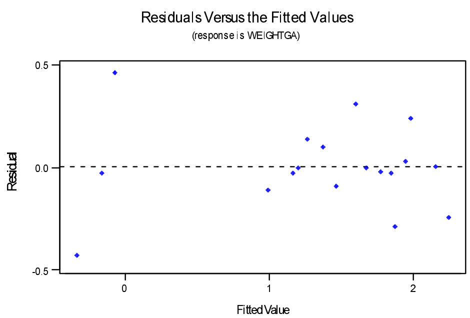
Reject H_0

b) 0.19 1.14 1.34 1.74 1.82 2.12

c)





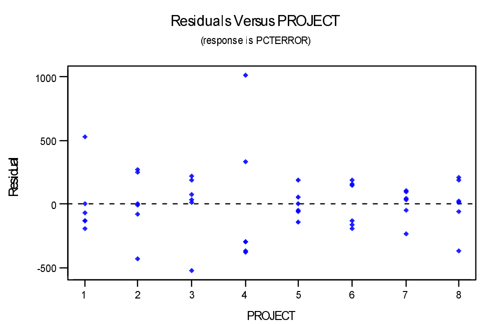


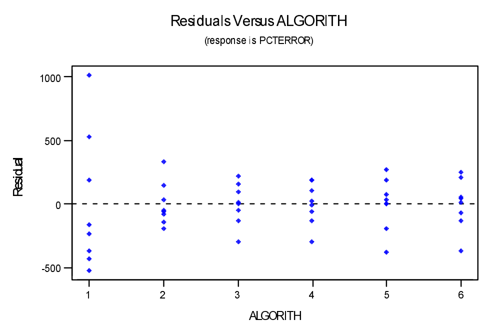
12-52. a)

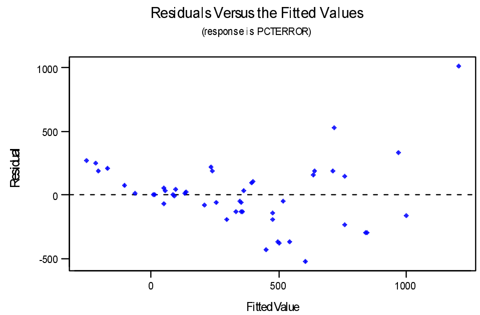
Source	DF	SS	MS	F	P
ALGORITH	5	2825746	565149	6.23	0.000
PROJECT	7	2710323	387189	4.27	0.002
Error	35	3175290	90723		
Total	47	8711358			

Reject H_0

b)







c) The best choice is algorithm 5.

12-53. a) $\mu = (1+5+8+4)/4 = 4.5$ and

$$\Phi^2 = \frac{4[(1-4.5)^2 + (5-4.5)^2 + (8-4.5)^2 + (4-4.5)^2]}{4(4)} = 1.5625$$

$\Phi = 1.25$

Numerator degrees of freedom = $a - 1 = 3 = \nu_1$

Denominator degrees of freedom = $a(n-1) = 12 = \nu_2$

From Chart VII, $\beta = 0.7$ and the power = $1 - \beta = 0.3$

b)

n	Φ^2	Φ	a(n-1)	β	Power = 1-β
6	2.344	1.531	20	0.5	0.5
10	3.906	1.976	36	0.12	0.88
11	4.297	2.073	40	0.10	0.9

The sample size should be approximately n = 11.

12-54. a) $\lambda = \sqrt{1 + \frac{4(2\sigma^2)}{\sigma^2}} = 3$

From Chart VIII with numerator degrees of freedom = a - 1 = 4, denominator degrees of freedom = a(n - 1) = 15, $\beta = 0.15$, and the power = $1 - \beta = 0.85$.

b)

n	λ	a(n - 1)	β	Power = 1 - β
5	3.317	20	0.10	0.90

The sample size should be approximately n = 5

Mind-Expanding Exercises

12-55. $MS_E = \frac{\sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij} - \bar{y}_{i.})^2}{a(n-1)}$ and $y_{ij} = \mu + \alpha_i + \varepsilon_{ij}$. Then, $y_{ij} - \bar{y}_{i.} = \varepsilon_{ij} - \bar{\varepsilon}_{i.}$ and

$\frac{\sum_{j=1}^{n}(\varepsilon_{ij} - \bar{\varepsilon}_{i.})^2}{n-1}$ is recognized to be the sample variance of the independent random variables $\varepsilon_{i1}, \varepsilon_{i2}, \ldots, \varepsilon_{in}$.

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

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.

12-56. The two sample t-test rejects equality of means if the statistic

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

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

12-57. $MS_E = \dfrac{\sum_{i=1}^{2}\sum_{j=1}^{n}(y_{ij} - \bar{y}_{i.})^2}{2(n-1)}$ 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.

12-58. $V(\sum_{i=1}^{a} c_i Y_{i.}) = \sum_{i=1}^{a} c_i^2 V(Y_{i.})$ from the independence of $Y_{1.}, Y_{2.}, \ldots, Y_{a.}$.

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

12-59. 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 $Q_1^2 + Q_2^2 + Q_3^2 = \sum_{i=1}^{a}\frac{y_{i.}}{n} - \frac{y_{..}^2}{N}$ 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.

12-60. Because $\Phi^2 = \dfrac{n\sum_{i=1}^{a}(\mu_i - \bar{\mu})^2}{a\sigma^2}$, we only need to shows that $\dfrac{D^2}{2} \le \sum_{i=1}^{a}(\mu_i - \bar{\mu})^2$.

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

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

12-61. $MS_E = \frac{\sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij}-\bar{y}_{i.})^2}{a(n-1)} = \frac{\sum_{i=1}^{a} s_i^2}{a}$ 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

$y_{i1}, y_{i2}, \ldots, y_{in}$, $\frac{(n-1)S_i^2}{\sigma^2}$ 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.

12-62. From Exercise 12-61, $\frac{(N-a)MS_E}{\sigma^2}$ has a chi-square distribution with N - a degrees of freedom. Now,

$V(\bar{Y}_{i.}) = \sigma^2 + \frac{\sigma^2}{n}$ and 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-square distribution with a - 1 degrees of freedom. Using the

independence of MS_T and MS_E, we conclude that $\frac{\frac{MS_T}{n\sigma_\tau^2+\sigma^2}}{\frac{MS_E}{\sigma^2}}$ has an $F_{(a-1),(N-a)}$ distribution.

Therefore,

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

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

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

12-63. As in Exercise 12-62, $\frac{MS_T}{MS_E}\frac{\sigma^2}{n\sigma_\tau^2+\sigma^2}$ has an $F_{(a-1),(N-a)}$ distribution.

and

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

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

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

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

12-64. From Exercise 12-63,

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

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

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

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

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

12-65. $MS_T = \frac{\sum_{i=1}^{a} n_i(\bar{y}_{i.}-\bar{y}_{..})^2}{a-1}$ 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, $\bar{Y}_{1.}-\bar{Y}_{..} = (\frac{1}{n_1}-\frac{1}{N})Y_{11}+...+(\frac{1}{n_1}-\frac{1}{N})Y_{1n_1} - \frac{1}{N}Y_{21}-...-\frac{1}{N}Y_{2n_2}-...-\frac{1}{N}Y_{a1}-...-\frac{1}{N}Y_{an_a}$

and

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

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

Then,

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

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

12-66. a) If A is the width 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 12-56, $t^2_{\frac{\alpha}{2},a(n-1)} = F_{\alpha,1,a(n-1)}$. 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

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

CHAPTER 13

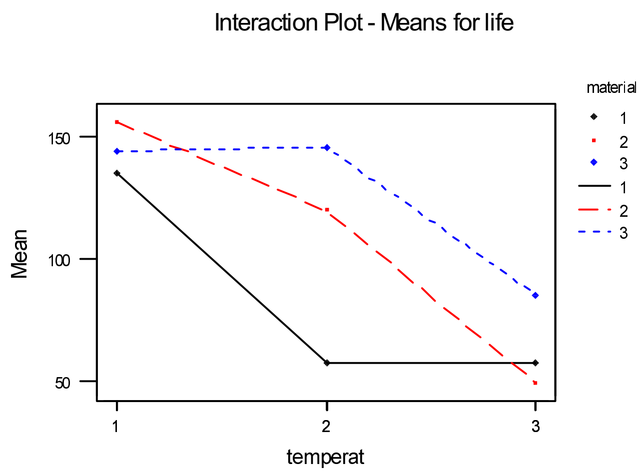
Section 13-4

13-1. a)

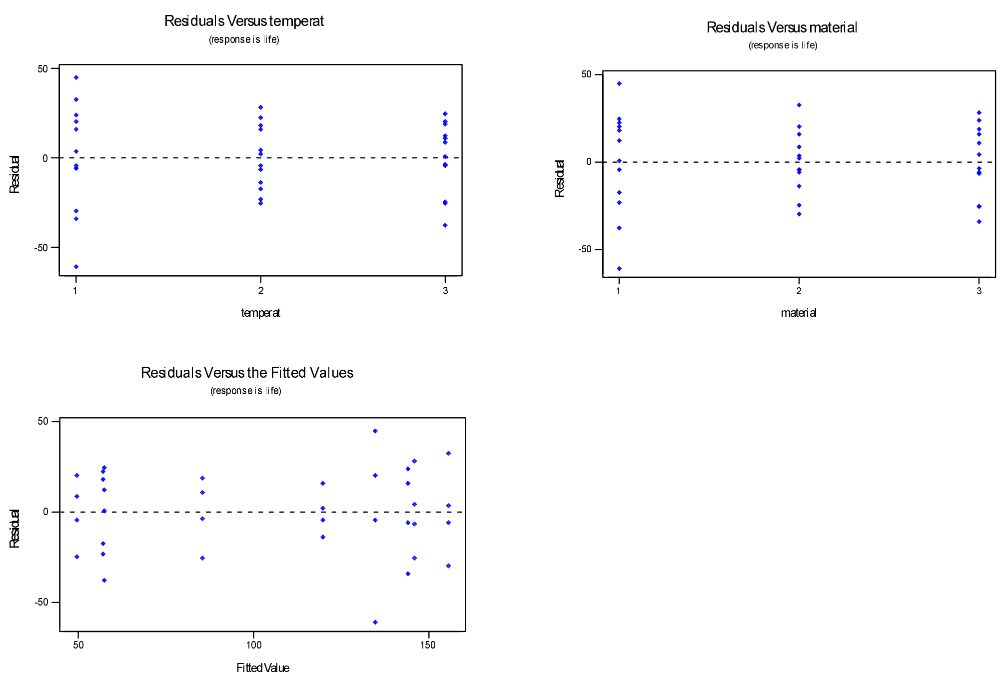
Source	DF	SS	MS	F	P
material	2	10683.7	5341.9	7.91	0.002
temperat	2	39118.7	19559.4	28.97	0.000
material*temperat	4	9613.8	2403.4	3.56	0.019
Error	27	18230.7	675.2		
Total	35	77647.0			

Main factors and interaction are all significant.

b)



c)

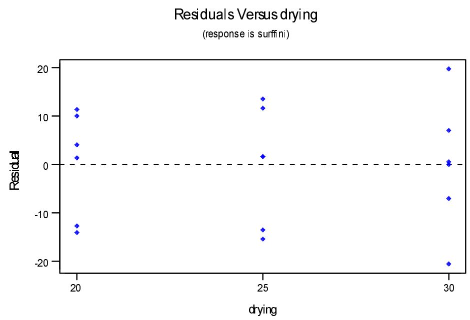


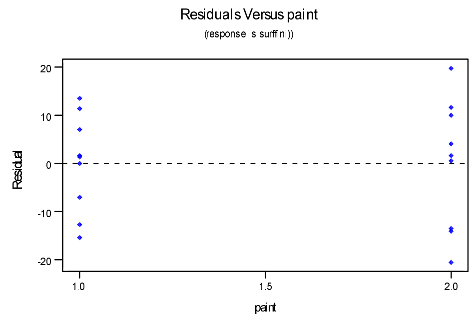
13-2. a)

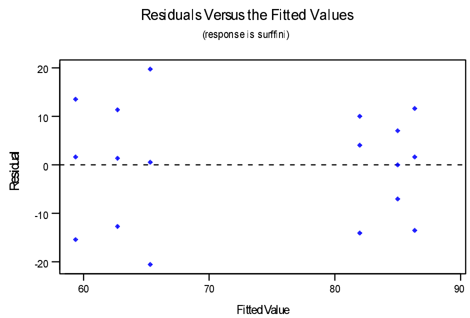
Source	DF	SS	MS	F	P
paint	1	355.6	355.6	1.90	0.193
drying	2	27.4	13.7	0.07	0.930
paint*drying	2	1878.8	939.4	5.03	0.026
Error	12	2242.7	186.9		
Total	17	4504.4			

Only the interaction is significant.

b)





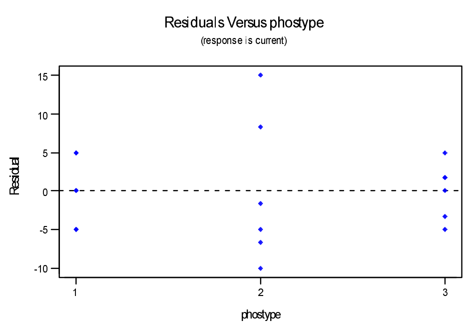


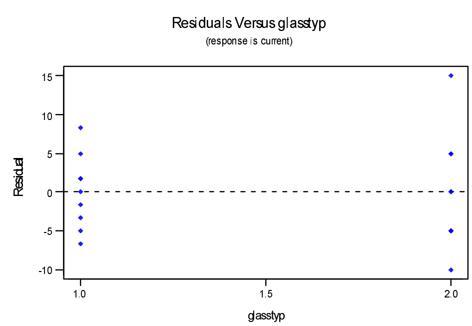
13-3. b)

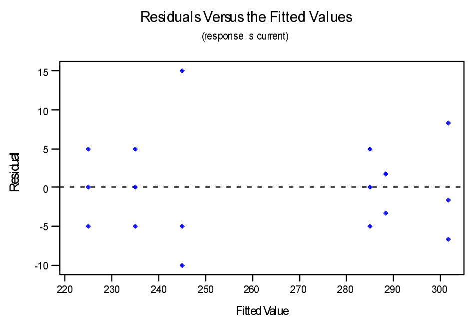
Source	DF	SS	MS	F	P
glasstyp	1	14450.0	14450.0	273.79	0.000
phostype	2	933.3	466.7	8.84	0.004
glasstyp*phostype	2	133.3	66.7	1.26	0.318
Error	12	633.3	52.8		
Total	17	16150.0			

Main effects are significant, the interaction is not significant.

c)





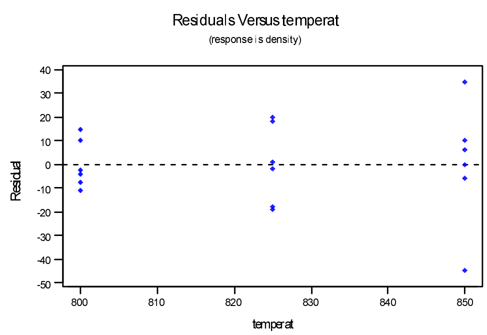


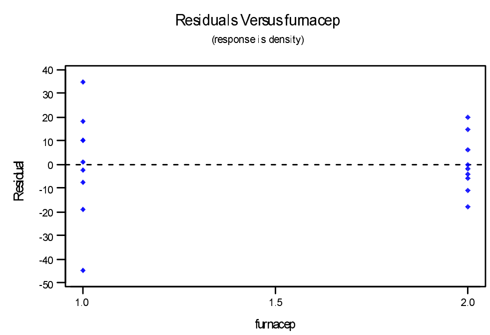
13-4. b)

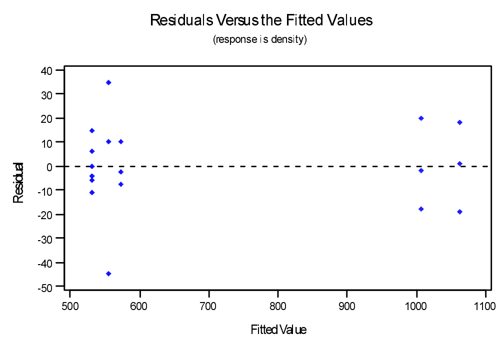
Source	DF	SS	MS	F	P
furnacep	1	7160	7160	16.00	0.002
temperat	2	945342	472671	1056.12	0.000
furnacep*temperat	2	818	409	0.91	0.427
Error	12	5371	448		
Total	17	958691			

Both main factors are significant.

c)



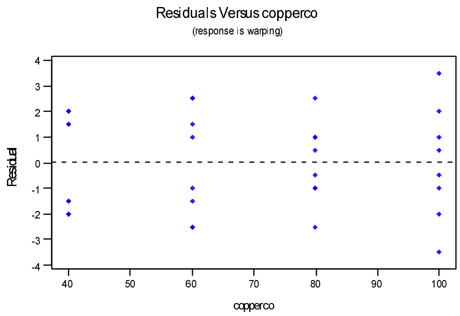


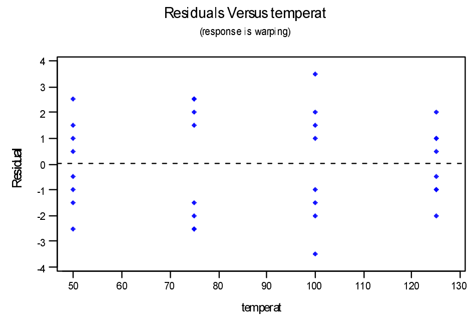


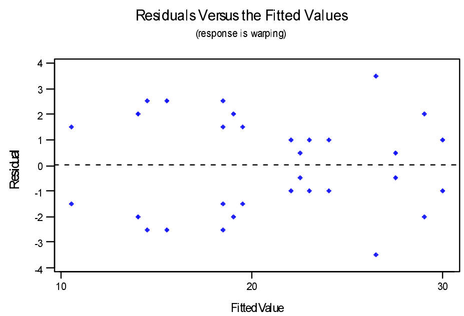
13-5.

```
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)
                 800         825
     825      -518.4
              -445.0
     850       -27.9       453.8
                45.5       527.2
```

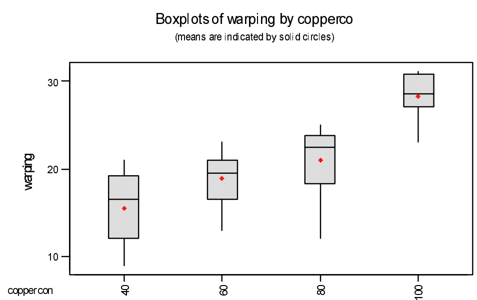
13-6. b)







c)




```
Fisher's pairwise comparisons
Family error rate = 0.195
Individual error rate = 0.0500
Critical value = 2.048
Intervals for (column level mean) - (row level mean)
                  40           60           80
      60       -7.139
                0.389
      80       -9.264       -5.889
               -1.736        1.639
     100      -16.514      -13.139      -11.014
               -8.986       -5.611       -3.486
```

d) No, because the factors do not interact.

13-7. $\hat{\sigma}_{\overline{X}} = \sqrt{\frac{6.78125}{8}} = 0.92$ Same conclusions.

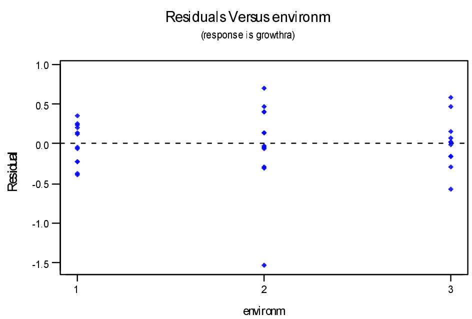
13-8. a)

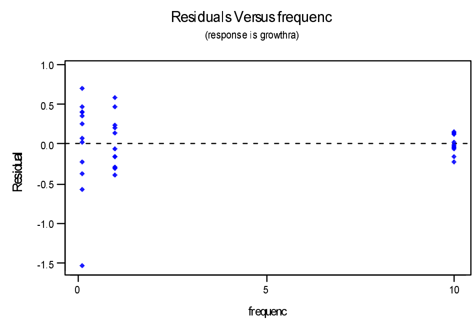
Source	DF	SS	MS	F	P
frequenc	2	209.893	104.946	522.40	0.000
environm	2	64.252	32.126	159.92	0.000
frequenc*environm	4	101.966	25.491	126.89	0.000
Error	27	5.424	0.201		
Total	35	381.535			

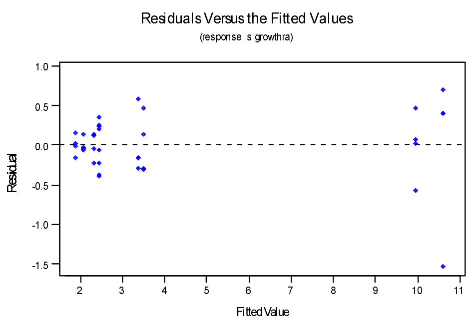
Both main factors and interaction are significant.

b)

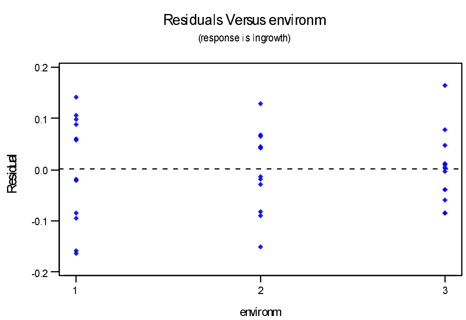
Source	DF	SS	MS	F	P
frequenc	2	7.5702	3.7851	404.09	0.000
environm	2	2.3576	1.1788	125.85	0.000
frequenc*environm	4	3.5284	0.8821	94.17	0.000
Error	27	0.2529	0.0094		
Total	35	13.7092			

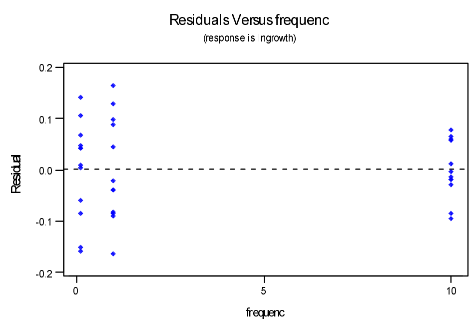


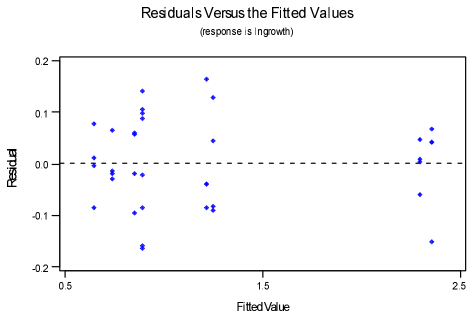




c) Residual plots on the log scale are improved.





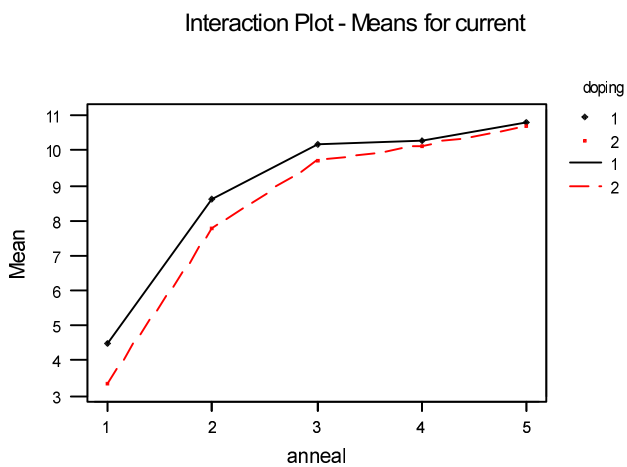


13-9. a)

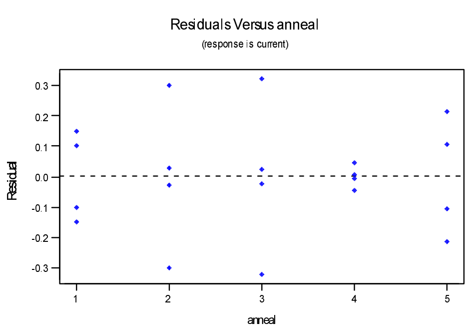
Source	DF	SS	MS	F	P
doping	1	1.442	1.442	25.23	0.000
anneal	4	124.238	31.059	543.52	0.000
doping*anneal	4	0.809	0.202	3.54	0.048
Error	10	0.571	0.057		
Total	19	127.060			

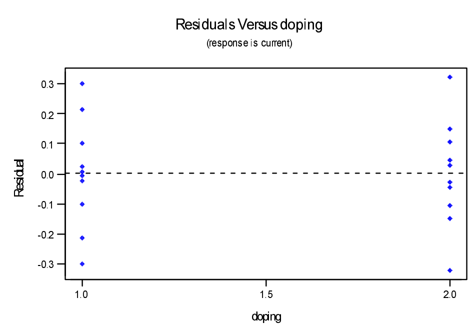
Both main factors and the interaction are significant.

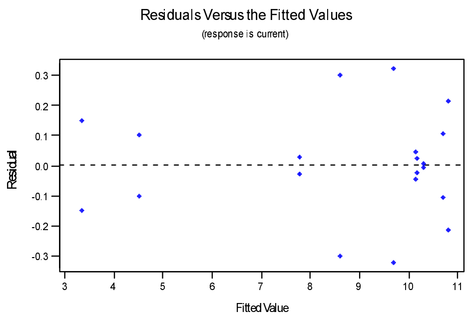
b)



c)







13-10. a)

$$\sigma_{\overline{X}} = \sqrt{\frac{0.057145}{4}} = 0.12$$

$$\pm 3\sigma_{\overline{X}} = \pm 0.36$$

b)

```
Fisher's pairwise comparisons
Family error rate = 0.258
Individual error rate = 0.0500
Critical value = 2.131
Intervals for (column level mean) - (row level mean)
                   1            2            3            4
       2     -4.9187
             -3.6113
       3     -6.6662      -2.4012
             -5.3588      -1.0938
       4     -6.9487      -2.6837      -0.9362
             -5.6413      -1.3763       0.3712
       5     -7.4787      -3.2137      -1.4662      -1.1837
             -6.1713      -1.9063      -0.1588       0.1237
```

c) The conclusions are the same.

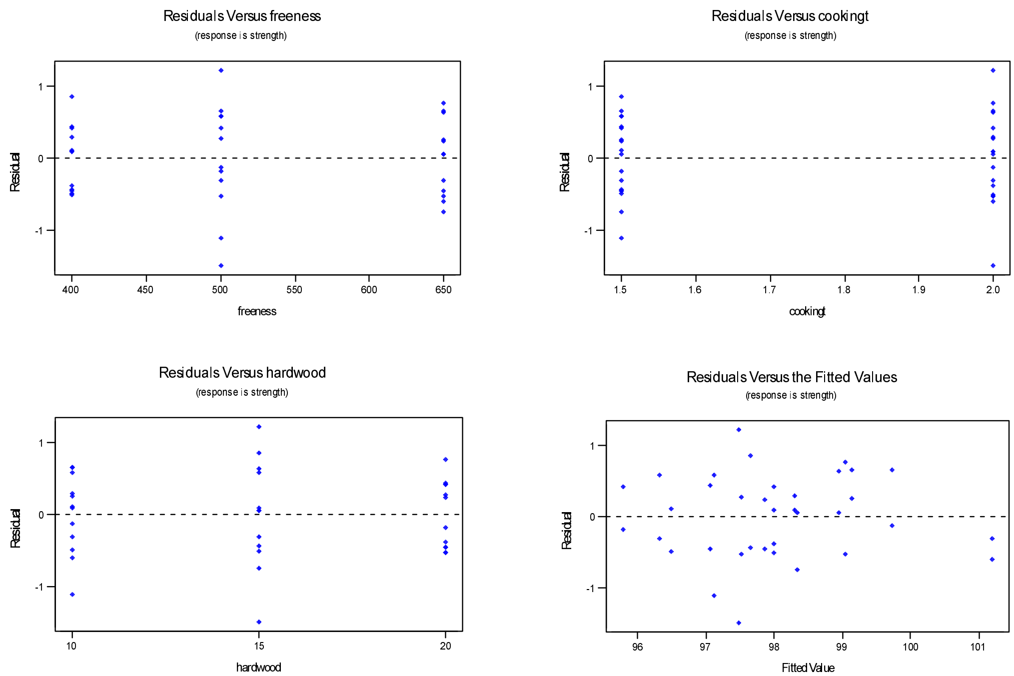
Section 13-5

13-11. a) & b)

Source	DF	SS	MS	F	P
hardwood	2	8.3750	4.1875	7.64	0.003
cookingt	1	17.3611	17.3611	31.66	0.000
freeness	2	21.8517	10.9258	19.92	0.000
hardwood*cookingt	2	3.2039	1.6019	2.92	0.075
hardwood*freeness	4	6.5133	1.6283	2.97	0.042
cookingt*freeness	2	1.0506	0.5253	0.96	0.399
Error	22	12.0644	0.5484		
Total	35	70.4200			

All main factors are significant. The interaction of hardwood*freeness is also significant.

c)

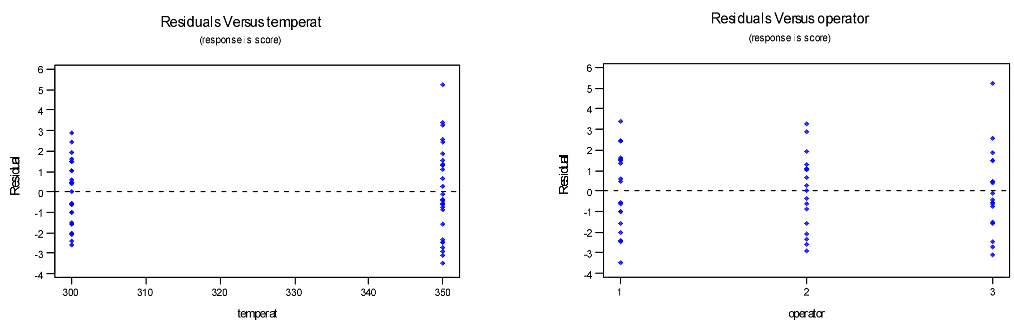


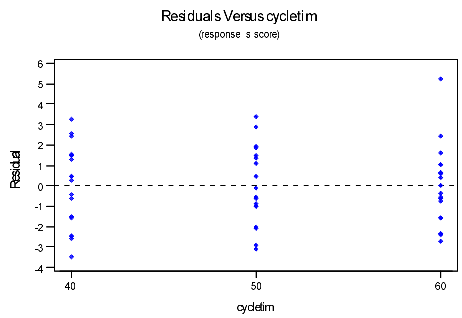
13-12. a)

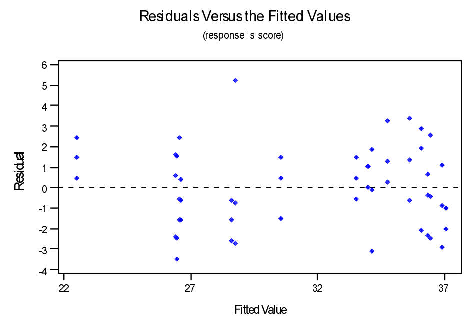
Source	DF	SS	MS	F	P
cycletim	2	393.037	196.519	40.25	0.000
operator	2	256.148	128.074	26.23	0.000
temperat	1	71.185	71.185	14.58	0.000
cycletim*operator	4	307.519	76.880	15.75	0.000
cycletim*temperat	2	67.704	33.852	6.93	0.003
operator*temperat	2	13.481	6.741	1.38	0.263
Error	40	195.296	4.882		
Total	53	1304.370			

Only the operator*temperature interaction is insignificant.

b)







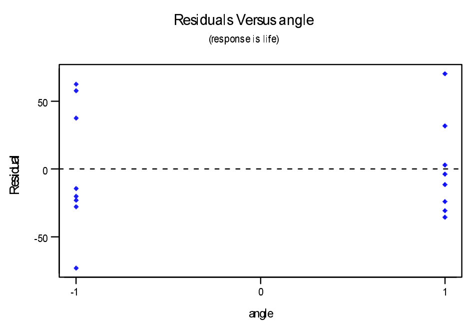
Section 13-6

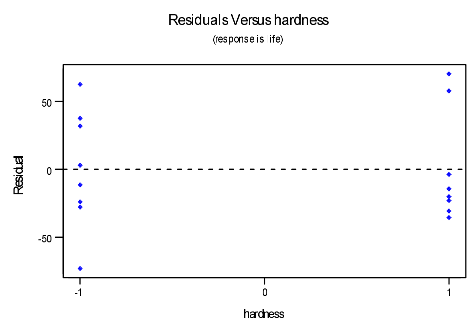
13-13. a)

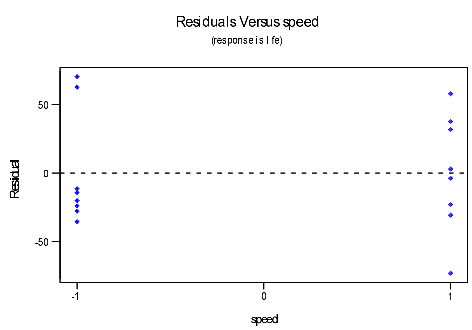
Source	DF	SS	MS	F	P
speed	1	1332	1332	0.49	0.502
hardness	1	28392	28392	10.42	0.010
angle	1	20592	20592	7.56	0.023
speed*hardness	1	506	506	0.19	0.677
speed*angle	1	56882	56882	20.87	0.000
hardness*angle	1	2352	2352	0.86	0.377
Error	9	24530	2726		
Total	15	134588			

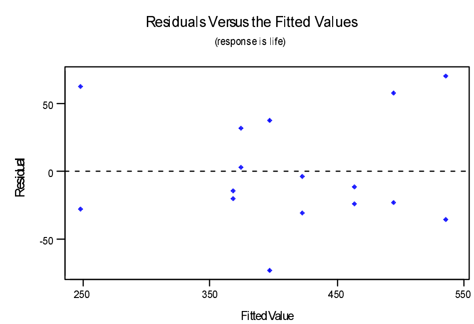
b) $y = 413.125 + 9.125x_{speed} + 45.125x_{hardness} + 35.875x_{angle} - 59.625x_{speed*angle}$

c)







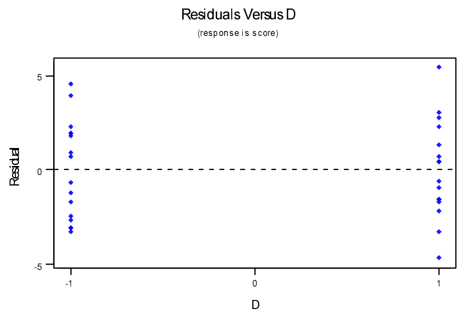


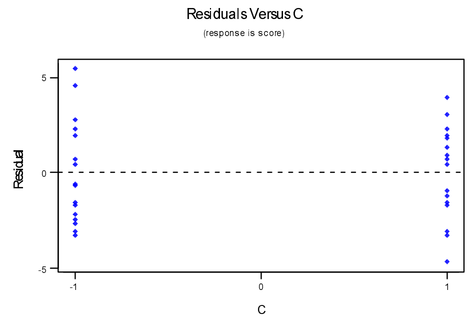
13-14.

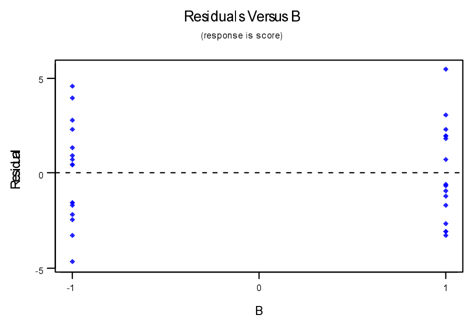
Estimated Effects and Coefficients for score

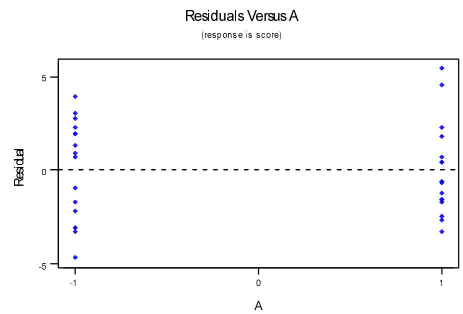
Term	Effect	Coef	StDev Coef	T	P
Constant		175.250	0.5467	320.59	0.000
A	8.250	4.125	0.5467	7.55	0.000
B	-10.375	-5.187	0.5467	-9.49	0.000
C	2.125	1.062	0.5467	1.94	0.070
D	-0.375	-0.187	0.5467	-0.34	0.736
A*B	-0.125	-0.063	0.5467	-0.11	0.910
A*C	-0.625	-0.313	0.5467	-0.57	0.575
A*D	9.125	4.562	0.5467	8.35	0.000
B*C	-0.250	-0.125	0.5467	-0.23	0.822
B*D	1.250	0.625	0.5467	1.14	0.270
C*D	-1.250	-0.625	0.5467	-1.14	0.270
A*B*C	-8.000	-4.000	0.5467	-7.32	0.000
A*B*D	-9.250	-4.625	0.5467	-8.46	0.000
A*C*D	-8.750	-4.375	0.5467	-8.00	0.000
B*C*D	-8.625	-4.312	0.5467	-7.89	0.000
A*B*C*D	-1.625	-0.812	0.5467	-1.49	0.157

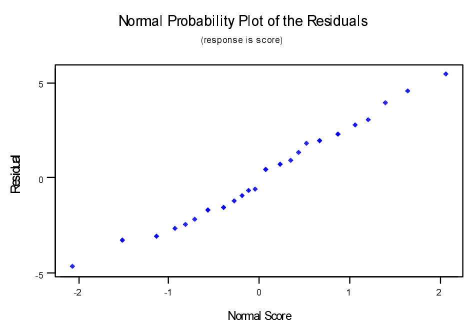
13-5.









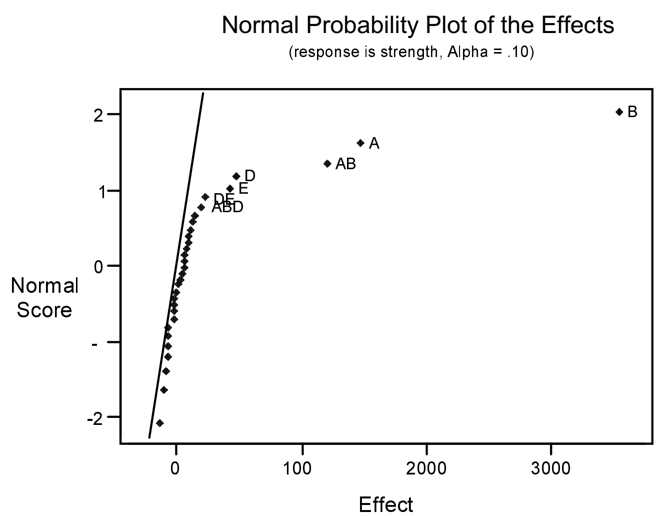


13-16. standard error of the effects = –1.09. Construct approximate 95% confidence intervals on each estimated effect: estimated effect ± 2(standard error). From the intervals, A, C, D, AD, ABC, ABD, ACD, and BCD.

13-17. a)

```
Estimated Effects and Coefficients for strength
Term           Effect
A             1462.13
B             3537.87
C             -137.12
D              474.62
E              425.38
A*B           1199.62
A*C            124.62
A*D             62.87
A*E             62.12
B*C            -99.63
B*D            -12.88
B*E            -12.12
C*D            112.13
C*E            -62.13
D*E            224.62
A*B*C          -62.88
A*B*D          200.38
A*B*E           49.63
A*C*D           75.38
A*C*E           99.63
A*D*E          -87.12
B*C*D           99.62
B*C*E          -74.63
B*D*E          -62.88
C*D*E           37.13
A*B*C*D        -12.12
A*B*C*E         12.13
A*B*D*E          0.37
A*C*D*E        150.37
B*C*D*E        -25.38
A*B*C*D*E       62.88
```

b)



The effects that appear to be important are A, B, D, E, and the interactions AB, DE, and ABD.

c) The regression analysis and final model are

The regression equation is

strength = 2888 + 731 A + 1769 B + 237 D + 213 E + 600 AB + 112 DE + 100 ABD

Predictor	Coef	StDev	T	P
Constant	2887.69	39.10	73.86	0.000
A	731.06	39.10	18.70	0.000
B	1768.94	39.10	45.24	0.000
D	237.31	39.10	6.07	0.000
E	212.69	39.10	5.44	0.000
AB	599.81	39.10	15.34	0.000
DE	112.31	39.10	2.87	0.008
ABD	100.19	39.10	2.56	0.017

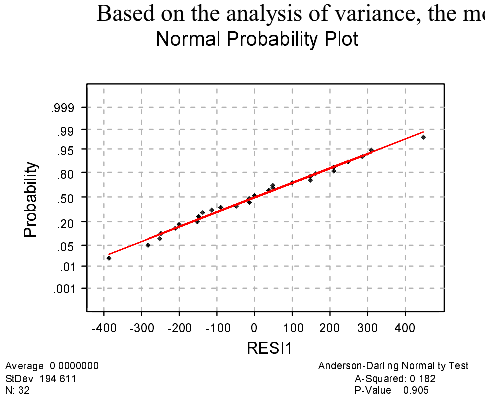
S = 221.2 R-Sq = 99.1% R-Sq(adj) = 98.9%

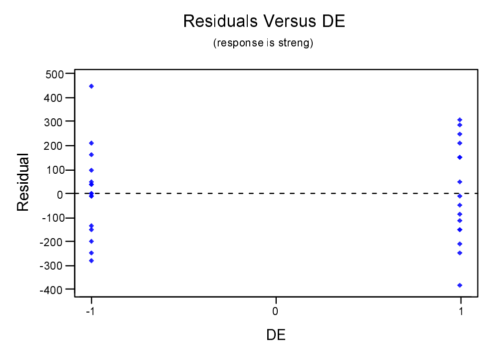
Analysis of Variance

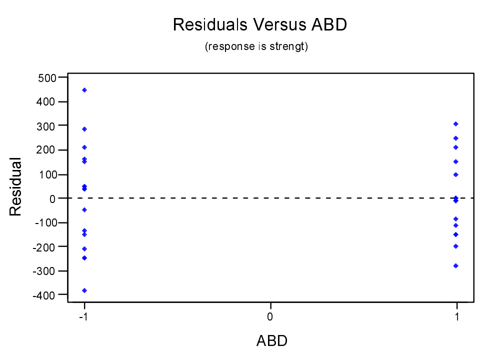
Source	DF	SS	MS	F	P
Regression	7	132722308	18960330	387.58	0.000
Error	24	1174077	48920		
Total	31	133896385			

Source	DF	Seq SS
A	1	17102476
B	1	100132476
D	1	1802151
E	1	1447551
AB	1	11512801
DE	1	403651
ABD	1	321201

Based on the analysis of variance, the model appears to be adequate.







The normal probability plot of the residuals indicates the assumption of normality is not violated. The model appears to be adequate.

d) To maximize strength, the variables A, B, D, and E should be increased. Variable C is not significant thus any level of C would be acceptable.

13-18. a) & b)

Estimated Effects and Coefficients for charge

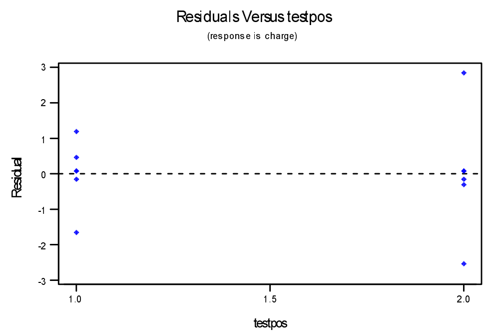
Term	Effect	Coef	StDev Coef	T	P
Constant		-1.000	0.4462	-2.24	0.055
cleanmet	-5.593	-2.797	0.4462	-6.27	0.000
testpos	-1.280	-0.640	0.4462	-1.43	0.189
cleanmet*testpos	-1.220	-0.610	0.4462	-1.37	0.209

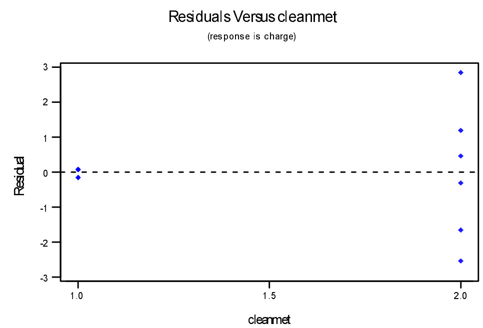
Analysis of Variance for charge

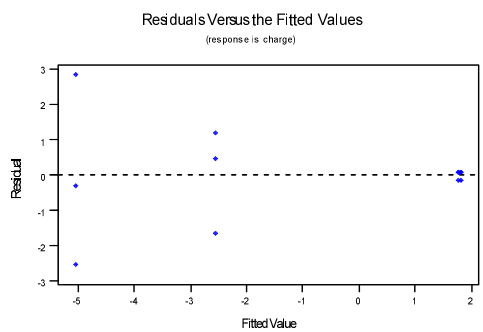
Source	DF	Seq SS	Adj SS	Adj MS	F	P
Main Effects	2	98.771	98.7713	49.386	20.67	0.001
2-Way Interactions	1	4.465	4.4652	4.465	1.87	0.209
Residual Error	8	19.110	19.1101	2.389		
Pure Error	8	19.110	19.1101	2.389		
Total	11	122.347				

Test position is the significant factor.

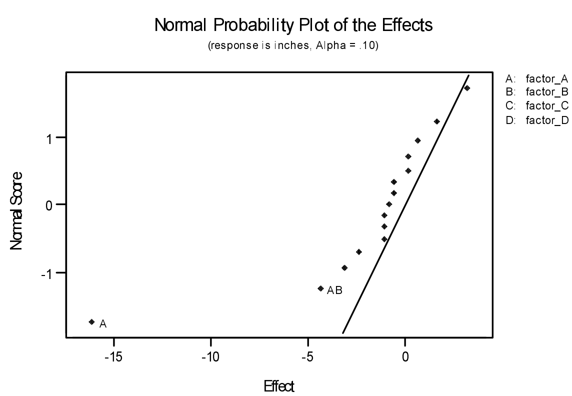
c)







13-19. a)



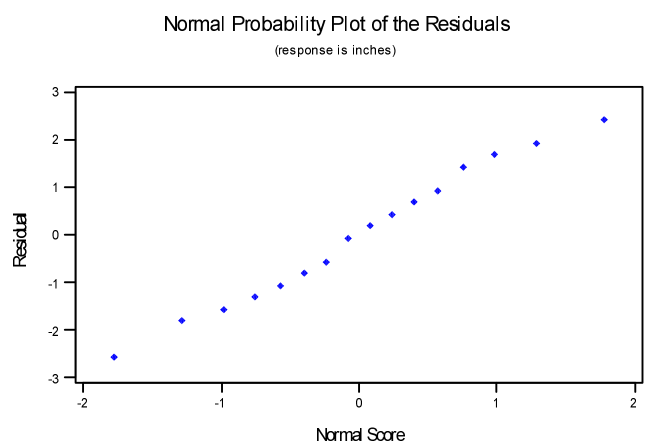
Estimated Effects and Coefficients for inches

Term	Effect	Coef
Constant		35.938
factor_A	-16.125	-8.063
factor_B	3.125	1.562
factor_C	-1.125	-0.562
factor_D	-1.125	-0.562
factor_A*factor_B	-4.375	-2.187
factor_A*factor_C	-0.625	-0.312
factor_A*factor_D	-3.125	-1.562
factor_B*factor_C	1.625	0.812
factor_B*factor_D	0.125	0.063
factor_C*factor_D	-0.625	-0.313
factor_A*factor_B*factor_C	0.625	0.313
factor_A*factor_B*factor_D	-2.375	-1.188
factor_A*factor_C*factor_D	-1.125	-0.563
factor_B*factor_C*factor_D	-0.875	-0.437
factor_A*factor_B*factor_C *factor_D	0.125	0.063

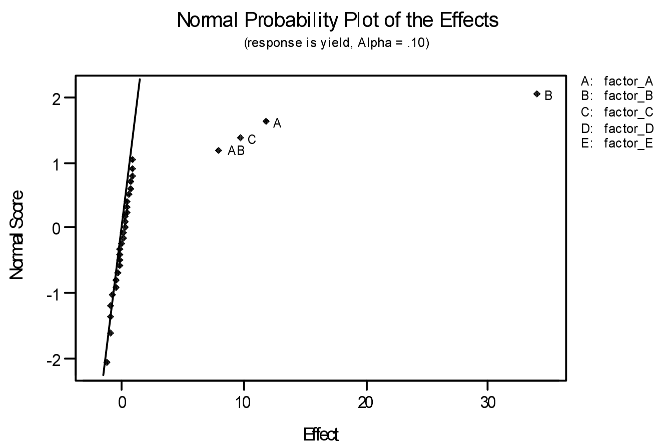
According to the normal probability plot, factors A, B, and AB appear to be significant.

b) & c) Removing the three and four factor interactions the analysis and residual plot are found:

Term	Effect	Coef	StDev Coef	T	P
Constant		35.938	0.6355	56.55	0.000
factor_A	-16.125	-8.063	0.6355	-12.69	0.000
factor_B	3.125	1.562	0.6355	2.46	0.057
factor_C	-1.125	-0.562	0.6355	-0.89	0.417
factor_D	-1.125	-0.562	0.6355	-0.89	0.417
factor_A*factor_B	-4.375	-2.187	0.6355	-3.44	0.018
factor_A*factor_C	-0.625	-0.312	0.6355	-0.49	0.644
factor_A*factor_D	-3.125	-1.562	0.6355	-2.46	0.057
factor_B*factor_C	1.625	0.812	0.6355	1.28	0.257
factor_B*factor_D	0.125	0.063	0.6355	0.10	0.925
factor_C*factor_D	-0.625	-0.313	0.6355	-0.49	0.644



13-20. a)



Factors A, B, C, and the interaction AB appear to be significant.

b)

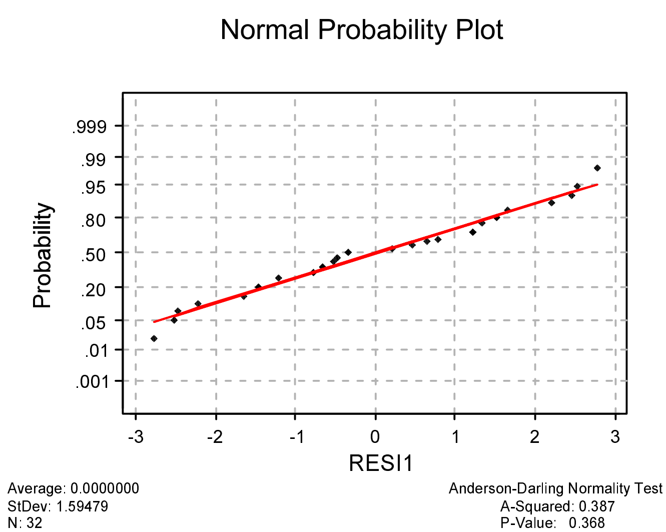
Term	Effect	Coef	StDev Coef	T	P
Constant		30.5312	0.2786	109.57	0.000
factor_A	11.8125	5.9063	0.2786	21.20	0.000
factor_B	33.9375	16.9687	0.2786	60.90	0.000
factor_C	9.6875	4.8437	0.2786	17.38	0.000
factor_D	-0.8125	-0.4063	0.2786	-1.46	0.164
factor_E	0.4375	0.2187	0.2786	0.79	0.444
factor_A*factor_B	7.9375	3.9687	0.2786	14.24	0.000
factor_A*factor_C	0.4375	0.2187	0.2786	0.79	0.444
factor_A*factor_D	-0.0625	-0.0313	0.2786	-0.11	0.912
factor_A*factor_E	0.9375	0.4687	0.2786	1.68	0.112
factor_B*factor_C	0.0625	0.0313	0.2786	0.11	0.912
factor_B*factor_D	-0.6875	-0.3437	0.2786	-1.23	0.235
factor_B*factor_E	0.5625	0.2813	0.2786	1.01	0.328
factor_C*factor_D	0.8125	0.4062	0.2786	1.46	0.164
factor_C*factor_E	0.3125	0.1563	0.2786	0.56	0.583
factor_D*factor_E	-1.1875	-0.5938	0.2786	-2.13	0.049

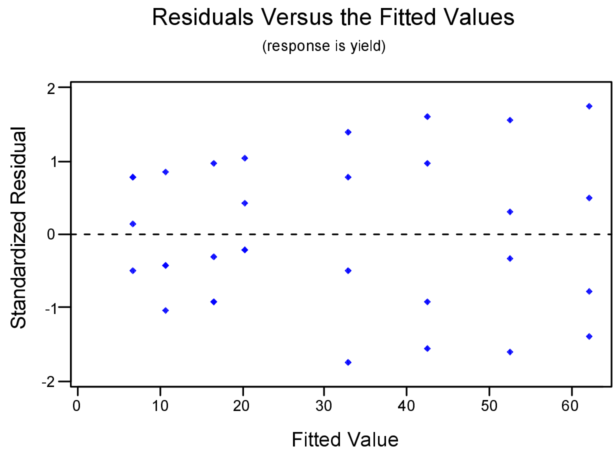
Analysis of Variance for yield

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Main Effects	5	11087.9	11087.9	2217.58	892.61	0.000
2-Way Interactions	10	536.3	536.3	53.63	21.59	0.000
Residual Error	16	39.7	39.7	2.48		
Total	31	11664.0				

The analysis confirms our findings from part a)

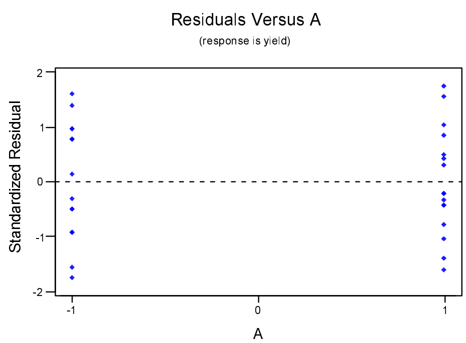
c)

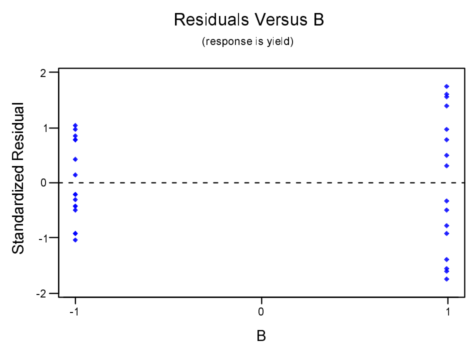


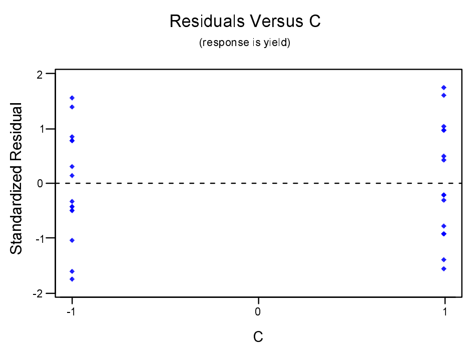


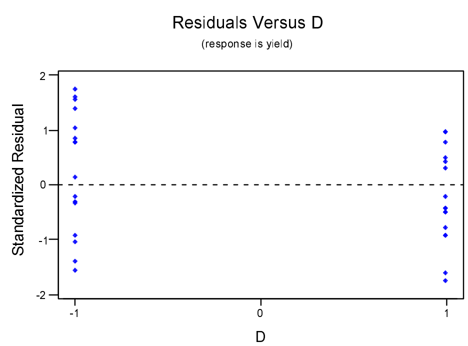
The residuals appear normally distributed, however their variance appears to increase as the fitted value increases.

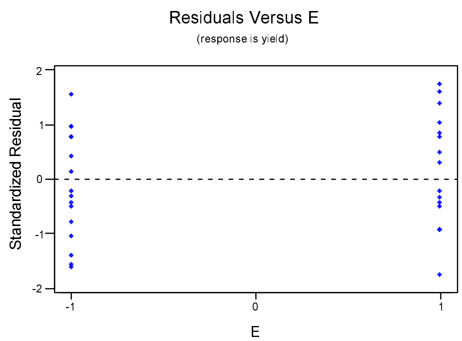
d)





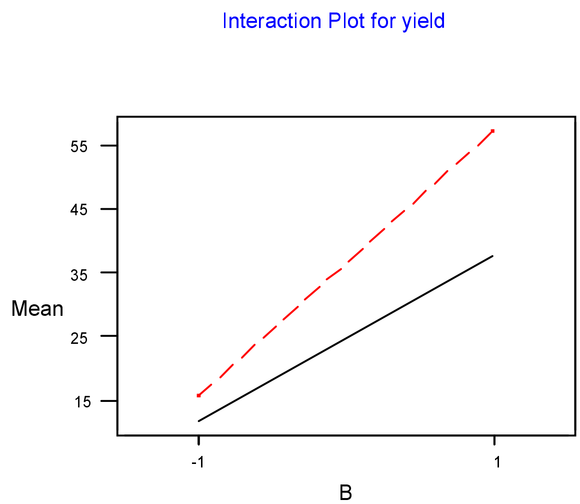






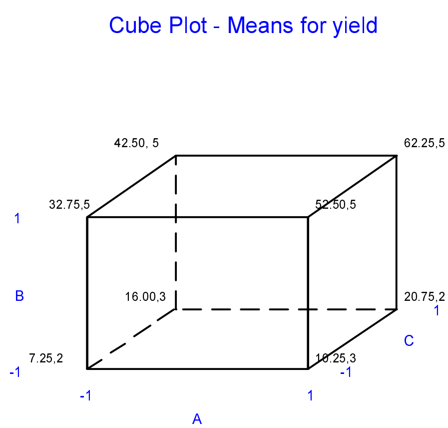
All plots support the constant variance assumption except B.

e) The interaction AB appears to be significant. The interaction plot from MINITAB indicates that a high level of A and of B increases the mean yield. While low levels of each would lead to a reduction in the mean yield.



f) To increase yield, set A, B, and C at their high levels.

g)



It is evident from the plot that we should run the process with all factors set at their high level.

13-21. a) Note to Instructor: If insignificant terms are removed, Minitab will include a 'Lack of Fit' term in the analysis output; and, the $MSE_{curvature}$ computed by hand will not be equivalent to $MSE_{curvature}$ computed by Minitab.

```
Analysis of Variance for inches
Source                DF      Seq SS     Adj SS     Adj MS       F       P
Main Effects           5     11087.9    11087.9    2217.58   15.25   0.005
2-Way Interactions    10       536.3      536.3      53.63    0.37   0.916
3-Way Interactions    10        24.3       24.3       2.43    0.02   1.000
4-Way Interactions     5        15.2       15.2       3.03    0.02   1.000
5-Way Interactions     1         0.3        0.3       0.28    0.00   0.967
Residual Error         5       726.8      726.8     145.37
  Curvature            1       694.0      694.0     694.04   84.64   0.001
  Pure Error           4        32.8       32.8       8.20
Total                 36     12390.8
```

The hand calculation of $\hat{\sigma}_c^2$ is

$$\hat{\sigma}_c^2 = \sum_{\substack{\text{center}\\ \text{pts}}} \frac{(y_i - \overline{y}_c)^2}{(n_c - 1)} = \frac{32.8}{4} = 8.2 \qquad \text{where } \overline{y}_c = 43.2$$

The difference between the this estimate and the estimate in Exercise 13-20 is due to the addition of centerpoints.

13-22. a)

```
Estimated Effects and Coefficients for y
Term                          Effect      Coef  StDev Coef        T       P
Constant                                414.58       4.316    96.06   0.000
factor_A                       13.75      6.87       5.286     1.30   0.263
factor_B                      127.75     63.87       5.286    12.08   0.000
factor_C                       97.75     48.87       5.286     9.25   0.001
factor_A*factor_B             -21.25    -10.62       5.286    -2.01   0.115
factor_A*factor_C            -137.25    -68.62       5.286   -12.98   0.000
factor_B*factor_C             -52.25    -26.13       5.286    -4.94   0.008
factor_A*factor_B*factor_C    -68.25    -34.12       5.286    -6.46   0.003

Analysis of Variance for y
Source              DF      Seq SS     Adj SS     Adj MS       F       P
Main Effects         3       52128    52128.4    17376.1   77.74   0.001
2-Way Interactions   3       44038    44038.4    14679.5   65.68   0.001
3-Way Interactions   1        9316     9316.1     9316.1   41.68   0.003
Residual Error       4         894      894.0      223.5
  Curvature          1         176      176.0      176.0    0.74   0.454
  Pure Error         3         718      718.0      239.3
Total               11      106377
```

Section 13-7

13-23.

BLOCK	A	B	C	y
1	-1	-1	-1	221
1	1	1	-1	552
1	1	-1	1	406
1	-1	1	1	605
2	1	-1	-1	325
2	-1	1	-1	354
2	-1	-1	1	440
2	1	1	1	392

```
Term                  Effect      Coef
Constant                        411.87
Block                            34.12
factor_A               13.75      6.87
factor_B              127.75     63.87
factor_C               97.75     48.87
factor_A*factor_B     -21.25    -10.63
factor_A*factor_C    -137.25    -68.63
factor_B*factor_C     -52.25    -26.13
Term                  Effect      Coef  StDev Coef       T      P
Constant                        411.87       19.94   20.65  0.002
Block                            34.12       19.94    1.71  0.229
factor_A               13.75      6.87       19.94    0.34  0.763
factor_B              127.75     63.87       19.94    3.20  0.085
factor_C               97.75     48.87       19.94    2.45  0.134
factor_A*factor_C    -137.25    -68.63       19.94   -3.44  0.075

Analysis of Variance for y
Source                DF      Seq SS      Adj SS      Adj MS       F       P
Blocks                 1        9316        9316        9316    2.93   0.229
Main Effects           3       52128       52128       17376    5.46   0.159
2-Way Interactions     1       37675       37675       37675   11.84   0.075
Residual Error         2        6363        6363        3182
Total                  7      105483
```

13-24.

BLOCK	A	B	C	D	var_1
1	1	-1	-1	-1	174
1	-1	1	-1	-1	181
1	-1	-1	1	-1	177
1	1	1	1	-1	173
1	-1	-1	-1	1	198
1	1	1	-1	1	185
1	1	-1	1	1	179
1	-1	1	1	1	187
2	-1	-1	-1	-1	190
2	1	1	-1	-1	183
2	1	-1	1	-1	181
2	-1	1	1	-1	188
2	1	-1	-1	1	172
2	-1	1	-1	1	187
2	-1	-1	1	1	199
2	1	1	1	1	180

```
Term                           Effect       Coef
Constant                                 183.375
Block                                     -1.625
factor_A                      -10.000     -5.000
factor_B                       -0.750     -0.375
factor_C                       -0.750     -0.375
factor_D                        5.000      2.500
factor_A*factor_B               4.500      2.250
factor_A*factor_C               0.500      0.250
factor_A*factor_D              -3.750     -1.875
factor_B*factor_C              -1.250     -0.625
factor_B*factor_D              -1.500     -0.750
factor_C*factor_D               1.500      0.750
factor_A*factor_B*factor_C     -6.000     -3.000
factor_A*factor_B*factor_D      4.750      2.375
factor_A*factor_C*factor_D     -0.250     -0.125
factor_B*factor_C*factor_D     -2.000     -1.000
```

```
Term         Effect      Coef  StDev Coef       T      P
Constant              183.375       1.607  114.14  0.000
Block                  -1.625       1.607   -1.01  0.336
factor_A    -10.000    -5.000       1.607   -3.11  0.011
factor_B     -0.750    -0.375       1.607   -0.23  0.820
factor_C     -0.750    -0.375       1.607   -0.23  0.820
factor_D      5.000     2.500       1.607    1.56  0.151
Analysis of Variance for var_1
Source              DF      Seq SS     Adj SS     Adj MS      F       P
Blocks               1       42.25      42.25      42.25   1.02   0.336
Main Effects         4      504.50     504.50     126.13   3.05   0.069
Residual Error      10      413.00     413.00      41.30
Total               15      959.75
```

13-25.

BLOCK	A	B	C	D	var_1
1	-1	-1	-1	-1	190
1	1	1	1	-1	173
1	1	-1	-1	1	172
1	-1	1	1	1	187
2	-1	1	-1	-1	181
2	1	-1	1	-1	181
2	1	1	-1	1	185
2	-1	-1	1	1	199
3	1	1	-1	-1	183
3	-1	-1	1	-1	177
3	-1	1	-1	1	187
3	1	-1	1	1	179
4	1	-1	-1	-1	174
4	-1	1	1	-1	188
4	-1	-1	-1	1	198
4	1	1	1	1	189

```
Term                            Effect       Coef  StDev Coef       T      P
Constant                                  183.938       1.958   93.96  0.000
Block  1                                   -3.437       3.391   -1.01  0.417
       2                                    2.563       3.391    0.76  0.529
       3                                   -2.438       3.391   -0.72  0.547
factor_A                        -8.875     -4.438       1.958   -2.27  0.152
factor_B                         0.375      0.187       1.958    0.10  0.932
factor_C                         0.375      0.187       1.958    0.10  0.932
factor_D                         6.125      3.062       1.958    1.56  0.258
factor_A*factor_B                5.625      2.812       1.958    1.44  0.287
factor_A*factor_C                1.625      0.813       1.958    0.42  0.718
factor_A*factor_D               -2.625     -1.312       1.958   -0.67  0.572
factor_B*factor_D               -0.375     -0.188       1.958   -0.10  0.932
factor_B*factor_C*factor_D      -0.875     -0.438       1.958   -0.22  0.844
factor_A*factor_B*factor_C
*factor_D                        4.375      2.187       1.958    1.12  0.380

Analysis of Variance for var_1
Source               DF     Seq SS     Adj SS     Adj MS      F       P
Blocks                3    141.187    141.187     47.062   0.77   0.608
Main Effects          4    466.250    466.250    116.563   1.90   0.373
2-Way Interactions    4    165.250    165.250     41.313   0.67   0.670
3-Way Interactions    1      3.063      3.063      3.063   0.05   0.844
4-Way Interactions    1     76.562     76.562     76.562   1.25   0.380
Residual Error        2    122.625    122.625     61.312
Total                15    974.938
```

13-26.

Run	Block	A	B	C	D	E
1	1	-	-	-	-	-
2	1	+	+	-	-	-
3	1	+	-	+	-	-
4	1	-	+	+	-	-
5	1	+	-	-	+	-
6	1	-	+	-	+	-
7	1	-	-	+	+	-
8	1	+	+	+	+	-
9	1	+	-	-	-	+
10	1	-	+	-	-	+
11	1	-	-	+	-	+
12	1	+	+	+	-	+
13	1	-	-	-	+	+
14	1	+	+	-	+	+
15	1	+	-	+	+	+
16	1	-	+	+	+	+
17	2	+	-	-	-	-
18	2	-	+	-	-	-
19	2	-	-	+	-	-
20	2	+	+	+	-	-
21	2	-	-	-	+	-
22	2	+	+	-	+	-
23	2	+	-	+	+	-
24	2	-	+	+	+	-
25	2	-	-	-	-	+
26	2	+	+	-	-	+
27	2	+	-	+	-	+
28	2	-	+	+	-	+
29	2	+	-	-	+	+
30	2	-	+	-	+	+
31	2	-	-	+	+	+
32	2	+	+	+	+	+

13-27.

Run	Block	A	B	C	D	E
1	1	-	-	-	-	-
2	1	+	+	-	-	-
3	1	+	-	+	+	-
4	1	-	+	+	+	-
5	1	+	-	+	-	+
6	1	-	+	+	-	+
7	1	-	-	-	+	+
8	1	+	+	-	+	+
9	2	+	-	-	-	-
10	2	-	+	-	-	-
11	2	-	-	+	+	-
12	2	+	+	+	+	-
13	2	-	-	+	-	+
14	2	+	+	+	-	+
15	2	+	-	-	+	+
16	2	-	+	-	+	+
17	3	+	-	+	-	-
18	3	-	+	+	-	-
19	3	-	-	-	+	-
20	3	+	+	-	+	-
21	3	-	-	-	-	+
22	3	+	+	-	-	+
23	3	+	-	+	+	+
24	3	-	+	+	+	+
25	4	-	-	+	-	-
26	4	+	+	+	-	-
27	4	+	-	-	+	-
28	4	-	+	-	+	-
29	4	+	-	-	-	+
30	4	-	+	-	-	+
31	4	-	-	+	+	+
32	4	+	+	+	+	+

13-28.

Term	Effect	Coef
Constant		35.938
Block		-0.063
factor_A	-16.125	-8.062
factor_B	3.125	1.563
factor_C	-1.125	-0.563
factor_D	-1.125	-0.563
factor_A*factor_B	-4.375	-2.187
factor_A*factor_C	-0.625	-0.313
factor_A*factor_D	-3.125	-1.563
factor_B*factor_C	1.625	0.813
factor_B*factor_D	0.125	0.063
factor_C*factor_D	-0.625	-0.312
factor_A*factor_B*factor_C	0.625	0.313
factor_A*factor_B*factor_D	-2.375	-1.187
factor_A*factor_C*factor_D	-1.125	-0.562
factor_B*factor_C*factor_D	-0.875	-0.438

Estimated Effects and Coefficients for var_1

Term	Effect	Coef	StDev Coef	T	P
Constant		35.938	0.7354	48.87	0.000
Block		-0.063	0.7354	-0.08	0.934
factor_A	-16.125	-8.062	0.7354	-10.96	0.000
factor_B	3.125	1.563	0.7354	2.12	0.057
factor_A*factor_B	-4.375	-2.187	0.7354	-2.97	0.013

Analysis of Variance for var_1

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Blocks	1	0.06	0.06	0.063	0.01	0.934
Main Effects	2	1079.12	1079.12	539.562	62.35	0.000
2-Way Interactions	1	76.56	76.56	76.562	8.85	0.013
Residual Error	11	95.19	95.19	8.653		
Lack of Fit	3	9.69	9.69	3.229	0.30	0.823
Pure Error	8	85.50	85.50	10.687		
Total	15	1250.94				

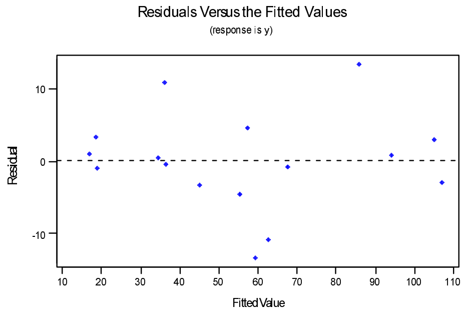
13-29. a) Estimated Effects and Coefficients for y

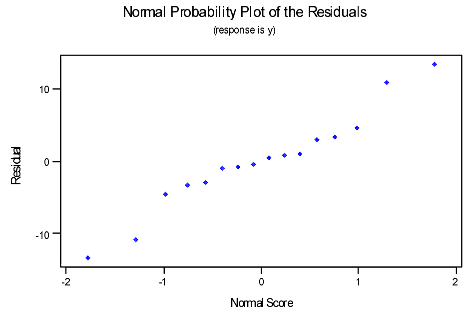
Term	Effect	Coef	StDev Coef	T	P
Constant		56.37	2.633	21.41	0.000
Block 1		15.63	4.560	3.43	0.014
2		-3.38	4.560	-0.74	0.487
3		-10.88	4.560	-2.38	0.054
A	-45.25	-22.62	2.633	-8.59	0.000
B	-1.50	-0.75	2.633	-0.28	0.785
C	14.50	7.25	2.633	2.75	0.033
A*B	19.00	9.50	2.633	3.61	0.011
A*C	-14.50	-7.25	2.633	-2.75	0.033
B*C	-9.25	-4.63	2.633	-1.76	0.130

Analysis of Variance for y

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Blocks	3	1502.8	1502.8	500.9	4.52	0.055
Main Effects	3	9040.2	9040.2	3013.4	27.17	0.001
2-Way Interactions	3	2627.2	2627.2	875.7	7.90	0.017
Residual Error	6	665.5	665.5	110.9		
Total	15	13835.7				

b)



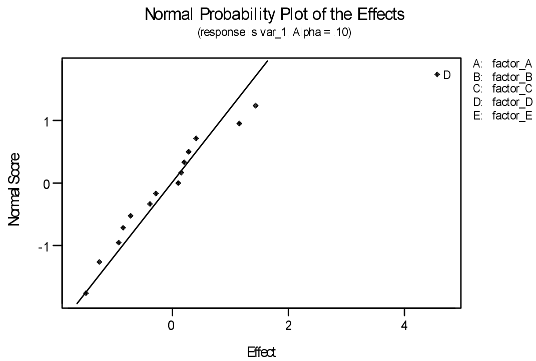


13-30.

Run	Block	A	B	C	D	E	F
1	1	-	-	-	+	+	-
2	1	+	-	+	+	-	-
3	1	-	+	-	-	+	+
4	1	-	+	+	+	-	-
5	1	+	+	-	+	+	-
6	1	-	-	+	+	+	+
7	1	+	-	+	-	+	-
8	1	+	+	+	-	-	+
9	1	-	-	+	-	-	+
10	1	+	+	-	-	-	-
11	1	-	+	+	-	+	-
12	1	+	-	-	-	+	+
13	1	-	-	-	-	-	-
14	1	-	+	-	+	-	+
15	1	+	+	+	+	+	+
16	1	+	-	-	+	-	+
17	2	-	+	+	-	+	+
18	2	+	+	+	-	-	-
19	2	-	-	-	-	-	+
20	2	-	+	+	+	-	+
21	2	+	+	-	-	-	+
22	2	+	-	-	+	-	-
23	2	-	-	-	+	+	+
24	2	-	-	+	+	+	-
25	2	-	-	+	-	-	-
26	2	+	-	+	-	+	+
27	2	-	+	-	+	-	-
28	2	+	-	+	+	-	+
29	2	+	-	-	-	+	-
30	2	+	+	+	+	+	-
31	2	-	+	-	-	+	-
32	2	+	+	-	+	+	+
33	3	+	+	+	+	-	+
34	3	-	-	-	-	+	-
35	3	-	+	+	-	-	-
36	3	-	-	+	-	+	+
37	3	-	+	-	-	-	+
38	3	-	-	+	+	-	+
39	3	+	-	-	+	+	+
40	3	+	-	+	+	+	-
41	3	+	-	+	-	-	-
42	3	+	+	-	-	+	-
43	3	-	+	+	+	+	-
44	3	-	+	-	+	+	+
45	3	+	-	-	-	-	+
46	3	+	+	-	+	-	-
47	3	-	-	-	+	-	-
48	3	+	+	+	-	+	+
49	4	+	-	+	+	+	+
50	4	-	-	-	-	+	+
51	4	+	-	-	-	-	-
52	4	-	-	+	+	-	-
53	4	-	+	+	+	+	+
54	4	-	+	-	-	-	-
55	4	+	-	+	-	-	+
56	4	+	+	-	+	-	+
57	4	-	+	-	+	+	-
58	4	+	+	+	-	+	-
59	4	+	+	-	-	+	+
60	4	-	-	+	-	+	-
61	4	-	+	+	-	-	+
62	4	+	-	-	+	+	-
63	4	-	-	-	+	-	+
64	4	+	+	+	+	-	-

Section 13-8

13-31. a)



c)

Estimated Effects and Coefficients for var_1

Term	Effect	Coef	StDev Coef	T	P
Constant		2.7700	0.2674	10.36	0.000
factor_A	1.4350	0.7175	0.2674	2.68	0.025
factor_B	-1.4650	-0.7325	0.2674	-2.74	0.023
factor_D	4.5450	2.2725	0.2674	8.50	0.000
factor_A*factor_B	1.1500	0.5750	0.2674	2.15	0.060
factor_A*factor_D	-1.2300	-0.6150	0.2674	-2.30	0.047
factor_B*factor_D	0.1200	0.0600	0.2674	0.22	0.827

Analysis of Variance for var_1

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Main Effects	3	99.450	99.4499	33.1500	28.97	0.000
2-Way Interactions	3	11.399	11.3992	3.7997	3.32	0.071
Residual Error	9	10.300	10.2997	1.1444		
Lack of Fit	1	0.533	0.5329	0.5329	0.44	0.527
Pure Error	8	9.767	9.7668	1.2208		
Total	15	121.149				

Collapse using A, B, D only:

Estimated Effects and Coefficients for var_1

Term	Effect	Coef	StDev Coef	T	P
Constant		2.7700	0.2762	10.03	0.000
factor_A	1.4350	0.7175	0.2762	2.60	0.032
factor_B	-1.4650	-0.7325	0.2762	-2.65	0.029
factor_D	4.5450	2.2725	0.2762	8.23	0.000
factor_A*factor_B	1.1500	0.5750	0.2762	2.08	0.071
factor_A*factor_D	-1.2300	-0.6150	0.2762	-2.23	0.057
factor_B*factor_D	0.1200	0.0600	0.2762	0.22	0.833
factor_A*factor_B*factor_D	-0.3650	-0.1825	0.2762	-0.66	0.527

Analysis of Variance for var_1

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Main Effects	3	99.450	99.4499	33.1500	27.15	0.000
2-Way Interactions	3	11.399	11.3992	3.7997	3.11	0.088
3-Way Interactions	1	0.533	0.5329	0.5329	0.44	0.527
Residual Error	8	9.767	9.7668	1.2208		
Pure Error	8	9.767	9.7668	1.2208		
Total	15	121.149				

13-32. a)Design Generators: D = ABC

Alias Structure

I + ABCD
A + BCD
B + ACD
C + ABD
D + ABC
AB + CD
AC + BD
AD + BC

b)
```
Term          Effect      Coef
Constant                70.750
A            19.000      9.500
B             1.500      0.750
C            14.000      7.000
D            16.500      8.250
A*B          -1.000     -0.500
A*C         -18.500     -9.250
A*D          19.000      9.500
```
A, C, D, AC, and AD have large estimated effects.

c)
```
Estimated Effects and Coefficients for rate
Term          Effect      Coef  StDev Coef       T       P
Constant                70.750      0.6374  111.00  0.000
A            19.000      9.500      0.6374   14.90  0.004
C            14.000      7.000      0.6374   10.98  0.008
D            16.500      8.250      0.6374   12.94  0.006
A*C         -18.500     -9.250      0.6374  -14.51  0.005
A*D          19.000      9.500      0.6374   14.90  0.004

Analysis of Variance for rate
Source              DF      Seq SS     Adj SS     Adj MS      F       P
Main Effects         3     1658.50    1658.50    552.833 170.10  0.006
2-Way Interactions   2     1406.50    1406.50    703.250 216.38  0.005
Residual Error       2        6.50       6.50      3.250
Total                7     3071.50
```

13-34. The generators are E=ABC, F=BCD. Complete defining relation is I =ABCE=BCDF=ADEF. Since the remaining factors A, B, D, and F do not form a word in the complete defining relation it can be verified that the resulting design is a full-factorial.

13-35. Since factors A, B, C, and E form a word in the complete defining relation, it can be verified that the resulting design is two replicates of a 2^{4-1} fractional factorial.

13-36.

Factor	Estimated Effect
A	15.75
B	0.75
C	9.75
D	8.75
AB+CD	-3.75
AC+BD	0.25
AD+BC	9.25

13-37. Generators D=AB, E=AC for 2^{5-2}, Resolution III

A	B	C	D	E	var_1
-1	-1	-1	1	1	1900
1	-1	-1	-1	-1	900
-1	1	-1	-1	1	3500
1	1	-1	1	-1	6100
-1	-1	1	1	-1	800
1	-1	1	-1	1	1200
-1	1	1	-1	-1	3000
1	1	1	1	1	6800

```
Term                   Effect      Coef
Constant                        3025.00
factor_A              1450.00    725.00
factor_B              3650.00   1825.00
factor_C              -150.00    -75.00
factor_D              1750.00    875.00
factor_E               650.00    325.00
factor_B*factor_C      250.00    125.00
factor_B*factor_E      -50.00    -25.00
```

13_38. Design Generators: D = AB E = AC F = BC

Defining Relation: I = ABD = ACE = BCF = BCDE = ACDF = ABEF = DEF

A + BD + CE
B + AD + CF
C + AE + BF
D + AB + EF
E + AC + DF
F + BC + DE
AF + BE + CD

Data Matrix

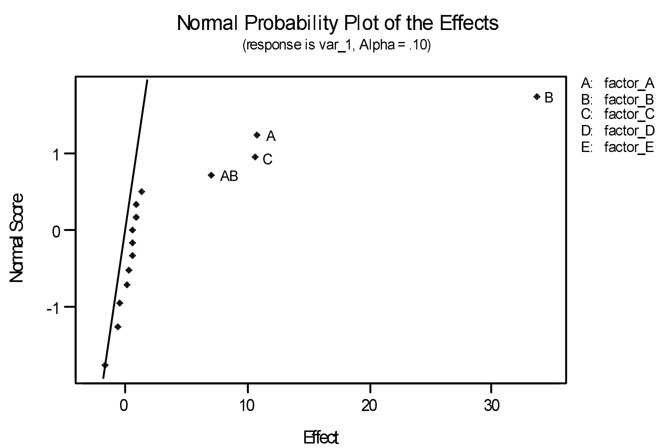
Run	A	B	C	D	E	F
1	-	-	-	+	+	+
2	+	-	-	-	-	+
3	-	+	-	-	+	-
4	+	+	-	+	-	-
5	-	-	+	+	-	-
6	+	-	+	-	+	-
7	-	+	+	-	-	+
8	+	+	+	+	+	+

13-40. b) E=ABCD

c)

Term	Effect	Coef
Constant		30.4375
factor_A	10.8750	5.4375
factor_B	33.6250	16.8125
factor_C	10.6250	5.3125
factor_D	-0.6250	-0.3125
factor_E	0.3750	0.1875
factor_A*factor_B	7.1250	3.5625
factor_A*factor_C	0.6250	0.3125
factor_A*factor_D	0.8750	0.4375
factor_A*factor_E	1.3750	0.6875
factor_B*factor_C	0.8750	0.4375
factor_B*factor_D	-0.3750	-0.1875
factor_B*factor_E	0.1250	0.0625
factor_C*factor_D	0.6250	0.3125
factor_C*factor_E	0.6250	0.3125
factor_D*factor_E	-1.6250	-0.8125

d)



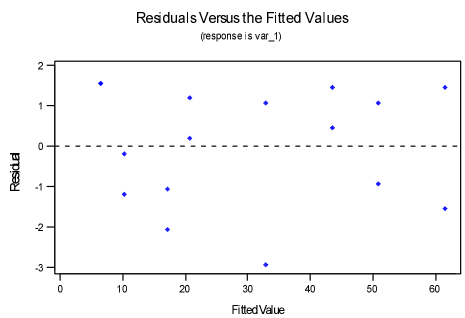
e)

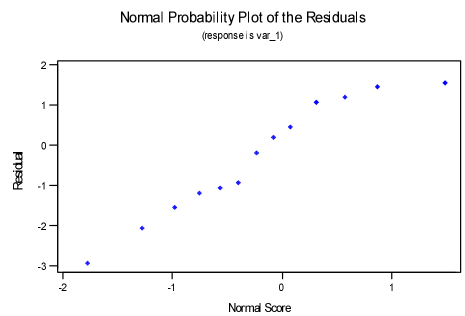
Term	Effect	Coef	StDev Coef	T	P
Constant		30.438	0.4243	71.73	0.000
factor_A	10.875	5.438	0.4243	12.81	0.000
factor_B	33.625	16.812	0.4243	39.62	0.000
factor_C	10.625	5.313	0.4243	12.52	0.000
factor_A*factor_B	7.125	3.562	0.4243	8.40	0.000

Analysis of Variance for var_1

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Main Effects	3	5447.19	5447.19	1815.73	630.31	0.000
2-Way Interactions	1	203.06	203.06	203.06	70.49	0.000
Residual Error	11	31.69	31.69	2.88		
Lack of Fit	3	15.19	15.19	5.06	2.45	0.138
Pure Error	8	16.50	16.50	2.06		
Total	15	5681.94				

f)

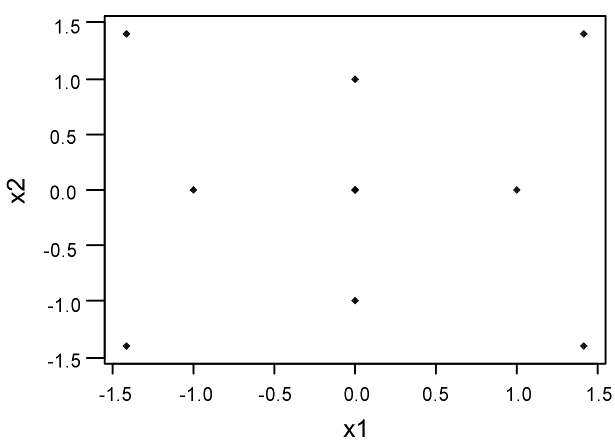




g) The main factors A, B, C, and the interaction AB are all significant.

Section 13-9

13-41. a)

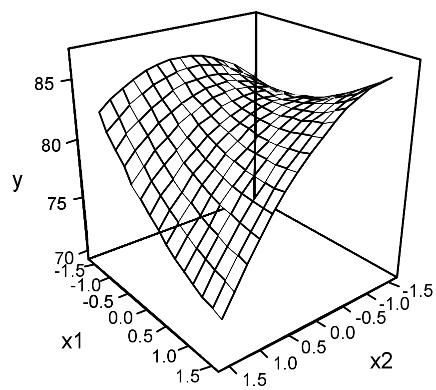


b)

```
Estimated Regression Coefficients for y
Term             Coef       StDev          T      P
Constant       82.024      0.5622    145.905  0.000
x1             -1.115      0.4397     -2.536  0.044
x2             -2.408      0.4397     -5.475  0.002
x1*x1           0.861      0.7343      1.172  0.286
x2*x2          -1.590      0.7342     -2.165  0.074
x1*x2          -1.801      0.3477     -5.178  0.002
S = 1.390       R-Sq = 92.0%     R-Sq(adj) = 85.3%

Analysis of Variance for y
Source          DF    Seq SS    Adj SS    Adj MS     F       P
Regression       5   132.837   132.837   26.5674  13.74  0.003
  Linear         2    70.393    70.391   35.1957  18.21  0.003
  Square         2    10.602    10.610    5.3048   2.74  0.142
  Interaction    1    51.842    51.842   51.8425  26.82  0.002
Residual Error   6    11.600    11.600    1.9333
  Lack-of-Fit    3    10.052    10.052    3.3507   6.50  0.079
  Pure Error     3     1.548     1.548    0.5158
Total           11   144.437
```

The second order model appears to be significant for the interaction term (p = 0.002). However, the square terms are not significant (p = 0.142).

c)



There appears to be a saddle point in the experimental region. The yield increases as x_1 is decreased and x_2 is near the zero level.

13-42. Move 1.5 units in the direction of x_1 for every –0.8 unit in the direction of x_2. Thus, the path of steepest ascent passes through the point (0,0) and has a slope -0.8/1.5 = -0.533.

13-43. a) $10 + 5x_1 + 2x_2 > 12$ $\quad$ $23 + 3x_1 + 4x_2 > 27.50$

$x_2 > -\frac{5}{2}x_1 + 1$ $\quad$ $x_2 < -0.75x_1 + 1.125$

The feasible region is between thses two lines, which can be shown graphically on the x_1-x_2 plane.

b) Operating the process with $x_1 = 1.5$ and $x_2 = -1.5$ results in y_1 and y_2 comfortably within the feasible region.

13-44. a) A central composite design has been used but it is not rotatable.

b)

Term	Coef	StDev	T	P
Constant	150.04	7.821	19.184	0.000
x1	-58.47	5.384	-10.861	0.000
x2	3.35	5.384	0.623	0.556
x1*x1	-6.53	5.693	-1.147	0.295
x2*x2	10.58	5.693	1.859	0.112
x1*x2	0.50	7.848	0.064	0.951

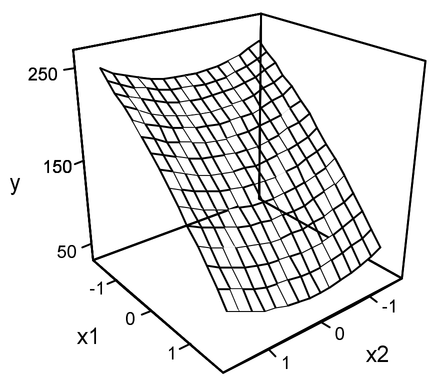
S = 15.70 R-Sq = 95.4% R-Sq(adj) = 91.6%

Analysis of Variance for y

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Regression	5	30688.7	30688.7	6137.7	24.91	0.001
Linear	2	29155.4	29155.4	14577.7	59.17	0.000
Square	2	1532.3	1532.3	766.1	3.11	0.118
Interaction	1	1.0	1.0	1.0	0.00	0.951
Residual Error	6	1478.2	1478.2	246.4		
Lack-of-Fit	3	4.2	4.2	1.4	0.00	1.000
Pure Error	3	1474.0	1474.0	491.3		
Total	11	32166.9				

The linear terms appear to be significant ($p = 0.001$) while both the square terms and interaction terms are insignificant ($p = 0.118$ and $p = 0.951$, respectively).

Since x_1 is the only significant factor, to minimize ash increase the value of x_1.



13-45. a)

Response Surface Regression

Term	Coef	StDev	T	P
Constant	311.190	31.11	10.002	0.000
x1	158.983	14.35	11.083	0.000
x2	138.830	14.94	9.293	0.000
x3	160.870	14.94	10.768	0.000
x1*x1	27.252	26.21	1.040	0.314
x2*x2	7.020	25.19	0.279	0.784
x3*x3	0.346	25.19	0.014	0.989
x1*x2	39.001	17.57	2.220	0.041
x1*x3	48.443	17.57	2.757	0.014
x2*x3	87.690	18.65	4.702	0.000

S = 60.86 R-Sq = 95.5% R-Sq(adj) = 93.0%

Analysis of Variance for y1

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Regression	9	1267693	1267693	140855	38.03	0.000
Linear	3	1130334	1150835	383612	103.56	0.000
Square	3	9052	4431	1477	0.40	0.756
Interaction	3	128307	128307	42769	11.55	0.000
Residual Error	16	59267	59267	3704		
Total	25	1326960				

The quadratic model for y_1 is

$$y_1 = 331.190 + 158.983x_1 + 138.830x_2 + 160.870x_3 + 39.001x_1x_2 + 48.443x_1x_3 + 87.69x_2x_3$$

b) **Response Surface Regression**

Term	Coef	StDev	T	P
Constant	49.604	8.148	6.088	0.000
x1	6.893	9.769	0.706	0.488
x2	17.736	10.083	1.759	0.092
x3	31.605	10.083	3.135	0.005

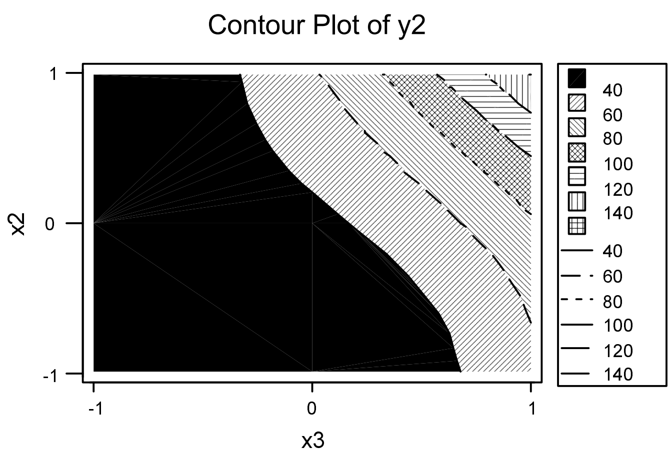
S = 41.45   R-Sq = 36.8%   R-Sq(adj) = 28.1%

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Regression	3	21968.9	21968.9	7322.96	4.26	0.016
Linear	3	21968.9	21968.9	7322.96	4.26	0.016
Residual Error	22	37793.3	37793.3	1717.88		
Total	25	59762.2				

The linear model for y_2 is given by

$$y_2 = 49.604 + 17.736x_2 + 31.605x_3$$

c)



The darkened section represents the smallest standard deviation for the combination of x_2 and x_3. If the level of 0 is chosen for both x_2 and x_3, then to attain the desired level of 500 for the mean, solve the quadratic equation found in part a using y = 500, $x_2 = 0$ and $x_3 = 0$ for x_1. The level of x_1 for this particular situation is approximately 1.

Supplemental Exercises

13-46. a)

Estimated Effects and Coefficients for viscosit

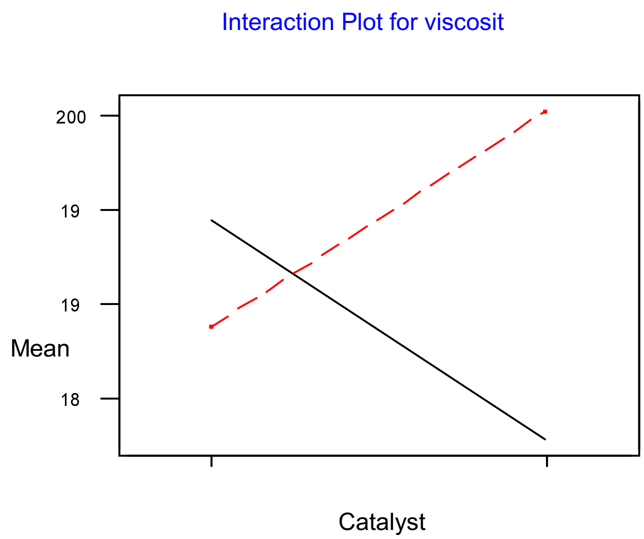
Term	Effect	Coef	StDev Coef	T	P
Constant		191.563	1.158	165.49	0.000
pH	5.875	2.937	1.158	2.54	0.026
Catalyst	-0.125	-0.062	1.158	-0.05	0.958
pH*Catalyst	11.625	5.812	1.158	5.02	0.000

Analysis of Variance for viscosit

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Main Effects	2	138.125	138.125	69.06	3.22	0.076
2-Way Interactions	1	540.563	540.563	540.56	25.22	0.000
Residual Error	12	257.250	257.250	21.44		
Pure Error	12	257.250	257.250	21.44		
Total	15	935.938				

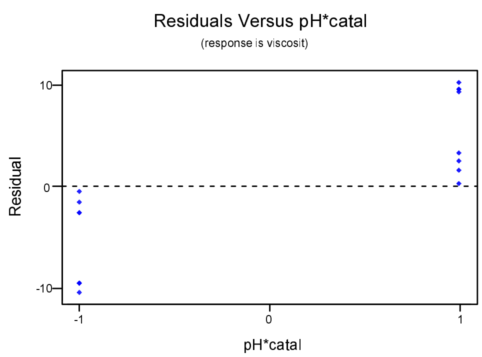
The main effect of pH and the interaction of pH and Catalyst are significant at the 0.05 level of significance.

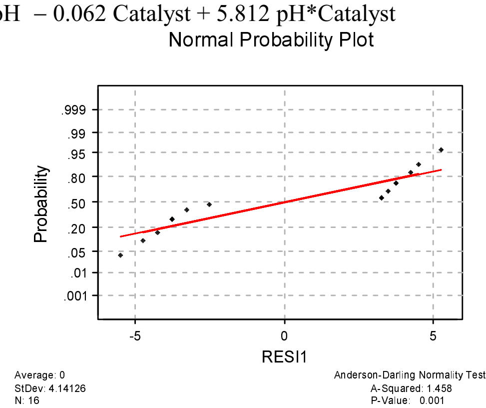
b)



The interaction plot supports the information found in the analysis of part a. There appears to be a significant interaction between pH and Catalyst.

c) The model used is viscosity = 191.563 + 2.937pH – 0.062 Catalyst + 5.812 pH*Catalyst





The residual plots appear to be adequate. There is a significant difference between the two levels of the interaction pH*Catalyst.

13-47. a)

Estimated Effects and Coefficients for flatness

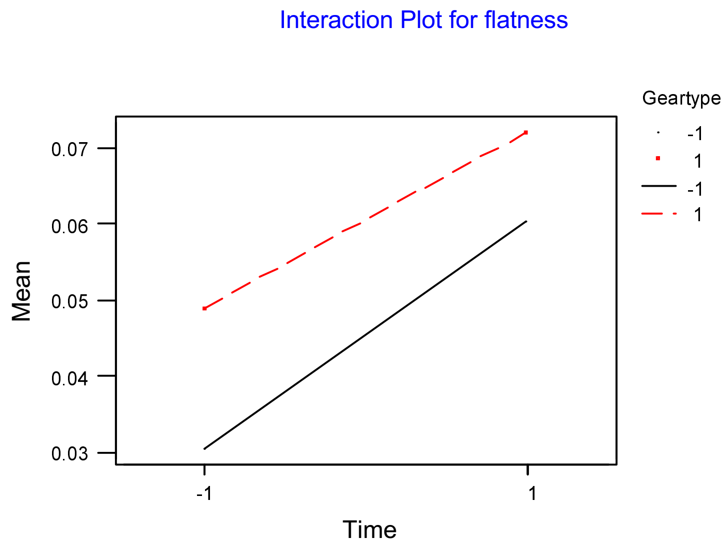
Term	Effect	Coef	StDev Coef	T	P
Constant		0.053000	0.003637	14.57	0.000
Geartype	0.015250	0.007625	0.003637	2.10	0.104
Time	0.026750	0.013375	0.003637	3.68	0.021
Geartype*Time	-0.003500	-0.001750	0.003637	-0.48	0.656

Analysis of Variance for flatness

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Main Effects	2	0.00189625	0.00189625	0.00094812	8.96	0.033
2-Way Interactions	1	0.00002450	0.00002450	0.00002450	0.23	0.656
Residual Error	4	0.00042325	0.00042325	0.00010581		
Pure Error	4	0.00042325	0.00042325	0.00010581		
Total	7	0.00234400				

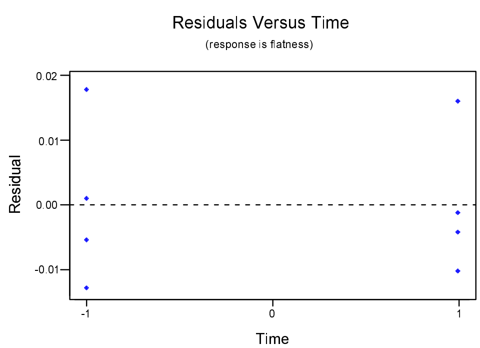
There is no evidence that flatness distortion is different for the different gear types ($p = 0.104$). Heat treating time affects the flatness distortion ($p = 0.021$). The factors do not interact ($p = 0.656$).

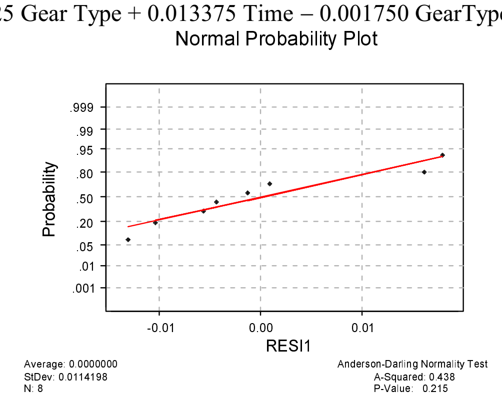
b)



The interaction plot for the effects indicates that there is no interaction between gear type and time. There is also no significant difference between the low and high levels of gear type. The interaction plot does indicate however, there may be some significant difference between the low and high levels of time.

c) The model used is flatness = 0.053 + 0.007625 Gear Type + 0.013375 Time – 0.001750 GearType*Time





The residual plots are adequate. There does not appear to be any serious departure from normality or violation of the assumption of constant variance.

13-48. a)

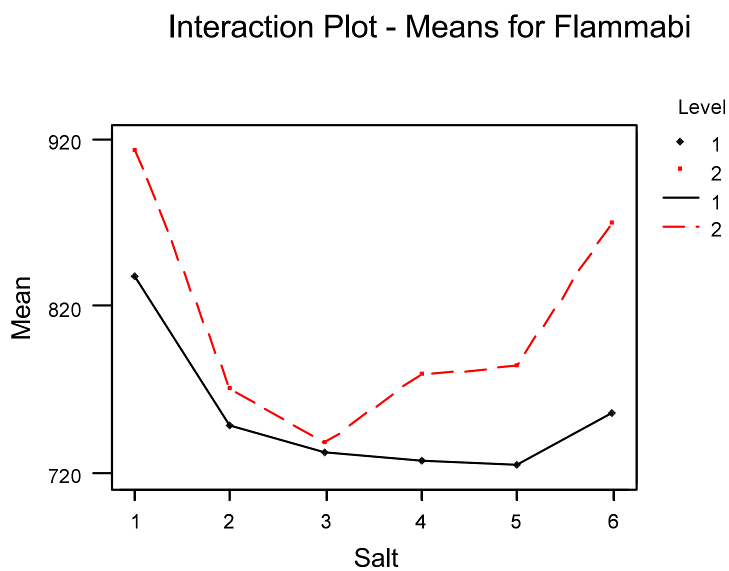
Factor	Type	Levels	Values					
Level	fixed	2	1	2				
Salt	fixed	6	1	2	3	4	5	6

Analysis of Variance for Flammabi

Source	DF	SS	MS	F	P
Level	1	27390	27390	63.24	0.000
Salt	5	86087	17217	39.75	0.000
Level*Salt	5	11459	2292	5.29	0.002
Error	24	10395	433		
Total	35	135332			

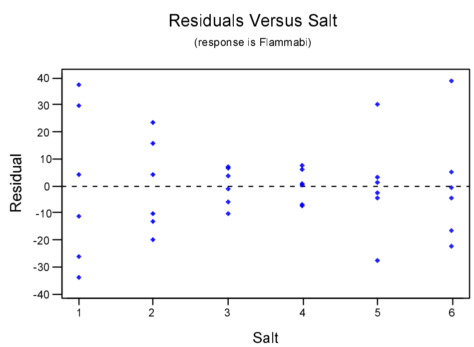
There is a significant difference between the application levels, the salts, and there is a significant difference between the levels of the interaction of the two.

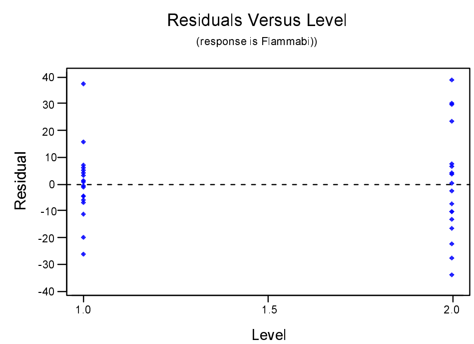
b)

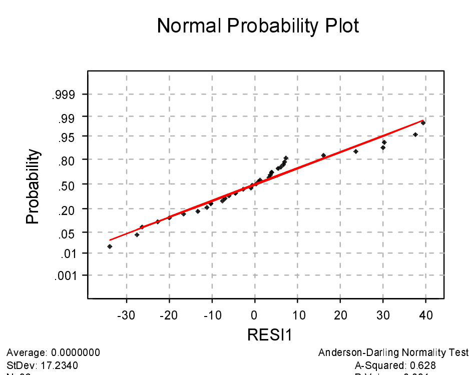


From the interaction plot, we see that the untreated salt has a higher flammability average than any of the other five levels. The remaining five levels ($MgCl_2$, $NaCl$, $CaCO_3$, $CaCl_2$, Na_2CO_3) appear to have similar flammability averages. Overall, application level 1 increases the flammability average.

c)







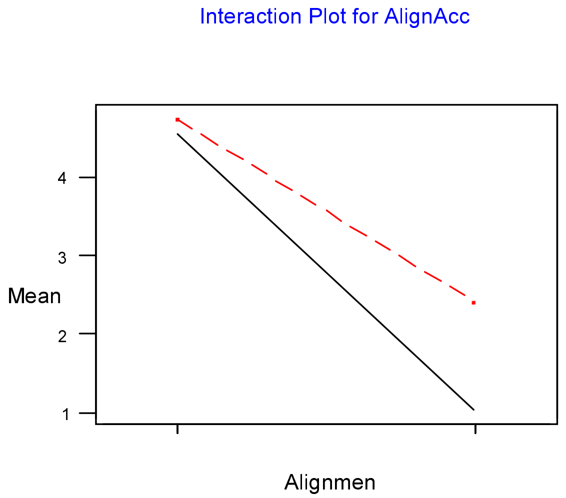
The residual plots indicate a significant difference among the salt types, while there does not appear to be a significant difference between residuals for the two application levels. There is no apparent violation of the normality assumption.

13-49. a)

```
Estimated Effects and Coefficients for AlignAcc
Term                  Effect      Coef  StDev Coef       T      P
Constant                         3.187     0.05265   60.54  0.000
SolderSi               0.785     0.393     0.05265    7.46  0.002
Alignmen              -2.940    -1.470     0.05265  -27.92  0.000
SolderSi*Alignmen      0.600     0.300     0.05265    5.70  0.005

Analysis of Variance for AlignAcc
Source                DF      Seq SS     Adj SS     Adj MS      F      P
Main Effects           2     18.5197    18.5197    9.25983 417.58  0.000
2-Way Interactions     1      0.7200     0.7200    0.72000  32.47  0.005
Residual Error         4      0.0887     0.0887    0.02218
  Pure Error           4      0.0887     0.0887    0.02217
Total                  7     19.3283
```

The analysis indicates that both solder size and alignment method significantly affect alignment accuracy. The interaction between solder size and alignment method is also significant in affecting alignment accuracy

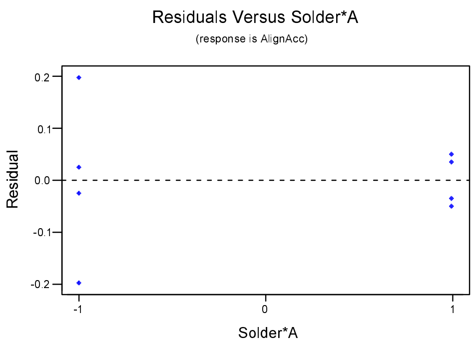
b)

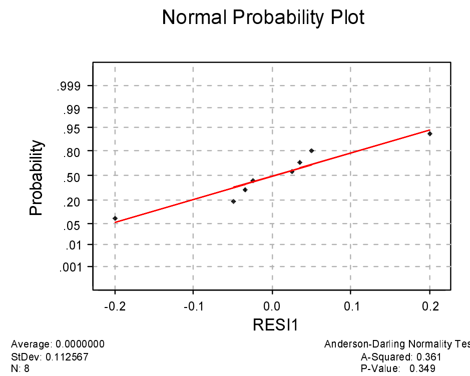



```
The regression equation is
AlignAcc = 3.19 + 0.392 SolderSize - 1.47 Alignment + 0.300 Solder*Align
Predictor        Coef       StDev          T        P
Constant      3.18750     0.05265      60.54    0.000
SolderSi      0.39250     0.05265       7.46    0.002
Alignmen     -1.47000     0.05265     -27.92    0.000
Solder*A      0.30000     0.05265       5.70    0.005
S = 0.1489      R-Sq = 99.5%     R-Sq(adj) = 99.2%

Analysis of Variance
Source        DF          SS          MS          F         P
Regression     3     19.2397      6.4132     289.21     0.000
Error          4      0.0887      0.0222
Total          7     19.3283

Source        DF      Seq SS
SolderSi       1      1.2324
Alignmen       1     17.2872
Solder*A       1      0.7200
```

Since the smaller value is preferred, then, to improve alignment accuracy, it seems that solder size should be set at its small level (75µm) while the alignment method used should be method 3.

c)





The normal probability plot does not suggest a departure from normality. The assumption of constant variance may be of concern. It appears that the variability is lower for the high level of the interaction.

13-50. a)

Estimated Effects

Term	Effect
A	-2.74
B	-6.66
C	3.49
A*B	-8.71
A*C	7.04
B*C	11.46
A*B*C	-6.49

b)

Estimated Effects and Coefficients for deltalin

Term	Effect	Coef	StDev Coef	T	P
Constant		-11.82	1.328	-8.90	0.001
A	-2.74	-1.37	1.626	-0.84	0.447
B	-6.66	-3.33	1.626	-2.05	0.110
C	3.49	1.75	1.626	1.07	0.343
A*B	-8.71	-4.35	1.626	-2.68	0.055
A*C	7.04	3.52	1.626	2.17	0.096
B*C	11.46	5.73	1.626	3.52	0.024
A*B*C	-6.49	-3.25	1.626	-2.00	0.117

Analysis of Variance for deltalin

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Main Effects	3	128.082	128.082	42.694	2.02	0.254
2-Way Interactions	3	513.383	513.383	171.128	8.09	0.036
3-Way Interactions	1	84.305	84.305	84.305	3.99	0.117
Residual Error	4	84.608	84.608	21.152		
Curvature	1	59.441	59.441	59.441	7.09	0.076
Pure Error	3	25.167	25.167	8.389		
Total	11	810.379				

The estimate of pure error using the centerpoints is given in Minitab as 8.389. Hand calculations may differ slightly due to significant digits.

c) Using the output given in part b), it appears from the Minitab output that the 2-way interactions are significant ($p = 0.036$), while the main effects and 3-way interaction are not ($p = 0.254$ and 0.117, respectively). Curvature is also insignificant with $p = 0.076$.

d) Using 'Regression' in Minitab the resulting model should be

The regression equation is
deltaline = - 11.8 - 1.37 A - 3.33 B + 1.75 C - 4.35 AB + 3.52 AC + 5.73 BC

Predictor	Coef	StDev	T	P
Constant	-11.822	1.678	-7.05	0.001
A	-1.371	2.055	-0.67	0.534
B	-3.329	2.055	-1.62	0.166
C	1.746	2.055	0.85	0.434
ab	-4.354	2.055	-2.12	0.088
ac	3.521	2.055	1.71	0.147
bc	5.729	2.055	2.79	0.039

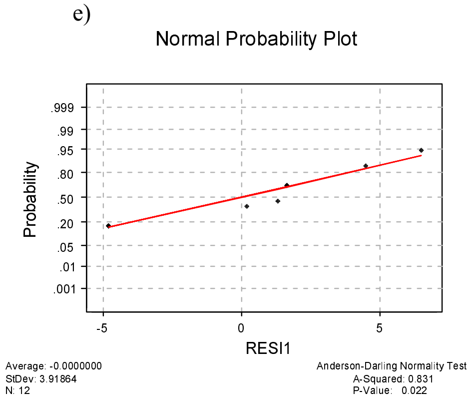
S = 5.812 R-Sq = 79.2% R-Sq(adj) = 54.1%

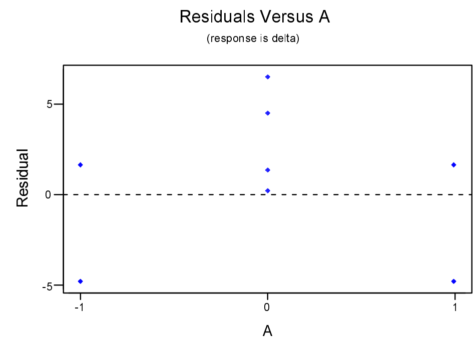
Analysis of Variance

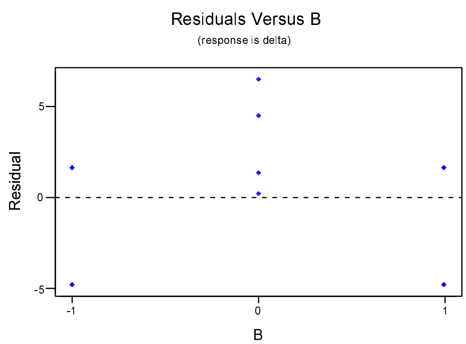
Source	DF	SS	MS	F	P
Regression	6	641.47	106.91	3.16	0.113
Error	5	168.91	33.78		
Total	11	810.38			

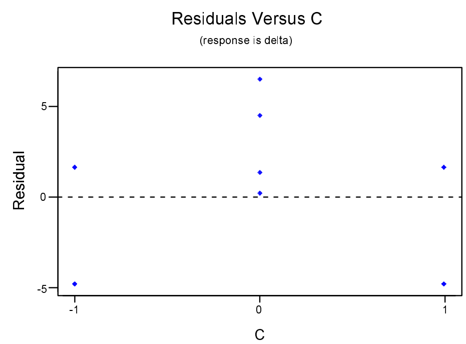
Source	DF	Seq SS
A	1	15.04
B	1	88.64
C	1	24.40
ab	1	151.64
ac	1	99.19
bc	1	262.55

e)









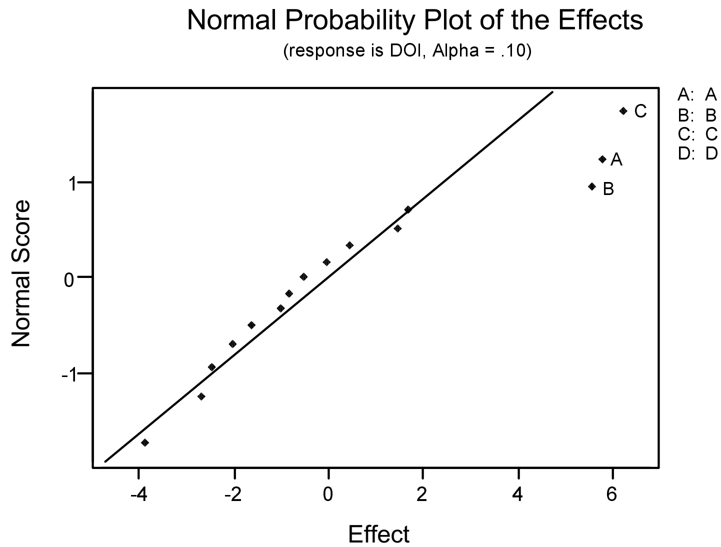
The residual plots appear to be adequate.

13-51. a)

Estimated Effects and Coefficients for DOI

Term	Effect	Coef
Constant		72.894
A	5.763	2.881
B	5.563	2.781
C	6.238	3.119
D	-3.862	-1.931
A*B	-2.013	-1.006
A*C	-0.837	-0.419
A*D	-2.488	-1.244
B*C	-0.537	-0.269
B*D	1.462	0.731
C*D	-0.063	-0.031
A*B*C	0.437	0.219
A*B*D	-1.013	-0.506
A*C*D	1.663	0.831
B*C*D	-2.687	-1.344
A*B*C*D	-1.612	-0.806

b)



Based on the normal probability plot it appears that factors A, B, and C are significant.

c) Run an analysis using the main factors A, B, and C and interactions among these variables to see if any are significant.

Predictor	Coef	StDev	T	P
Constant	72.894	1.073	67.92	0.000
A	2.881	1.073	2.68	0.028
B	2.781	1.073	2.59	0.032
C	3.119	1.073	2.91	0.020
ab	-1.006	1.073	-0.94	0.376
ac	-0.419	1.073	-0.39	0.707
bc	-0.269	1.073	-0.25	0.809
abc	0.219	1.073	0.20	0.844

S = 4.293 R-Sq = 74.6% R-Sq(adj) = 52.4%

Based on this analysis, only the main factors A, B, and C are significant. Run a regression analysis using these important factors.

d)

The regression equation is
DOI = 72.9 + 2.88 A + 2.78 B + 3.12 C

Predictor	Coef	StDev	T	P
Constant	72.8937	0.9365	77.84	0.000
A	2.8813	0.9365	3.08	0.010
B	2.7813	0.9365	2.97	0.012
C	3.1188	0.9365	3.33	0.006

S = 3.746 R-Sq = 71.0% R-Sq(adj) = 63.7%

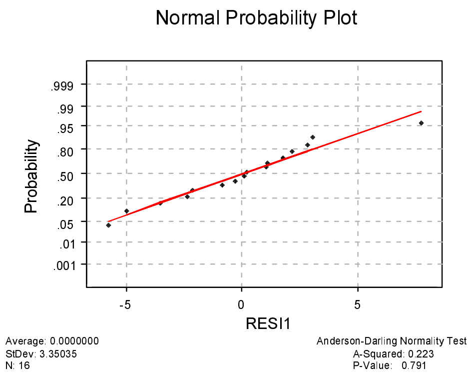
Analysis of Variance

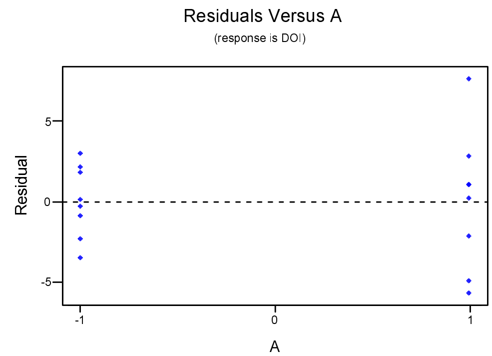
Source	DF	SS	MS	F	P
Regression	3	412.22	137.41	9.79	0.002
Error	12	168.37	14.03		
Total	15	580.59			

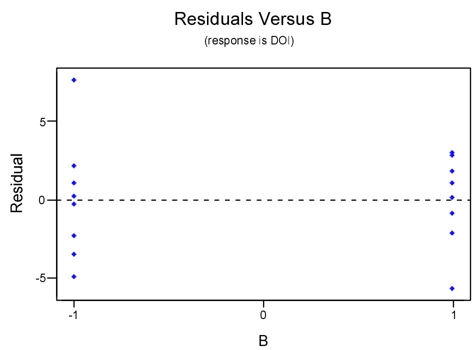
Source	DF	Seq SS
A	1	132.83
B	1	123.77
C	1	155.63

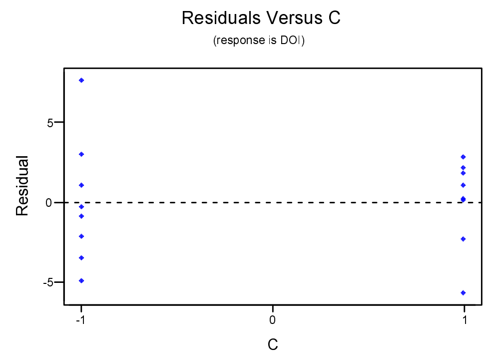
DOI = 72.9 + 2.88 A + 2.78 B + 3.12 C

e)









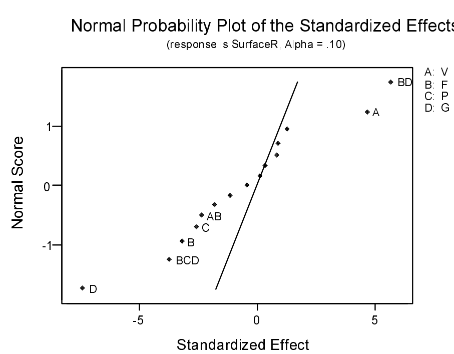
The residual plots appear to be adequate.

13-52. a)

Estimated Effects

Term	Effect
V	15.75
F	-10.75
P	-8.75
G	-25.00
V*F	-8.00
V*P	3.00
V*G	2.75
F*P	-6.00
F*G	19.25
P*G	-3.75
V*F*P	1.25
V*F*G	-1.50
V*P*G	0.50

b)



According to the probability plot, factors V, F, P, and G along with the interactions VF and FPG. are possibly significant.

c) Running a regression analysis using these factors we find

SurfaceRough = 101 + 7.88 V - 5.37 F - 4.38 P - 12.5 G - 6.25 FPG - 4.00VF

Predictor	Coef	StDevCoeff	T	P
Constant	100.522	2.420	41.53	0.000
V	7.875	2.902	2.71	0.015
F	-5.375	2.902	-1.85	0.083
P	-4.375	2.902	-1.51	0.151
G	-12.500	2.902	-4.31	0.001
fpg	-6.250	2.902	-2.15	0.047
vf	-4.000	2.902	-1.38	0.187

S = 11.61   R-Sq = 70.5%   R-Sq(adj) = 59.4%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	6	5141.8	857.0	6.36	0.001
Error	16	2156.0	134.7		
Total	22	7297.7			

Source	DF	Seq SS
V	1	992.3
F	1	462.2
P	1	306.3
G	1	2500.0
fpg	1	625.0
vf	1	256.0

Based on this initial analysis we see that P, F, and the interaction VF are insignificant at the 0.05 level of significance.

Thus, the regression analysis and final model are

SurfaceRough = 101 + 7.88 V - 5.37 F - 4.38 P - 12.5 G - 6.25 fpg

Predictor	Coef	StDevCoef	T	P
Constant	100.522	2.484	40.47	0.000
V	7.875	2.978	2.64	0.017
F	-5.375	2.978	-1.80	0.089
P	-4.375	2.978	-1.47	0.160
G	-12.500	2.978	-4.20	0.001
fpg	-6.250	2.978	-2.10	0.051

S = 11.91   R-Sq = 66.9%   R-Sq(adj) = 57.2%

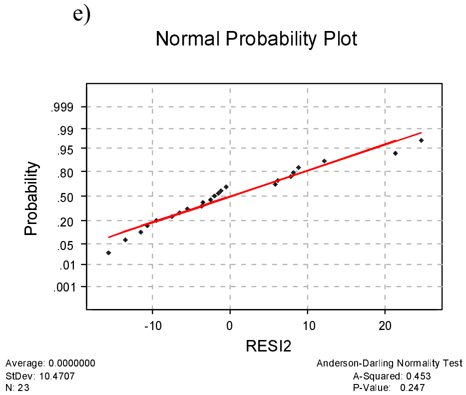
Analysis of Variance

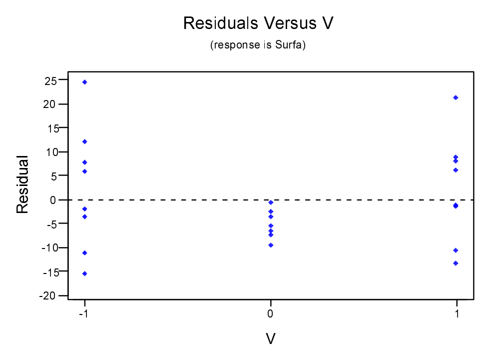
Source	DF	SS	MS	F	P
Regression	5	4885.8	977.2	6.89	0.001
Error	17	2412.0	141.9		
Total	22	7297.7			

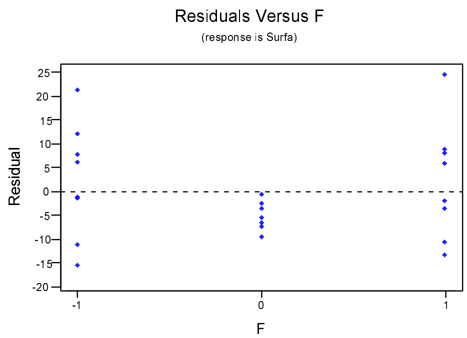
Source	DF	Seq SS
V	1	992.3
F	1	462.2
P	1	306.3
G	1	2500.0
fpg	1	625.0

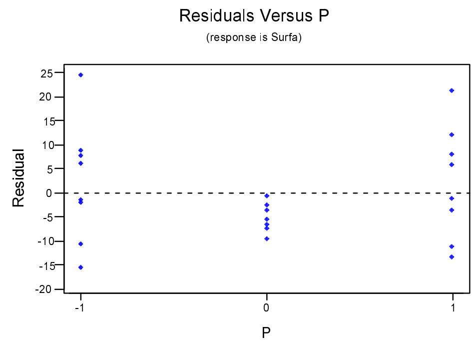
d) From the analysis, we see that jet traverse speed and abrasive grain size are significant along with the interaction of abrasive flow rate, abrasive grain size, and waterjet pressure. Since we have adopted a hierarchical modeling approach, flow rate and waterjet pressure are also included in the model.

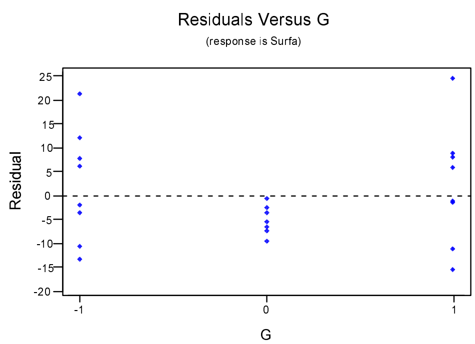
e)











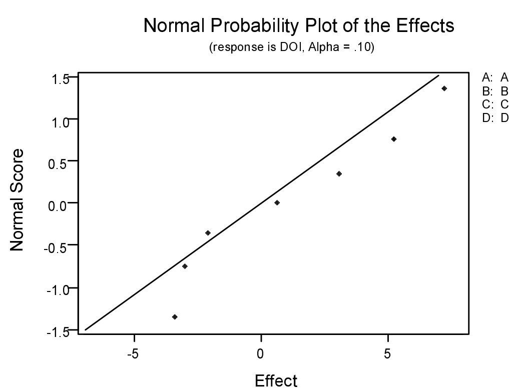
The residual plots appear to indicate the assumption of constant variance may not be met.

13-53. One possible design is

A	B	C	D	DOI
-1	-1	-1	-1	63.8
1	-1	-1	1	64.9
-1	1	-1	1	72.7
1	1	-1	-1	76.5
-1	-1	1	1	68.0
1	-1	1	-1	77.2
-1	1	1	-1	77.7
1	1	1	1	75.9

Estimated Effects

Term	Effect
A	3.075
B	7.225
C	5.225
D	-3.425
AB	-2.075
AC	0.625
AD	-3.025



It may be useful to run analysis with the main effects only to see which main effect is significant.

Term	Effect	Coef	StDev Coef	T	P
Constant		72.088	1.074	67.11	0.000
A	3.075	1.538	1.074	1.43	0.248
B	7.225	3.612	1.074	3.36	0.044
C	5.225	2.612	1.074	2.43	0.093
D	-3.425	-1.712	1.074	-1.59	0.209

```
Analysis of Variance for DOI
Source              DF      Seq SS     Adj SS     Adj MS       F       P
Main Effects         4      201.38     201.38     50.344    5.45  0.097
Residual Error       3       27.69      27.69      9.231
Total                7      229.07
```

Based on this analysis, only factor B appears to be significant at the 0.05 level of significance. Comparing this result to that of problem 7-38, we see that the Factors A and C were not revealed as significant in the smaller design.

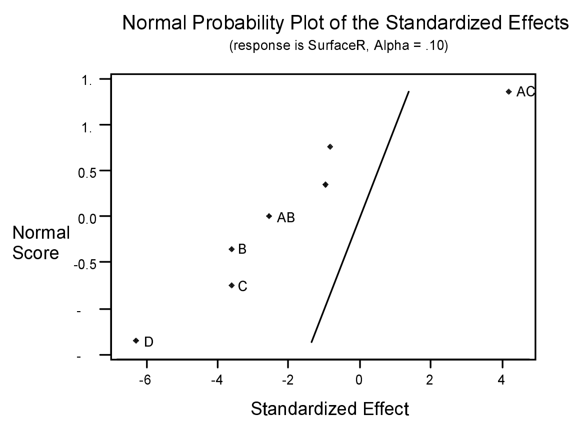
13-54. One possible design is

A	B	C	D	SurfaceRough
-1	-1	-1	-1	143
1	-1	-1	1	98
-1	1	-1	1	110
1	1	-1	-1	103
-1	-1	1	1	76
1	-1	1	-1	137
-1	1	1	-1	98
1	1	1	1	70
0	0	0	0	95
0	0	0	0	98
0	0	0	0	100
0	0	0	0	97
0	0	0	0	94
0	0	0	0	93
0	0	0	0	91

Where A = V, B = F, C = P, and D = G

Estimated Effects

Term	Effect
A	-4.75
B	-18.25
C	-18.25
D	-31.75
A*B	-12.75
A*C	21.25
A*D	-4.25



From the estimated effects, we could tentatively identify F, P, G, and the interactions VF and VP as significant.

The regression analysis gives

```
SurfaceRough = 100 - 2.37V - 9.13F - 9.13P - 15.9G - 6.37VF + 10.6VP- 2.13VG
Predictor       Coef       StDev          T        P
Constant     100.200       1.843      54.38    0.000
V             -2.375       2.523      -0.94    0.378
F             -9.125       2.523      -3.62    0.009
P             -9.125       2.523      -3.62    0.009
G            -15.875       2.523      -6.29    0.000
VF            -6.375       2.523      -2.53    0.039
VP            10.625       2.523       4.21    0.004
VG            -2.125       2.523      -0.84    0.428
```

From the analysis we see that factors F, P, G, VF, and VP are significant. In the analysis of 7-39, we found only V, G, and FPG significant.

13-55.

```
Fractional Factorial Design
Factors:   8    Base Design:      8, 16     Resolution: IV
Runs:     16    Replicates:          1      Fraction: 1/16
Design Generators:  E = BCD  F = ACD  G = ABC  H = ABD
Defining Relation:  I = BCDE = ACDF = ABCG = ABDH = ABEF = ADEG = ACEH = BDFG =
BCFH = CDGH = CEFG = DEFH = BEGH = AFGH = ABCDEFGH
Alias Structure (up to order 4)
I + ABCG + ABDH + ABEF + ACDF + ACEH + ADEG + AFGH + BCDE + BCFH + BDFG + BEGH +
CDGH + CEFG + DEFH

A + BCG + BDH + BEF + CDF + CEH + DEG + FGH
B + ACG + ADH + AEF + CDE + CFH + DFG + EGH
C + ABG + ADF + AEH + BDE + BFH + DGH + EFG
D + ABH + ACF + AEG + BCE + BFG + CGH + EFH
E + ABF + ACH + ADG + BCD + BGH + CFG + DFH
F + ABE + ACD + AGH + BCH + BDG + CEG + DEH
G + ABC + ADE + AFH + BDF + BEH + CDH + CEF
H + ABD + ACE + AFG + BCF + BEG + CDG + DEF
AB + CG + DH + EF + ACDE + ACFH + ADFG + AEGH + BCDF + BCEH + BDEG + BFGH
AC + BG + DF + EH + ABDE + ABFH + ADGH + AEFG + BCDH + BCEF + CDEG + CFGH
AD + BH + CF + EG + ABCE + ABFG + ACGH + AEFH + BCDG + BDEF + CDEH + DFGH
AE + BF + CH + DG + ABCD + ABGH + ACFG + ADFH + BCEG + BDEH + CDEF + EFGH
AF + BE + CD + GH + ABCH + ABDG + ACEG + ADEH + BCFG + BDFH + CEFH + DEFG
AG + BC + DE + FH + ABDF + ABEH + ACDH + ACEF + BDGH + BEFG + CDFG + CEGH
AH + BD + CE + FG + ABCF + ABEG + ACDG + ADEF + BCGH + BEFH + CDFH + DEGH
```

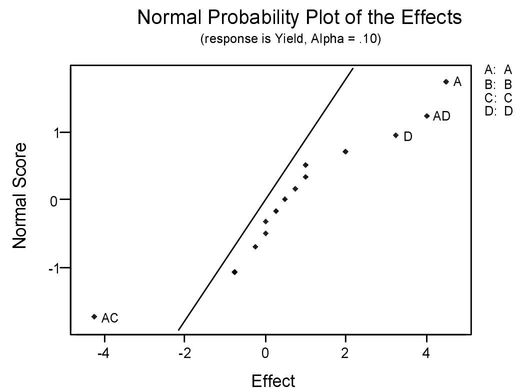
Run	A	B	C	D	E	F	G	H
1	−	−	−	−	−	−	−	−
2	+	−	−	−	−	+	+	+
3	−	+	−	−	+	−	+	+
4	+	+	−	−	+	+	−	−
5	−	−	+	−	+	+	+	−
6	+	−	+	−	+	−	−	+
7	−	+	+	−	−	+	−	+
8	+	+	+	−	−	−	+	−
9	−	−	−	+	+	+	−	+
10	+	−	−	+	+	−	+	−
11	−	+	−	+	−	+	+	−
12	+	+	−	+	−	−	−	+
13	−	−	+	+	−	−	+	+
14	+	−	+	+	−	+	−	−
15	−	+	+	+	+	−	−	−
16	+	+	+	+	+	+	+	+

13-56.

```
Fractional Factorial Design
Factors:   5    Base Design:       5, 8     Resolution: III
Runs:      8    Replicates:           1     Fraction: 1/4
Blocks: none    Center pts (total):   0
*** NOTE *** Some main effects are confounded with two-way interactions
Design Generators:  D = AB  E = AC
Defining Relation:  I = ABD = ACE = BCDE
Alias Structure     I + ABD + ACE + BCDE
A + BD + CE + ABCDE
B + AD + CDE + ABCE
C + AE + BDE + ABCD
D + AB + BCE + ACDE
E + AC + BCD + ABDE
BC + DE + ABE + ACD
BE + CD + ABC + ADE
```

Run	A	B	C	D	E
1	−	−	−	+	+
2	+	−	−	−	−
3	−	+	−	−	+
4	+	+	−	+	−
5	−	−	+	+	−
6	+	−	+	−	+
7	−	+	+	−	−
8	+	+	+	+	+

13-57. a)



The factors that appear to have large effects are A, D, AC, and AD.

b) The regression analysis is

Estimated Effects and Coefficients for Yield

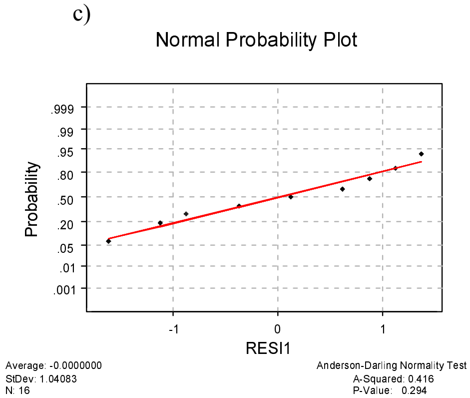
Term	Effect	Coef	StDev Coef	T	P
Constant		17.375	0.3187	54.52	0.000
A	4.500	2.250	0.3187	7.06	0.000
C	2.000	1.000	0.3187	3.14	0.011
D	3.250	1.625	0.3187	5.10	0.000
A*C	-4.250	-2.125	0.3187	-6.67	0.000
A*D	4.000	2.000	0.3187	6.28	0.000

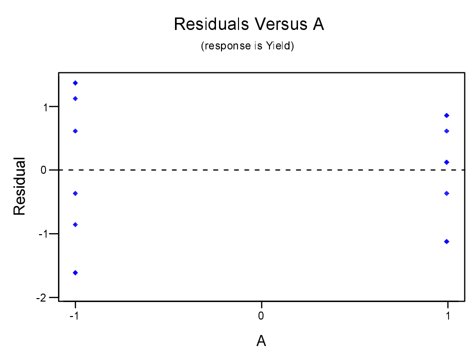
Analysis of Variance for Yield

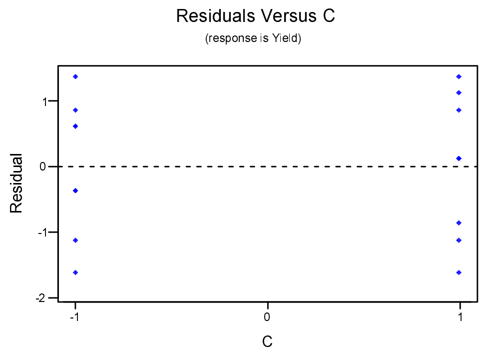
Source	DF	Seq SS	Adj SS	Adj MS	F	P
Main Effects	3	139.250	139.250	46.4167	28.56	0.000
2-Way Interactions	2	136.250	136.250	68.1250	41.92	0.000
Residual Error	10	16.250	16.250	1.6250		
Lack of Fit	2	0.250	0.250	0.1250	0.06	0.940
Pure Error	8	16.000	16.000	2.0000		
Total	15	291.750				

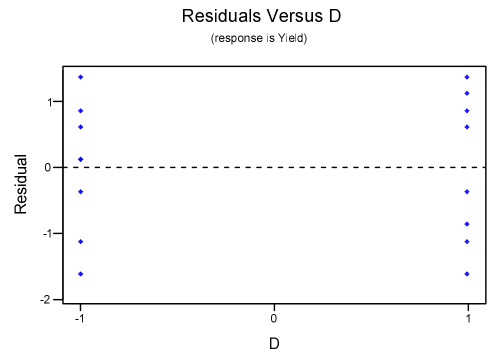
The conclusions are that factors A, C, and D are significant along with the interactions AC and AD.

c)



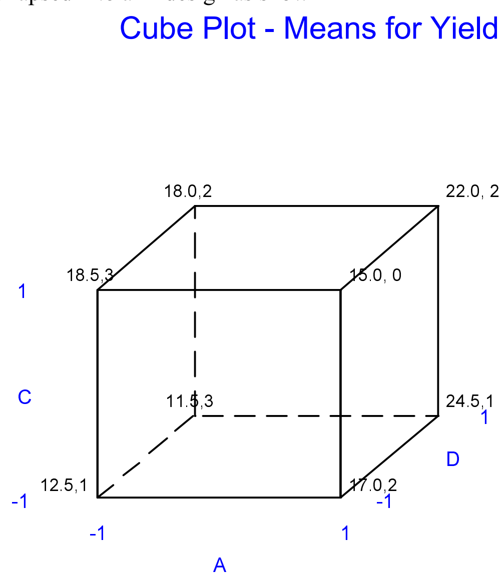






The residual plots appear to be adequate.

d) The design can be collapsed into a 2^3 design as shown



The data are Mean, Range for each location. It is evident that the mean yield is largest for high levels of A and D and low level of C.

13-58. If we consider the main effects A, C, and D as significant, we may want to investigate the interactions containing these factors.

Term	Effect	Coef	StDev Coef	T	P
A	4.500	2.250	0.3536	6.36	0.000
C	2.000	1.000	0.3536	2.83	0.022
D	3.250	1.625	0.3536	4.60	0.000
A*C	-4.250	-2.125	0.3536	-6.01	0.000
A*D	4.000	2.000	0.3536	5.66	0.000

Using the relationship 2(st.dev coef) = st. err. we have

2(0.3536) = 0.7072

For a 95% confidence interval, $t_{8,0.025} = 2.307$

Therefore the 95% confidence intervals for the effects are given by: effect $\pm$ 2.307(0.072)

or effect $\pm$ 1.6315.

13-59. The generator for this fraction was I=ABCD

Totally confounded terms were removed from the analysis

```
I + a*b*c*d
a + b*c*d
b + a*c*d
c + a*b*d
d + a*b*c
e + a*b*c*d*e
a*b + c*d
a*c + b*d
a*d + b*c
a*e + b*c*d*e
b*e + a*c*d*e
c*e + a*b*d*e
d*e + a*b*c*e
a*b*e + c*d*e
a*c*e + b*d*e
a*d*e + b*c*e
```

b) Estimated Effects and Coefficients for freeheig

Term	Effect	Coef	StDev Coef	T	P
Constant		7.6400	0.01901	401.97	0.000
a	0.2133	0.1067	0.01901	5.61	0.000
b	-0.1925	-0.0963	0.01901	-5.06	0.000
c	-0.0783	-0.0392	0.01901	-2.06	0.048
d	0.0625	0.0313	0.01901	1.64	0.110
e	-0.2100	-0.1050	0.01901	-5.52	0.000
a*b	-0.0008	-0.0004	0.01901	-0.02	0.983
a*c	0.0300	0.0150	0.01901	0.79	0.436
a*d	0.0058	0.0029	0.01901	0.15	0.879
a*e	0.0350	0.0175	0.01901	0.92	0.364
b*e	0.1242	0.0621	0.01901	3.27	0.003
c*e	-0.0617	-0.0308	0.01901	-1.62	0.115
d*e	0.0108	0.0054	0.01901	0.28	0.777
a*b*e	0.0308	0.0154	0.01901	0.81	0.423
a*c*e	0.0483	0.0242	0.01901	1.27	0.213
a*d*e	-0.0308	-0.0154	0.01901	-0.81	0.423

Analysis of Variance for freeheig

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Main Effects	5	1.64052	1.64052	0.32810	18.92	0.000
2-Way Interactions	7	0.25797	0.25797	0.03685	2.13	0.069
3-Way Interactions	3	0.05085	0.05085	0.01695	0.98	0.416
Residual Error	32	0.55487	0.55487	0.01734		
Pure Error	32	0.55487	0.55487	0.01734		
Total	47	2.50420				

Based on the analysis, factors A, B, C, and E are significant. The interaction of BE is also significant.

c)

a	b	c	d	e	Range
-1	-1	-1	-1	-1	0.03
1	-1	-1	1	-1	0.30
-1	1	-1	1	-1	0.06
1	1	-1	-1	-1	0.19
-1	-1	1	1	-1	0.46
1	-1	1	-1	-1	0.40
-1	1	1	-1	-1	0.12
1	1	1	1	-1	0.25
-1	-1	-1	-1	1	0.06
1	-1	-1	1	1	0.44
-1	1	-1	1	1	0.06
1	1	-1	-1	1	0.19
-1	-1	1	1	1	0.12
1	-1	1	-1	1	0.13
-1	1	1	-1	1	0.07
1	1	1	1	1	0.31

Estimated Effects and Coefficients for Range

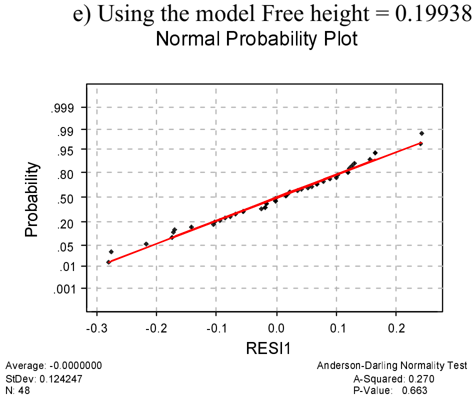
Term	Effect	Coef	StDev Coef	T	P
Constant		0.19938	0.02714	7.35	0.000
a	0.15375	0.07688	0.02714	2.83	0.018
b	-0.08625	-0.04313	0.02714	-1.59	0.143
c	0.06625	0.03312	0.02714	1.22	0.250
d	0.10125	0.05062	0.02714	1.87	0.092
e	-0.05375	-0.02687	0.02714	-0.99	0.345

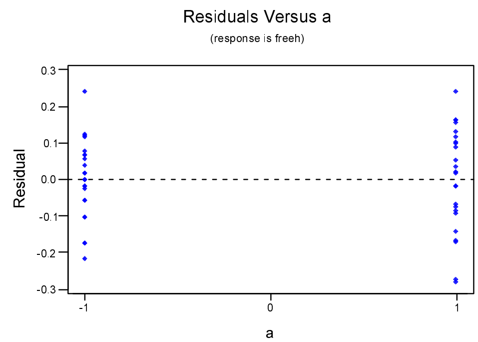
Analysis of Variance for Range

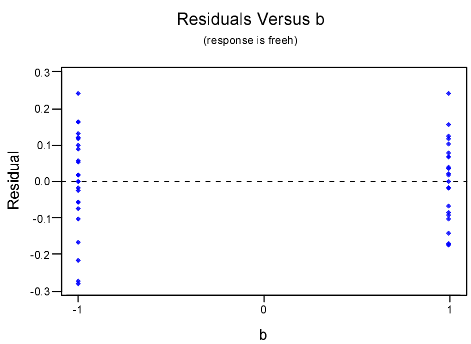
Source	DF	Seq SS	Adj SS	Adj MS	F	P
Main Effects	5	0.1944	0.1944	0.03889	3.30	0.051
Residual Error	10	0.1179	0.1179	0.01179		
Total	15	0.3123				

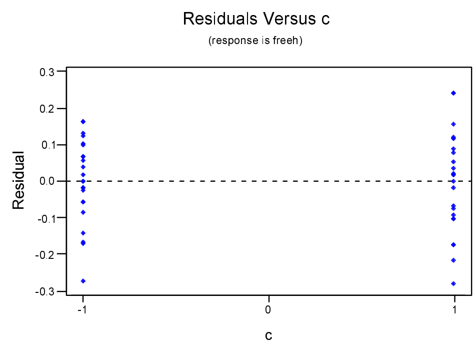
From the analysis we see that factor A is significant in affecting variability in free height.

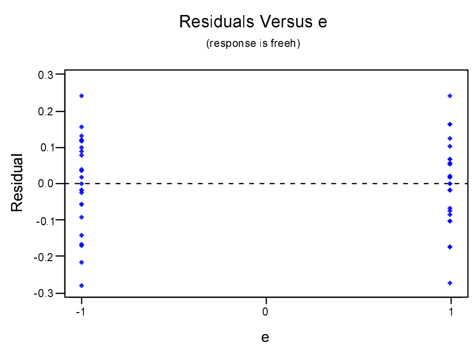
e) Using the model Free height = 0.19938 + 0.07688 A











The residual plots appear to be adequate.

13-60.

The analysis was done using coded units.

Estimated Regression Coefficients for y

Term	Coef	StDev	T	P
Constant	13.7280	0.04201	326.790	0.000
A	0.2966	0.03321	8.931	0.000
B	-0.4052	0.03321	-12.202	0.000
A*A	-0.1240	0.03561	-3.482	0.010
B*B	-0.0790	0.03561	-2.218	0.062
A*B	0.0550	0.04697	1.171	0.280

S = 0.09393   R-Sq = 97.2%   R-Sq(adj) = 95.2%

Analysis of Variance for y

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Regression	5	2.16451	2.16451	0.43290	49.06	0.000
Linear	2	2.01751	2.01751	1.00876	114.32	0.000
Square	2	0.13490	0.13490	0.06745	7.64	0.017
Interaction	1	0.01210	0.01210	0.01210	1.37	0.280
Residual Error	7	0.06177	0.06177	0.00882		
Lack-of-Fit	3	0.02949	0.02949	0.00983	1.22	0.412
Pure Error	4	0.03228	0.03228	0.00807		
Total	12	2.22628				

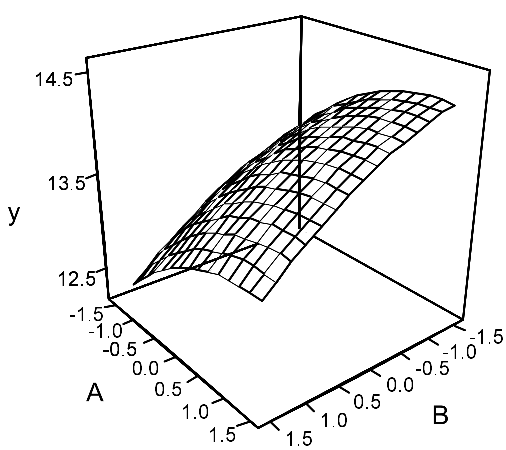
From the analysis, we see that the main effects and the square terms are significant, but the interactions are not. Running an analysis using only the significant values we have

Term	Coef	StDev	T	P
Constant	13.7280	0.04297	319.460	0.000
A	0.2966	0.03397	8.731	0.000
B	-0.4052	0.03397	-11.928	0.000
A*A	-0.1240	0.03643	-3.404	0.009
B*B	-0.0790	0.03643	-2.168	0.062

S = 0.09609   R-Sq = 96.7%   R-Sq(adj) = 95.0%

Analysis of Variance for y

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Regression	4	2.15241	2.15241	0.53810	58.28	0.000
Linear	2	2.01751	2.01751	1.00876	109.25	0.000
Square	2	0.13490	0.13490	0.06745	7.31	0.016
Residual Error	8	0.07387	0.07387	0.00923		
Lack-of-Fit	4	0.04159	0.04159	0.01040	1.29	0.406
Pure Error	4	0.03228	0.03228	0.00807		
Total	12	2.22628				



Maximum viscosity is found by computing the derivatives with respect to A and B of the model. Then setting these equations equal to 0 and solving:

Model: $13.728 + 0.2966\,A - 0.4052\,B - 0.1240\,A^2 - 0.079\,B^2 = y$

$$\frac{\partial y}{\partial A} = 0.2966 - 0.248A = 0$$

$A = 1.19$

$$\frac{\partial y}{\partial B} = -0.4052 - 0.158B = 0$$

$B = -2.56$

The maximum viscosity is given by:

$13.728 + 0.2966\,(1.19) - 0.4052\,(-2.56) - 0.1240\,(1.19)^2 - 0.079\,(-2.56)^2 = 14.425$

Mind-Expanding Exercises

13-61. The ABCD interaction is

$$\frac{1}{8}\left[(1)+ab+ac+bc+ad+bd+cd+abcd\right]-\left[a+b+c+d+abc+abd+acd+bcd\right]$$

If ab is missing, then ABCD interaction will be zero when

[550 + ab + 642 + 601 + 749 + 1052 + 1075 + 729] - [669 + 604 + 633 + 1037 + 635 + 868 + 860 + 1063]=0

Therefore, ab + 5398 - 6369 = 0 or ab = 971.

After estimating ab, only the A and AD effects appear significant.

13-62. Two three-factor interactions could be used to generate the blocks such as ABC and ACD. This would confound these effects and ABC • ACD = BD with blocks. Therefore, only one two-factor and no main effects are confounded with blocks.

13-63.

	A	**B**	**AB**	**block**
(1)	-	-	+	1
a	+	-	-	2
b	-	+	-	2
ab	+	+	+	1

The block effect is estimated by $\frac{a+b}{2} - \frac{(1)-ab}{2}$ which is the same as the estimate of the effect of AB.

13-64. a) A different effect can be confounded in each replicate as follows.

Replicate 1
ABC confounded

(1)	a
ab	b
ac	c
bc	abc

Replicate 2
AB confounded

(1)	a
c	b
ab	ac
abc	bc

Replicate 3
BC confounded

(1)	b
a	c
bc	ab
abc	ac

Replicate 4
AC confounded

(1)	a
b	c
ac	ab
abc	bc

b)

Source of Variation	Degrees of freedom
Replicates	3
Blocks with replicates	4
[or ABC (rep. 1) + AB (rep. 2) + BC (rep. 3) + AC (rep. 4)]	
A	1
B	1
C	1
AB (from replicates 1, 3, and 4)	1
AC (from replicates 1, 2, and 3)	1
BC (from replicates 1, 2, and 4)	1
ABC (from replicates 2, 3, and 4)	1
Error (by subtraction)	17
Total	31

In calculating an interaction sum of squares, only data from the replicates in which the interaction is unconfounded are used.

13-65.

A	B	C	D	E = ABCD	AB = CDE	block
-	-	-	-	+	+	1
+	-	-	-	-	-	2
-	+	-	-	-	-	2
+	+	-	-	+	+	1
-	-	+	-	-	+	1
+	-	+	-	+	-	2
-	+	+	-	+	-	2
+	+	+	-	-	+	1
-	-	-	+	-	+	1
+	-	-	+	+	-	2
-	+	-	+	+	-	2
+	+	-	+	-	+	1
-	-	+	+	+	+	1
+	-	+	+	-	-	2
-	+	+	+	-	-	2
+	+	+	+	+	+	1

This uses AB = CDE as the effect to confound with blocks.

13-66. The generators are F = ABCD and G = ABDE. The complete defining relation is
I = ABCDF = ABDEG = CEFG.
The design can be constructed in four blocks by confounding ACE = AFG and BCE = BFG with blocks. This also confounds AB = CDF = DEG with blocks.
Yes, a two-factor interaction is confounded with blocks. The best blocking scheme confounds only one two-factor interaction with blocks.

13-67. The generators are E = ABC, F = BCD, and G = ACD. The complete defining relation is I = ABCE = BCDF = ADEF = ACDG = BDEG = ABFG = CEFG. The alias set ABD = CDE = ACF = BEF = BCG = AEG = DFG can be used to construct the blocks. Then, only three-factor interaction are confounded with blocks.

13-68. a)

A	B	C	D = AB	E = AC	F = BC	G = ABC
-	-	-	+	+	+	-
+	-	-	-	-	+	+
-	+	-	-	+	-	+
+	+	-	+	-	-	-
-	-	+	+	-	-	+
+	-	+	-	+	-	-
-	+	+	-	-	+	-
+	+	+	+	+	+	+

The complete defining relation is I = ABD = ACE = BCDE = BCF = ACDF = ABEF = DEF = ABCG = CDG = BEG = ADEG = AFG = BDFG = CEFG = ABCDEFG.
The alias structure follows (including only one- and two-factor effects).
A = BD = CE = FG
B = AD = CF = EG
C = AE = BF = DG
D = AB = EF = CG
E = AC = DF = BG
F = BC = DE = AG
G = CD = BE = AF

b) The complete defining relation is I = -ABD = -ACE = BCDE = -BCF = ACDF = ABEF = -DEF = ABCG = -CDG = -BEG = ADEG = -AFG = BDFG = CEFG = -ABCDEFG.
The aliases (up to two-factor effects) are:
A = -BD = -CE = -FG
B = -AD = -CF = -EG
C = -AE = -BF = -DG
D = -AB = -EF = -CG
E = -AC = -DF = -BG
F = -BC = -DE = -AG
G = -CD = -BE = -AF

c) All the main effects can be estimated by averaging the estimates of an alias set containing a main effect from each fraction because the two-factor effects will cancel.

CHAPTER 14

Section 14-2

14-1. Sample Median = 7.385
Using the binomial distribution with n=10 and p=0.5, P-value = $2P(R^*\geq 8|p=0.5)=0.109$

14-2. Sample median = 8.44
Using the binomial distribution with n=19 and p=0.5, P-value = $2P(R^*\leq 7|p=0.5)=0.359$

14-3. $H_0: \tilde{\mu} = 2.5$
$H1: \tilde{\mu} < 2.5$
Using the binomial distribution with n=20 and p=0.5, P-value = $P(R^*\leq 2|p=0.5)=0.0002$

14-4. $\dfrac{r^* - 0.5n}{0.5\sqrt{n}} = 1.98$ P-value = $P(|Z|>1.98)=0.048$

14-6. P-value = $P(R^*\geq 3|p=0.5,n=6)=0.656$

14-7. $Z_0 = -1.15$, The P-value = $P(|Z|>1.15) = 0.250$. Using the continuity correction, the P-value=0.359

14-8. $Z_0 = 3.58$, P-value=$P(Z>3.58) = 0.00017$. Using the continuity correction, $Z_0 = 3.35$, P-value=0.0004

14-9. P-value = $2P(R^*\geq 6|p=0.5)= 0.289$. Because r=2 > 0, do not reject H_0.

14-10. Because $r=3 \leq 3$ for n = 18, reject H_0

14-11. $Z_0 = 2.83$, P-value=$P(|Z|>2.83)=0.0047$, reject H_0. Using the continuity correction, $Z_0 = 2.59$, P-value = 0.0095, reject H_0.

14-12. Because $r = 2 > r_{0.05} = 0$ for n=8, H_0 is not rejected.

14-13. Because $r=1\leq 3$, for n = 15, reject H_0.

14-14. $Z_0 = 1.41$, P-value = 0.159. Using the continuity correction $Z_0 = 1.06$, P-value = 0.289

14-15. $Z_0 = -3.56$, P-value = 0.0004. Using the continuity correction $Z_0 = 3.09$, P-value = 0.0019

14-16. a) $f(x) = \lambda e^{-\lambda x}$ for x > 0 and $\int_0^{3.5} \lambda e^{-\lambda x} = 0.5$. Solving for λ, we find $\lambda = 0.198$ and $E(X) = 1/\lambda = 5.05$
b) Because $r^* = 3 > 1$, do not reject H_0
c) β = P(Type II error)=$P(R^*>1|\ \tilde{\mu}=4.5)$
For $\tilde{\mu} = 4.5$, E(X) = 6.49 and $P(X<3.5|\ \tilde{\mu}=4.5) = \int_0^{3.5} 0.154e^{-0.154x} = 0.5$ 0.4167. Therefore, with n = 10, p = 0.4167, β = P(Type II error)=$P(R^*>1|\ \tilde{\mu}=4.5) = 0.963$

14-17. a) $\alpha = P(Z>1.96) = 0.025$
b) $\beta = P\left(\dfrac{X}{\sigma/\sqrt{n}} 1.96 | \mu = 1\right) = P(Z < -1.20) = 0.115$
c) The sign test that rejects if $R^- \leq 1$ has $\alpha = 0.011$ based on the binomial distribution.
d) $\beta = P\left(R^- > 1 | \mu = 1\right) = 0.1587$. Therefore, R^- has a binomial distribution with p=0.1587 and n = 10 when $\mu = 1$. Then $\beta = 0.487$. The value of β is greater for the sign test than for the normal test because the Z-test was designed for the normal distribution.

14-18. P-value = $2P(R^- \geq 6|p=0.5) = 0.289$

14-19. P-value = $2P(R^- \leq 3|p=0.5) = 0.0075$. The exact p-value computed here agrees with the normal approximation in Exercise 14-11 in the sense that both calculations would lead to the rejection of H_0. The continuity correction improves the accuracy of the normal approximation.

Section 14-3

14-20. Sample median = 7.385, $w = 4.5 < w^* = 8$, reject H_0

14-21. Because $w = 80.5 > 46$, do not reject H_0

14-22. From exercise 14-21, $Z_0 = 0.6036$, P-value=0.546,. do not reject H_0

14-23. Because $w = 5 < 60$, reject H_0

14-24. Because $w = 11.5 > 3$, do not reject H_0

14-25. Because $w = 17.5 < 27$, reject H_0

14-26. Because $w = 14.5 > 3$, do not reject H_0

14-27. Because $w = 1 < 25$, reject H_0

Section 14-4

14-28. Because $w_2 = 75 > w_{0.025} = 51$, do not reject H_0

14-29. Set $w=\min(w_1, w_2)$. Because $w = 38 > w_{0.01} = 23$, do not reject H_0

14-30. Set $w=\min(w_1, w_2)$. Because $w = 77 < w_{0.05} = 78$, reject H_0

14-31. $Z_0 = 0.577$, P-value = 0.716

14-32. $Z_0 = -2.12$ and P-value = 0.034

14-33. Set $w=\min(w_1, w_2)$. Because $w = 73 < w_{0.05} = 78$, reject H_0

14-34. Set $w=\min(w_1, w_2)$. Because $w = 207 > w_{0.05} = 185$, do not reject H_0

14-35. $Z_0 = 2.49$, P-value = 0.013

14-36. $Z_0 = 1.06$, P-value = 0.289

Section 14-5

14-37.
```
Kruskal-Wallis Test on strength
mixingte    N    Median   Ave Rank         Z
1           4      2945        9.6      0.55
2           4      3075       12.9      2.12
3           4      2942        9.0      0.24
4           4      2650        2.5     -2.91
Overall    16                  8.5
H = 10.00  DF = 3  P = 0.019
H = 10.03  DF = 3  P = 0.018 (adjusted for ties)
* NOTE * One or more small samples
```
Reject H_0

14-38.

```
Kruskal-Wallis Test on strength
method      N    Median   Ave Rank         Z
1           5     550.0        8.7      0.43
2           5     553.0       10.7      1.65
3           5     528.0        4.6     -2.08
Overall    15                  8.0
H = 4.83  DF = 2  P = 0.089
H = 4.84  DF = 2  P = 0.089 (adjusted for ties)
```

Do not reject H_0

14-39.

```
Kruskal-Wallis Test on angle
manufact    N    Median   Ave Rank         Z
1           5     39.00        7.6     -1.27
2           5     44.00       12.4      0.83
3           5     48.00       15.8      2.31
4           5     30.00        6.2     -1.88
Overall    20                 10.5
H = 8.37  DF = 3  P = 0.039
```

Do not reject H_0

14-40.

```
Kruskal-Wallis Test on UNIFORMI
FLOW        N    Median   Ave Rank         Z
125         6     3.100        5.8     -2.06
160         6     4.400       13.0      1.97
200         6     3.800        9.7      0.09
Overall    18                  9.5
H = 5.42  DF = 2  P = 0.067
H = 5.44  DF = 2  P = 0.066 (adjusted for ties)
```

Do not reject H_0

14-41. P-value = 0.018 (use the chi-square distribution)

14-42. P-value = 0.066 (use the chi-square distribution)

Supplemental Exercises

14-43. Because $w = 5 > w_{0.05} = 1$, do not reject H_0

14-44.

```
Sign test of median = 10.00 versus  not =  10.00
               N  Below  Equal  Above         P     Median
y             10      5      0      5    1.0000      10.01
```

Using continuity correction, P-value= 0.752

14-45.

```
Sign test of median = 6.000 versus  <  6.000
               N  Below  Equal  Above         P     Median
y             15      9      2      4    0.1334      4.000
```

Do not reject H_0

14-46. $Z_0 = -1.39$, P-value=0.082, using the continuity correction, P-value = 0.133

14-47.

```
test1      N =   8     Median =      1.4000
test2      N =   8     Median =      1.6500
Point estimate for ETA1-ETA2 is     -0.2000
95.9 Percent CI for ETA1-ETA2 is (-0.4999,0.1001)
W = 53.5
Test of ETA1 = ETA2  vs  ETA1 not = ETA2 is significant at 0.1415
The test is significant at 0.1377 (adjusted for ties)
Cannot reject at alpha = 0.05
```

14-48.

	N	N for Test	Wilcoxon Statistic	P	Estimated Median
y	10	10	28.5	0.959	10.00

Test of median = 10.00 versus median not = 10.00

Do not reject H_0

14-49.

Test of median = 6.000 versus median < 6.000

	N	N for Test	Wilcoxon Statistic	P	Estimated Median
y	15	13	19.0	0.035	5.000

Reject H_0

14-50. P-value = $P(R^+ \geq 10|p=0.5) = 0.00098$, where R^+ is the number of differences $X_{before} - X_{after}$

14-51. Set $w=\min(w_1, w_2)$, Because w = 55 < 71, reject H_0

14-52. $Z_0 = 3.77$, P-value = 0.00016. Reject H_0

14-53. $Z_0 = 4.75$, Reject H_0

14-54. From exercise 14-53, $Z_0 = 4.75$, P-value = $P(Z>4.75) = 0$

14-55. Kruskal-Wallis Test on RESISTAN

ALLOY	N	Median	Ave Rank	Z
1	10	98.00	5.7	-4.31
2	10	102.50	15.3	-0.09
3	10	138.50	25.5	4.40
Overall	30		15.5	

H = 25.30 DF = 2 P = 0.000
H = 25.45 DF = 2 P = 0.000 (adjusted for ties)

Reject H_0

14-56.

Kruskal-Wallis Test on SCORE

METHOD	N	Median	Ave Rank	Z
1	8	23.50	15.1	1.29
2	8	21.75	8.1	-2.14
3	8	23.05	14.3	0.86
Overall	24		12.5	

H = 4.66 DF = 2 P = 0.098
H = 4.66 DF = 2 P = 0.097 (adjusted for ties)

Do not reject H_0

14-57.

Kruskal-Wallis Test on VOLUME

TEMPERAT	N	Median	Ave Rank	Z
70	5	1245	12.4	2.69
75	5	1220	7.9	-0.06
80	5	1170	3.7	-2.63
Overall	15		8.0	

H = 9.46 DF = 2 P = 0.009
H = 9.57 DF = 2 P = 0.008 (adjusted for ties)

Reject H_0

Mind-Expanding Exercises

14-58. Under the null hypothesis each rank has probability of 0.5 of being either positive or negative. Define the random variable X_i as

$$X_i = \begin{cases} 1 & \text{rank is positive} \\ 0 & \text{rank is negative} \end{cases}$$

Then,

$$R^+ = \sum_{i=1}^{n} x_i i$$

and

$$E(R^+) = \sum_{i=1}^{n} E(x_i) i = \frac{1}{2}\sum i = \frac{n(n+1)}{4} \qquad \text{since} \qquad E(x_i) = \frac{1}{2}$$

where

$$V(R^+) = \sum V(x_i) i^2$$

by independence

$$V(X_i) = 1^2\frac{1}{2} + 0^2\frac{1}{2} - [E(X_i)]^2$$

$$= \frac{1}{2} - \frac{1}{4} = \frac{1}{4}$$

Then,

$$V(R^+) = \frac{n(n+1)2n+1)}{24}$$

14-59. a) 32 sequences are possible

b) Because each sequence has probability 1/32 under H_0, the distribution of W^* is obtained by counting the sequences that result in each value of W^*

14-60. b)

w^*	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Prob*1/32	1	1	1	2	2	3	3	3	3	3	3	2	2	1	1	1

c) $P(w^* > 13) = 2/32$

d) By enumerating the possible sequences, the probability that W^* exceeds any value can be calculated under the null hypothesis as in part (c). This approach can be used to determine the critical values for the test.

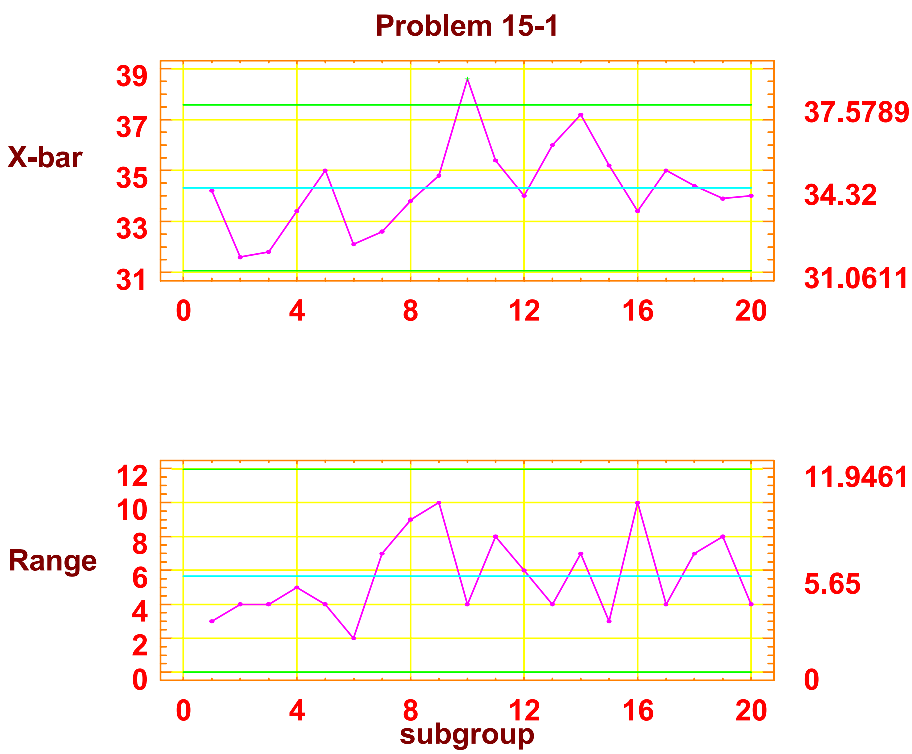
CHAPTER 15

Note to the Instructor: Many of the control charts and control chart summaries are created using Statgraphics. For Minitab users, the subroutine 'I and MR Charts' under "Control Charts" will provide similar results.

Section 15-5

15-1. a)

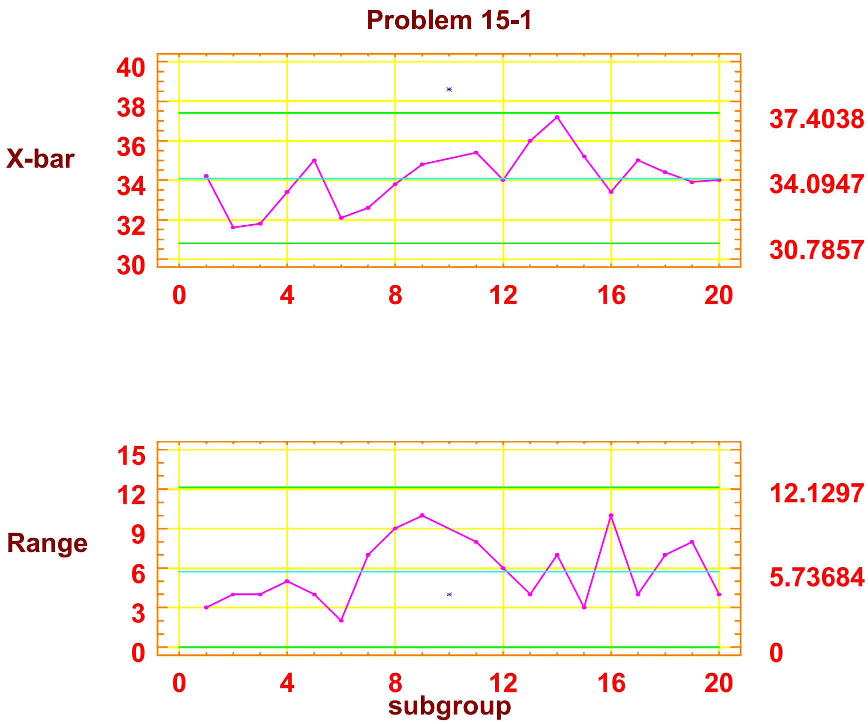
```
                    X-bar and Range - Initial Study
                            Problem 15-1

X-bar                                     | Range
-----                                     | -----
UCL:  +    3.0 sigma = 37.5789            | UCL:  +    3.0 sigma = 11.9461
Centerline           = 34.32              | Centerline           =  5.65
LCL:  -    3.0 sigma = 31.0611            | LCL:  -    3.0 sigma =  0
                                          |
out of limits = 1                         | out of limits = 0
--------------------------------------------------------------------------------
Chart: Both       Normalize: No

   20 subgroups, size 5                        0 subgroups excluded

Estimated
process mean  = 34.32
process sigma =  2.42906
mean Range    =  5.65
```



b)

```
                              Charting xbar
X-bar                                     | Range
-----                                     | -----
UCL:  +    3.0 sigma = 37.4038            | UCL:  +    3.0 sigma = 12.1297
Centerline           = 34.0947            | Centerline           =  5.73684
LCL:  -    3.0 sigma = 30.7857            | LCL:  -    3.0 sigma =  0
                                          |
out of limits = 0                         | out of limits = 0
```

```
--------------------------------------------------------------------------------
20 subgroups, size 5                          1 subgroups excluded
Estimated
process mean  = 34.0947
process sigma =  2.4664
mean Range    =  5.73684
```



15-2. $\bar{\bar{x}} = \frac{362.75}{25} = 14.510 \qquad \bar{r} = \frac{8.60}{25} = 0.344$

$\bar{x}$ chart

$UCL = CL + A_2\bar{r} = 14.510 + 0.577(0.344) = 14.708$

$CL = 14.510$

$LCL = CL - A_2\bar{r} = 14.510 - 0.577(0.344) = 14.312$

R chart

$UCL = D_4\bar{r} = 2.115(0.344) = 0.728$

$CL = 0.344$

$LCL = D_3\bar{r} = 0(0.344) = 0$

15-3. a)

```
                       X-bar and Range - Initial Study
                          Charting Problem 15-3
X-bar                                   | Range
-----                                   | -----
UCL:  +    3.0 sigma = 15.6299          | UCL:  +    3.0 sigma = 15.4047
Centerline           = 10.7125          | Centerline           =  6.75
LCL:  -    3.0 sigma =  5.79506         | LCL:  -    3.0 sigma =  0
                                        |
out of limits = 0                       | out of limits = 0
Estimated
process mean  = 10.7125
process sigma =  3.27829
mean Range    =  6.75
```

There are no points beyond the control limits. The process appears to be in control.

b) No points fell beyond the control limits, the limits do not need to be revised.

15-4.

$\overline{\overline{x}} = 20.0 \quad \frac{\overline{r}}{d_2} = 1.4 \quad d_2 = 2.534$

$\overline{r} = 1.4(2.534) = 3.5476$

$\overline{x}$ chart

$UCL = CL + A_2\overline{r} = 20.0 + 0.483(3.5476) = 21.71$

$CL = 20.0$

$LCL = CL - A_2\overline{r} = 20.0 - 0.483(3.5476) = 18.29$

R chart

$UCL = D_4\overline{r} = 2.004(3.53476) = 7.11$

$CL = 3.5476$

$LCL = D_3\overline{r} = 0(3.5476) = 0$

15-5. a) $\bar{\bar{x}} = \frac{7805}{35} = 223 \qquad \bar{r} = \frac{1200}{35} = 34.286$

$\bar{x}$ chart

$$UCL = CL + A_2\bar{r} = 223 + 0.577(34.286) = 242.78$$
$$CL = 223$$
$$LCL = CL - A_2\bar{r} = 223 - 0.577(34.286) = 203.22$$

R chart

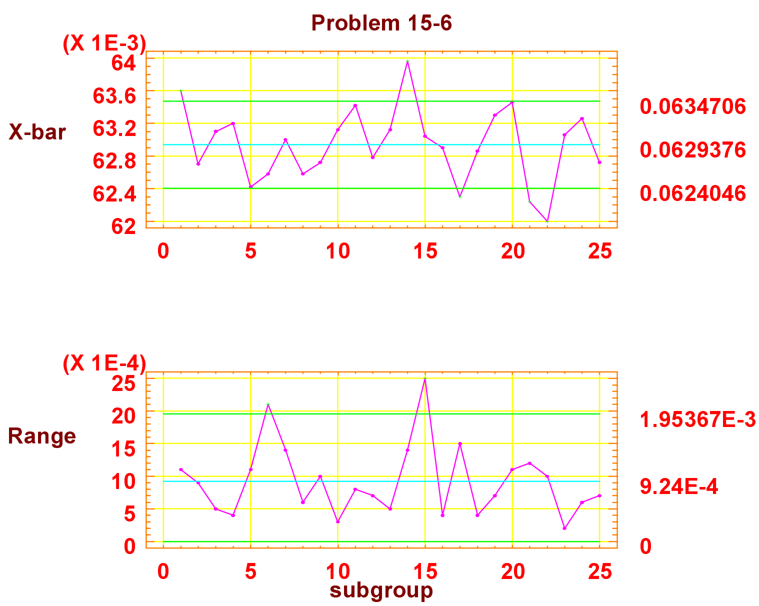
$$UCL = D_4\bar{r} = 2.115(34.286) = 72.51$$
$$CL = 34.286$$
$$LCL = D_3\bar{r} = 0(34.286) = 0$$

b)

$$\hat{\mu} = \bar{\bar{x}} = 223$$
$$\hat{\sigma} = \frac{\bar{r}}{d_2} = \frac{34.286}{2.326} = 14.74$$

15-6. a)

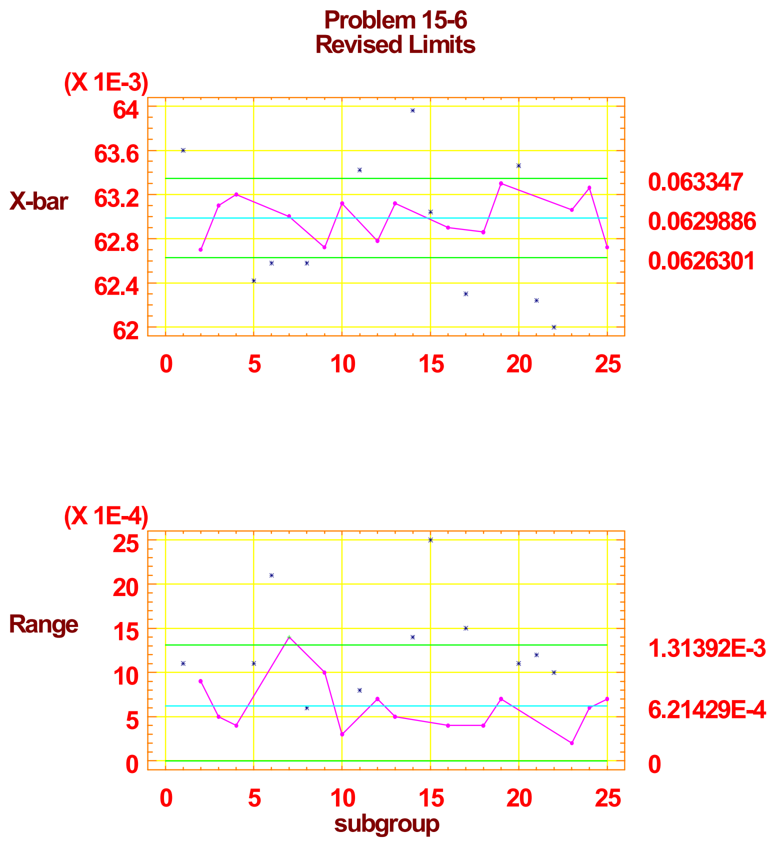
```
                  X-bar and Range - Initial Study
                     Charting Problem 15-6
X-bar                                   | Range
-----                                   | -----
UCL:  +    3.0 sigma = 0.0634706        | UCL:  +    3.0 sigma = 1.95367E-3
Centerline           = 0.0629376        | Centerline           = 9.24E-4
LCL:  -    3.0 sigma = 0.0624046        | LCL:  -    3.0 sigma = 0
                                        |
out of limits = 5                       | out of limits = 2
--------------------------------------------------------------------------------
Chart: Both      Normalize: No
25 subgroups, size 5                      0 subgroups excluded
Estimated
process mean  = 0.0629376
process sigma = 3.97248E-4
mean Range    = 9.24E-4
```



There are several points out of control. The control limits need to be revised. The points are 1, 5, 14,17, 20,21, and 22; or outside the control limits of the R chart: 6 and 15

b) X-bar and Range - Revised Limits

```
                                 Charting Problem 15-6
X-bar                                        | Range
-----                                        | -----
UCL:  +     3.0 sigma = 0.063347             | UCL:  +     3.0 sigma = 1.31392E-3
Centerline            = 0.0629886            | Centerline            = 6.21429E-4
LCL:  -     3.0 sigma = 0.0626301            | LCL:  -     3.0 sigma = 0
                                             |
out of limits = 0                            | out of limits = 1
Estimated
process mean  = 0.0629886
process sigma = 2.67166E-4
mean Range    = 6.21429E-4
```



15-7. a) X-bar and Range - Initial Study

```
Charting Problem 15-7
X-bar                                        | Range
-----                                        | -----
UCL:  +     3.0 sigma = 7.44253              | UCL:  +     3.0 sigma = 2.9153
Centerline            = 6.2836               | Centerline            = 1.1328
LCL:  -     3.0 sigma = 5.12467              | LCL:  -     3.0 sigma = 0
                                             |
out of limits = 0                            | out of limits = 0
Estimated
process mean  = 6.2836
process sigma = 0.669108
mean Range    = 1.1328
```

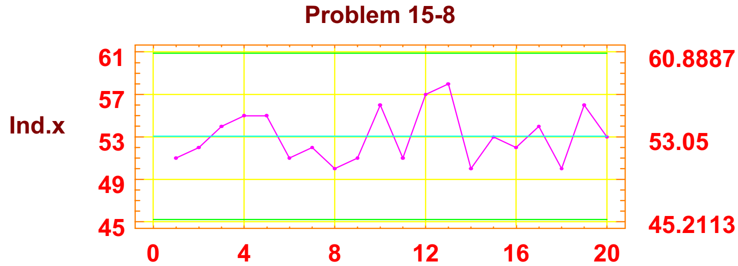
There are no points beyond the control limits. The process appears to be in control.

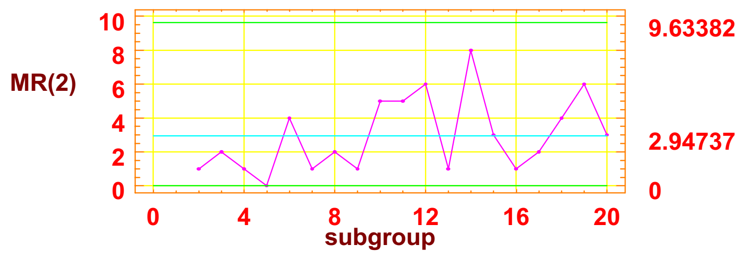
b) No points fell beyond the control limits, the limits do not need to be revised.

<u>Section 15-6</u>

15-8. a)

```
                        Individuals and MR(2) - Initial Study
--------------------------------------------------------------------------------
                             Charting Problem 15-8

Ind.x                                  | MR(2)
-----                                  | -----
UCL:  +    3.0 sigma = 60.8887         | UCL:  +    3.0 sigma = 9.63382
Centerline           = 53.05           | Centerline           = 2.94737
LCL:  -    3.0 sigma = 45.2113         | LCL:  -    3.0 sigma = 0
                                       |
out of limits = 0                      | out of limits = 0
--------------------------------------------------------------------------------
Chart: Both      Normalize: No

   20 subgroups, size 1                    0 subgroups excluded

Estimated
process mean  = 53.05
process sigma =  2.61292
mean MR(2)    =  2.94737
```





There are no points beyond the control limits. The process appears to be in control.

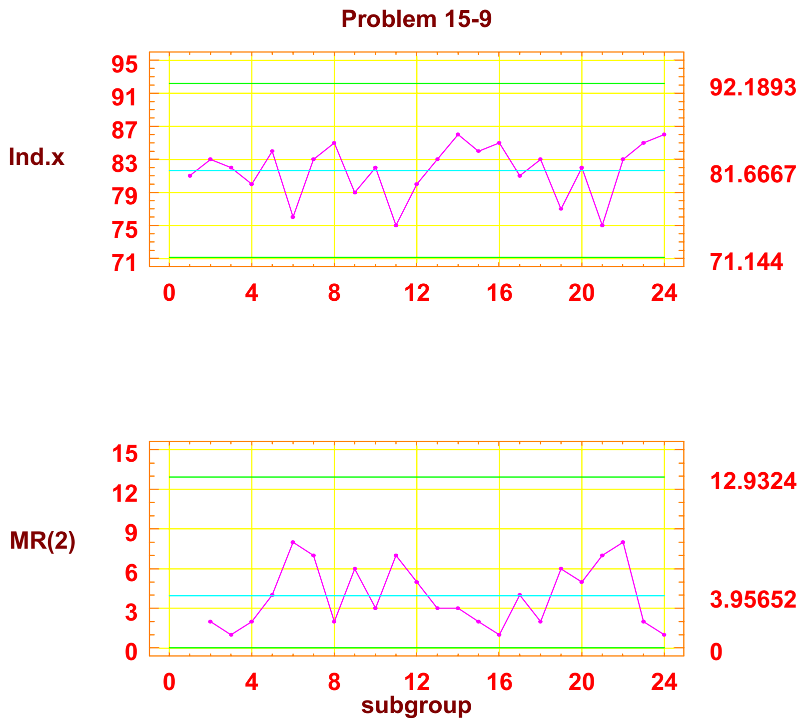
b)

$$\hat{\mu} = \overline{x} = 53.05$$

$$\hat{\sigma} = \frac{\overline{mr}}{d_2} = \frac{2.94737}{1.128} = 2.613$$

15-9. a)

```
                              Individuals and MR(2) - Initial Study
--------------------------------------------------------------------------------
                                Charting Problem 15-9

Ind.x                                  | MR(2)
-----                                  | -----
UCL:  +    3.0 sigma = 92.1893         | UCL:  +    3.0 sigma = 12.9324
Centerline           = 81.6667         | Centerline           =  3.95652
LCL:  -    3.0 sigma = 71.144          | LCL:  -    3.0 sigma =  0
                                       |
out of limits = 0                      | out of limits = 0
--------------------------------------------------------------------------------
Chart: Both      Normalize: No

   24 subgroups, size 1                    0 subgroups excluded

Estimated
process mean  = 81.6667
process sigma =  3.50755
mean MR(2)    =  3.95652
```



There are no points beyond the control limits. The process appears to be in control.

b)

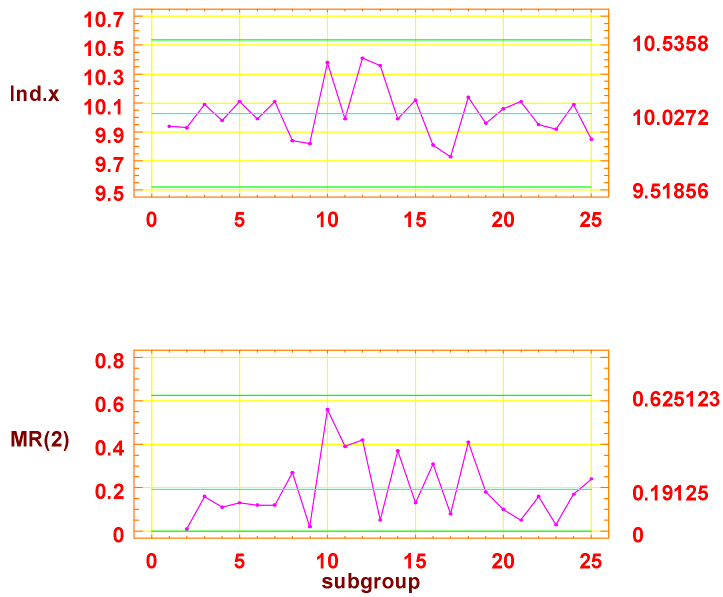
$\hat{\mu} = \overline{x} = 81.6667$

$$\hat{\sigma} = \frac{\overline{mr}}{d_2} = \frac{3.95652}{1.128} = 3.508$$

15-10. a) Ind.x and MR(2) - Initial Study

```
--------------------------------------------------------------------------------
                             Charting diameter

Ind.x                                   | MR(2)
-----                                   | -----
UCL:  +    3.0 sigma = 10.5358          | UCL:  +    3.0 sigma = 0.625123
Centerline           = 10.0272          | Centerline           = 0.19125
LCL:  -    3.0 sigma =  9.51856         | LCL:  -    3.0 sigma = 0
                                        |
out of limits = 0                       | out of limits = 0
--------------------------------------------------------------------------------
Chart: Both      Normalize: No

   25 subgroups, size 1                     0 subgroups excluded

Estimated
process mean  = 10.0272
process sigma =  0.169548
mean MR(2)    =  0.19125
```

Problem 15-10



There are no points beyond the control limits. The process appears to be in control.

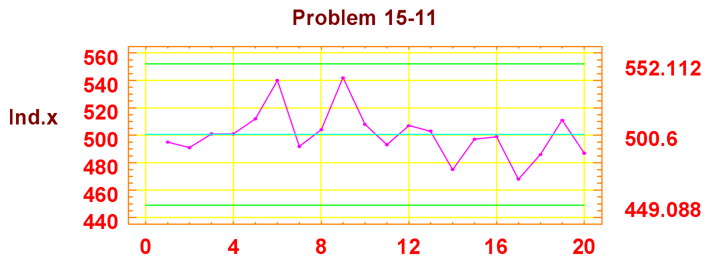
b)

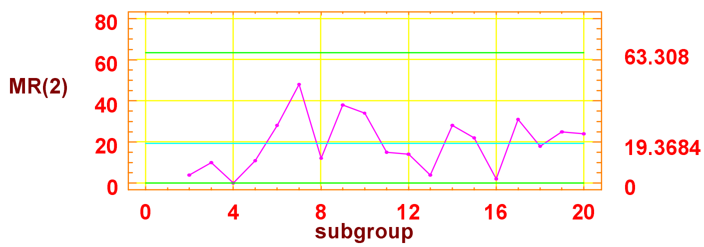
$$\hat{\mu} = \overline{x} = 10.0272$$

$$\hat{\sigma} = \frac{\overline{mr}}{d_2} = \frac{0.19125}{1.128} = 0.16955$$

15-11. a)

```
                              Ind.x and MR(2) - Initial Study
--------------------------------------------------------------------------------
                                   Charting Problem 15-11

Ind.x                                      | MR(2)
-----                                      | -----
UCL:  +    3.0 sigma = 552.112             | UCL:  +    3.0 sigma = 63.308
Centerline           = 500.6               | Centerline           = 19.3684
LCL:  -    3.0 sigma = 449.088             | LCL:  -    3.0 sigma =  0
                                           |
out of limits = 0                          | out of limits = 0
--------------------------------------------------------------------------------
Chart: Both       Normalize: No

   20 subgroups, size 1                          0 subgroups excluded

Estimated
process mean  = 500.6
process sigma =  17.1706
mean MR(2)    =  19.3684
```





There are no points beyond the control limits. The process appears to be in control.

b)

$$\hat{\mu} = \bar{x} = 500.6$$

$$\hat{\sigma} = \frac{\overline{mr}}{d_2} = \frac{19.3684}{1.128} = 17.17$$

Section 15-7

15-12. a) If the process uses 66.7% of the specification band, then $6\sigma = 0.667(USL\text{-}LSL)$ then assume $\bar{\bar{x}} = \mu$ since the process is centered

$$3\sigma = 0.667(USL - \bar{\bar{x}}) = 0.677(\bar{\bar{x}} - LSL) = 0.667(USL - \mu)$$

$$4.5\sigma = USL - \mu = LSU - \mu$$

$$PCR_K \; = \min\left[\frac{4.5\sigma}{3\sigma}, \frac{4.5\sigma}{3\sigma}\right] = 1.5$$

Since PCR and PCR_k exceeds unity, the natural tolerance limits lie inside the specification limits and very few defective units will be produced.

b) Assuming a normal distribution with $6\sigma = 0.667(USL - LSL)$ and a centered process, then $3\sigma = 0.667(USL - \mu)$. Consequently, $USL - \mu = 4.5\sigma$ and $\mu - LSL = 4.5\sigma$

$$\begin{aligned} P(X > USL) &= P\left(Z > \frac{4.5\sigma}{\sigma}\right) \\ &= P(Z > 4.5) \\ &= 1 - P(Z < 4.5) \\ &= 1 - 1 \\ &= 0 \end{aligned}$$

By symmetry, the fraction defective is $2[P(X > USL)] = 0$.

15-13. a) $\hat{\sigma} = \frac{\bar{r}}{d_2} = \frac{5.737}{2.326} = 2.466$ or $\hat{\sigma} = 0.0002466$

$$PCR = \frac{USL - LSL}{6(\hat{\sigma})} = \frac{0.5045 - 0.5025}{6(0.0002466)} = 1.35$$

$$\begin{aligned} PCR_K &= \min\left[\frac{USL - \bar{\bar{x}}}{3\hat{\sigma}}, \frac{\bar{\bar{x}} - LSL}{3\hat{\sigma}}\right] \\ &= \min\left[\frac{0.5045 - 0.5034}{3(0.0002466)}, \frac{0.5034 - 0.5025}{3(0.0002466)}\right] \\ &= \min[1.489, 1.217] \\ &= 1.217 \end{aligned}$$

Since PCR exceeds unity, the natural tolerance limits lie inside the specification limits and very few defective units will be produced.

Since $PCR_K \neq PCR$ the process is slightly off center.

b) Assuming a normal distribution with $\hat{\mu} = 0.5034$ and $\hat{\sigma} = 0.0002466$

$$\begin{aligned} P(X < LSL) &= P\left(Z < \frac{LSL - \hat{\mu}}{\hat{\sigma}}\right) \\ &= P(Z < -3.65) \\ &= 1 - P(Z < 3.65) \\ &= 1 - 0.99987 \\ &= 0.00013 \end{aligned}$$

$$\begin{aligned} P(X > USL) &= P\left(Z > \frac{USL - \hat{\mu}}{\hat{\sigma}}\right) \\ &= P(Z > 4.46) \\ &= 1 - P(Z < 4.46) \\ &= 1 - 1 \\ &= 0 \end{aligned}$$

Therefore, the proportion nonconforming is given by
$P(X<LSL) + P(X>USL) = 0.00013 + 0$
$= 0.00013$

15-14. a) Assuming a normal distribution with $\hat{\mu} = 0.14.510$ and $\hat{\sigma} = \frac{\bar{r}}{d_2} = \frac{0.344}{2.326} = 0.148$

$$\begin{aligned} P(X < LSL) &= P\left(Z < \frac{LSL - \hat{\mu}}{\hat{\sigma}}\right) \\ &= P\left(Z < \frac{14.00 - 14.51}{0.148}\right) \\ &= P(Z < -3.45) \\ &= 1 - P(Z < 3.45) \\ &= 1 - 0.99972 \\ &= 0.00028 \end{aligned}$$

$$P(X > USL) = P\left(Z > \frac{USL - \hat{\mu}}{\hat{\sigma}}\right)$$
$$= P\left(Z > \frac{15.00 - 14.51}{0.148}\right)$$
$$= P(Z > 3.31)$$
$$= 1 - P(Z < 3.31)$$
$$= 1 - 0.99953$$
$$= 0.00047$$

Therefore, the proportion nonconforming is given by
$P(X<LSL) + P(X>USL) = 0.00028 + 0.00047$
$= 0.00075$

b)

$$PCR = \frac{USL - LSL}{6(\hat{\sigma})} = \frac{15.00 - 14.00}{6(0.148)} = 1.13$$
$$PCR_K = \min\left[\frac{USL - \bar{\bar{x}}}{3\hat{\sigma}}, \frac{\bar{\bar{x}} - LSL}{3\hat{\sigma}}\right]$$
$$= \min\left[\frac{15.00 - 14.51}{3(0.148)}, \frac{14.51 - 14.00}{3(0.148)}\right]$$
$$= \min[1.104, 1.15]$$
$$= 1.104$$

Since PCR exceeds unity, the natural tolerance limits lie inside the specification limits and very few defective units will be produced.

$PCR_K \cong PCR$ the process appears to be centered.

15-15. $\hat{\sigma} = \frac{\bar{r}}{d_2} = \frac{6.75}{2.059} = 3.278$ or 0.0328 mm

$\bar{x} = 10.7125$

$$P(X > 15) + P(X < 5)$$
$$= P\left(Z > \frac{15 - 10.7125}{3.278}\right) + P\left(Z < \frac{5 - 10.7125}{3.278}\right)$$
$$= P(Z > 1.31) + P(Z < -1.74)$$
$$= 0.0951 + 0.04093$$
$$= 0.1360$$

$$PCR = \frac{15 - 5}{6(3.28)} = 0.508$$

With the PCR much less than unity, the process capability appears to be poor.

15-16. If the process uses 85% of the spec band then $6\sigma = 0.85(USL - LSL)$ and

$$PCR = \frac{USL - LSL}{0.85(USL - LSL)} = \frac{1}{0.85} = 1.18$$

Assume $\bar{\bar{x}} = \mu$ and $3\sigma = 0.85(USL - \bar{\bar{x}}) = 0.85(\mu - LSL)$

Therefore,

$$PCR_k = \min\left[\frac{3.53\hat{\sigma}}{3\hat{\sigma}}, \frac{3.53\hat{\sigma}}{3\hat{\sigma}}\right] = 1.18$$

Since PCR and PCR_k exceed unity, the natural tolerance limits lie inside the specification limits and very few defective units will be produced.

15-17. Assuming a normal distribution with $\hat{\mu}$ = 0.223 and $\hat{\sigma} = \frac{34.286}{2.326} = 14.74$

$$P(X < LSL) = P\left(Z < \frac{LSL - \hat{\mu}}{\hat{\sigma}}\right)$$
$$= P\left(Z < \frac{180 - 223}{14.74}\right)$$
$$= P(Z < -2.92)$$
$$= 1 - P(Z < 2.92)$$
$$= 1 - 0.99825$$
$$= 0.00175$$

$$P(X > USL) = P\left(Z > \frac{USL - \hat{\mu}}{\hat{\sigma}}\right)$$
$$= P\left(Z > \frac{260 - 223}{14.74}\right)$$
$$= P(Z > 2.51)$$
$$= 1 - P(Z < 2.51)$$
$$= 1 - 0.99396$$
$$= 0.00604$$

Therefore, the proportion nonconforming is given by
$P(X<LSL) + P(X>USL) = 0.00175 + 0.00604$
$= 0.00779$

$$PCR = \frac{USL - LSL}{6(\hat{\sigma})} = \frac{260 - 180}{6(14.74)} = 0.905$$

$$PCR_K = \min\left[\frac{USL - \bar{\bar{x}}}{3\hat{\sigma}}, \frac{\bar{\bar{x}} - LSL}{3\hat{\sigma}}\right]$$
$$= \min\left[\frac{260 - 223}{3(14.74)}, \frac{223 - 180}{3(14.74)}\right]$$
$$= \min[0.837, 0.972]$$
$$= 0.837$$

The process capability is marginal.

15-18. Assuming a normal distribution with $\hat{\mu}$ = 20.0 and $\hat{\sigma}$ = 1.4

$$PCR = \frac{USL - LSL}{6(\hat{\sigma})} = \frac{25 - 15}{6(1.4)} = 1.19$$

$$PCR_K = \min\left[\frac{USL - \bar{\bar{x}}}{3\hat{\sigma}}, \frac{\bar{\bar{x}} - LSL}{3\hat{\sigma}}\right]$$
$$= \min\left[\frac{25 - 20}{3(1.4)}, \frac{20 - 15}{3(1.4)}\right]$$
$$= \min[1.19, 1.19]$$
$$= 1.19$$

The process is capable.

15-19. Assuming a normal distribution with $\hat{\mu} = 6.284$ and $\hat{\sigma} = \dfrac{1.1328}{1.693} = 0.669$

$$PCR = \frac{USL - LSL}{6(\hat{\sigma})} = \frac{7-5}{6(0.669)} = 0.50$$

$$\begin{aligned} PCR_K &= \min\left[\frac{USL - \bar{\bar{x}}}{3\hat{\sigma}}, \frac{\bar{\bar{x}} - LSL}{3\hat{\sigma}}\right] \\ &= \min\left[\frac{7-6.284}{3(0.669)}, \frac{6.284-5}{3(0.669)}\right] \\ &= \min[0.357, 0.640] \\ &= 0.357 \end{aligned}$$

The process capability is poor.

15-20. Assuming a normal distribution with $\hat{\mu} = 0.0629$ and $\hat{\sigma} = \dfrac{0.000924}{2.326} = 0.00040$

The natural tolerance limits are then

$$\begin{aligned} \hat{\mu} \pm 3\hat{\sigma} &= 0.0629 \pm 3(0.00040) \\ &= (0.0617, 0.0641) \end{aligned}$$

15-21. Assuming a normal distribution with $\hat{\mu} = 500.6$ and $\hat{\sigma} = 17.17$

$$PCR = \frac{USL - LSL}{6(\hat{\sigma})} = \frac{525-475}{6(17.17)} = 0.49$$

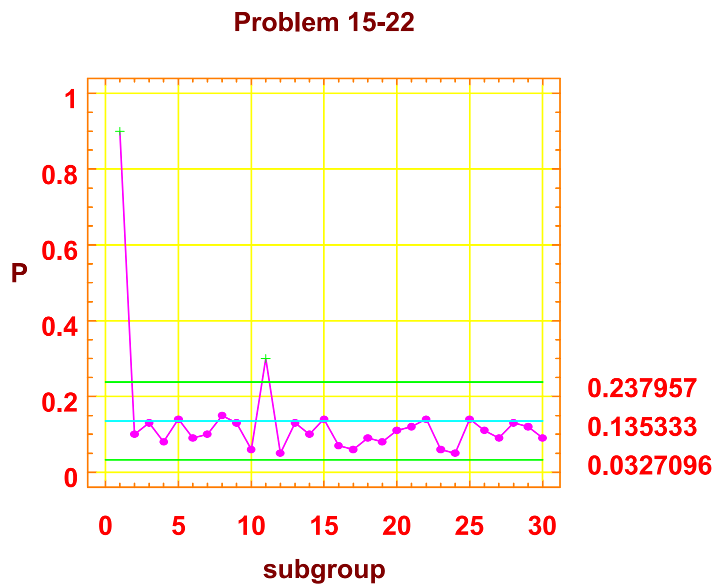
$$\begin{aligned} PCR_K &= \min\left[\frac{USL - \bar{\bar{x}}}{3\hat{\sigma}}, \frac{\bar{\bar{x}} - LSL}{3\hat{\sigma}}\right] \\ &= \min\left[\frac{525-500.6}{3(17.17)}, \frac{500.6-475}{3(17.17)}\right] \\ &= \min[0.474, 0.50] \\ &= 0.474 \end{aligned}$$

Since the process capability ratios are less than unity, the process capability appears to be poor.

Section 15-8

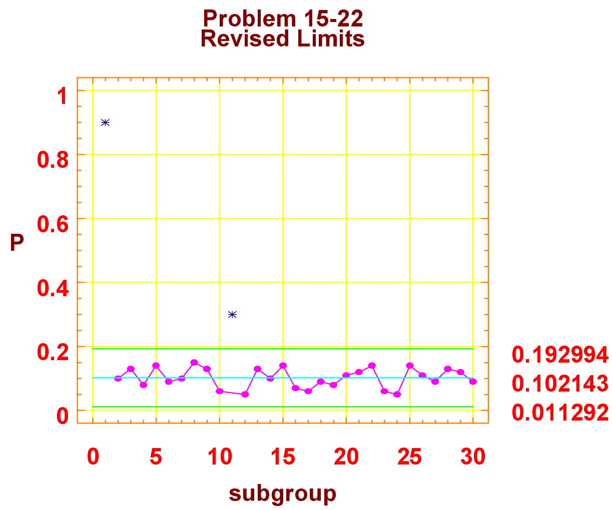
15-22.

```
                              P Chart - Initial Study
--------------------------------------------------------------------------------
                               Charting Problem 15-22

P Chart
-----
UCL:  +    3.0 sigma = 0.237957
Centerline           = 0.135333
LCL:  -    3.0 sigma = 0.0327096

out of limits = 2
--------------------------------------------------------------------------------
Normalize:  No

   30 subgroups, size 100                    0 subgroups excluded

Estimated
mean P  = 0.135333
sigma   = 0.0342079
```



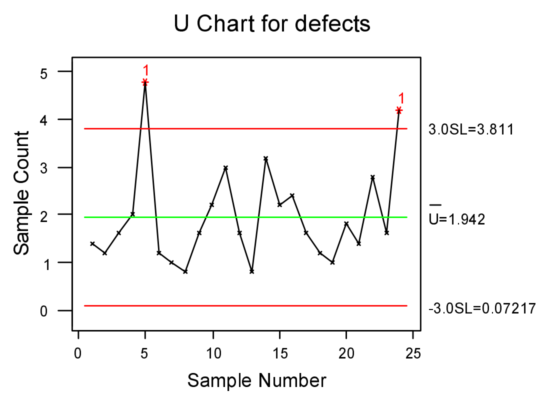
The process is out of control. The control limits need to be revised. The samples with out-of-control points are 1 and 11.

```
                              P Chart - Revised Limits
--------------------------------------------------------------------------------
                               Charting Problem 15-22

P Chart
-----
UCL:  +    3.0 sigma = 0.192994
Centerline           = 0.102143
LCL:  -    3.0 sigma = 0.011292

out of limits = 0
--------------------------------------------------------------------------------
Normalize:  No

   30 subgroups, size 100                    2 subgroups excluded

Estimated
mean P   = 0.102143
sigma    = 0.0302836
```



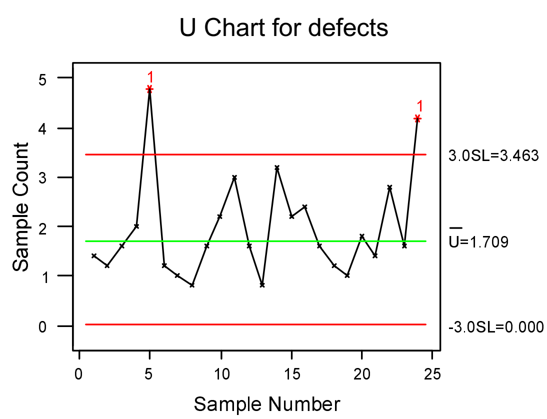
There are no points out of control for the revised limits.

15-23.



Samples 5 and 24 have out-of-control points. The limits need to be revised.

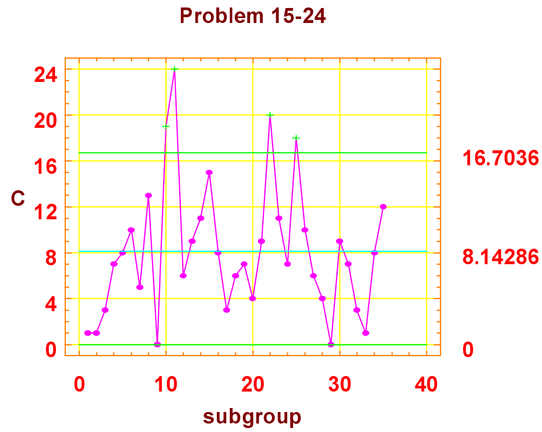
b)



The control limits are calculated without the out-of-control points. There are no points out of control for the revised limits.

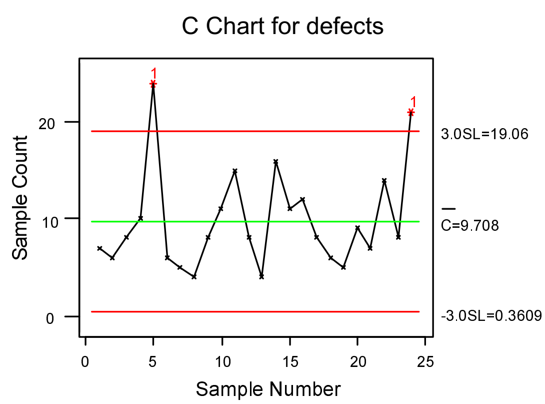
15-24.

```
                              C Chart - Initial Study
                              Charting Problem 15-24
C Chart
-----
UCL:  +    3.0 sigma = 16.7036
Centerline           =  8.14286
LCL:  -    3.0 sigma =  0

out of limits = 4
Estimated
mean  C = 8.14286
sigma   = 2.85357
```



The process is not in control. There are several points beyond the control limits.

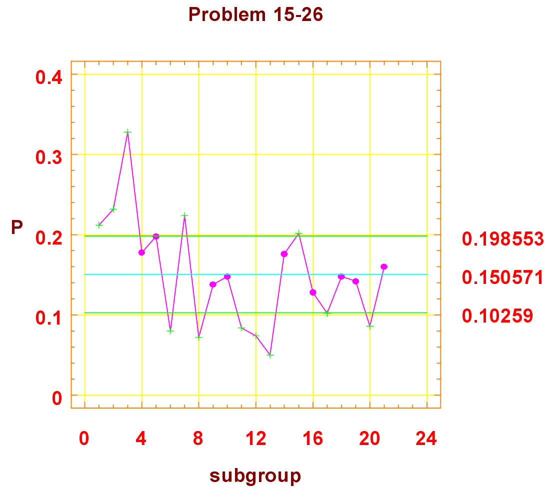
15-25.



There are two points beyond the control limits. They are samples 5 and 24.
The U chart and the C chart both detected out-of-control points at samples 5 and 24.

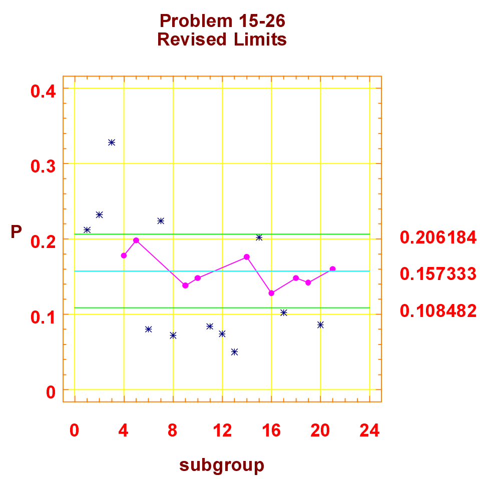
15-26.

```
                              P Chart - Initial Study
                                 Charting Problem 15-26
P Chart
-----
UCL:  +    3.0 sigma = 0.198553
Centerline           = 0.150571
LCL:  -    3.0 sigma = 0.10259

out of limits = 12
Estimated
mean P   = 0.150571
sigma    = 0.0159937
```



The samples with out-of-control points are 1, 2, 3, 6, 7, 8, 11, 12, 13, 15, 17, 20
There are several points are out of control. The control limits need to be revised.

```
                              P Chart - Revised Limits
                               Charting Problem 15-26
P Chart
-----
UCL:  +    3.0 sigma = 0.206184
Centerline           = 0.157333
LCL:  -    3.0 sigma = 0.108482
out of limits = 0
Estimated
mean P   = 0.157333
sigma    = 0.0162837
```



There are no points out of control for the revised limits.

Section 15-9

15-27. a) $\bar{x} = 74.01$ $\quad \sigma_{\bar{x}} = 0.0045$ $\quad \mu = 74.01$

$P(73.9865 < \overline{X} < 74.0135)$

$= P\left(\frac{73.9865 - 74.01}{0.0045} < \frac{X - \mu}{\hat{\sigma}_{\bar{x}}} < \frac{74.0135 - 74.01}{0.0045}\right)$

$= P(-5.22 < Z < 0.78)$

$= P(Z < 0.78) - P(Z < -5.22)$

$= P(Z < 0.78) - [1 - P(Z < 5.22)]$

$= 0.7823 - (1 - 1)$

$= 0.7823$

The probability that this shift will be detected on the next sample is p = 1–0.7823 = 0.2177.

b) $\text{ARL} = \frac{1}{p} = \frac{1}{0.2177} = 4.6$

15-28. a) $\mu + 3\frac{\sigma}{\sqrt{n}} = \text{UCL}$

$100 + 3\frac{\sigma}{\sqrt{4}} = 106$

$\sigma = \frac{2}{3}(106 - 100) = 4$

b) $\hat{\sigma}_{\bar{x}} = \frac{\hat{\sigma}}{\sqrt{n}} = \frac{4}{2} = 2, \quad \mu = 105$

P(94<X<106)

$= P\left(\frac{94 - 105}{2} < \frac{\overline{X} - \mu}{\sigma_{\bar{x}}} < \frac{106 - 105}{2}\right)$

$= P(-5.5 < Z < 0.5)$

$= P(Z < 0.5) - P(Z < -5.5)$

$= P(Z < 0.5) - [1 - P(Z < 5.5)]$

$= 0.69146 - 0$

$= 0.69146$

The probability that this shift will be detected on the next sample is p = 1–0.69146 = 0.30854.

c) $\text{ARL} = \frac{1}{p} = \frac{1}{0.30854} = 3.24$

15-29. a) $\hat{\sigma}_{\bar{x}} = \frac{\hat{\sigma}}{\sqrt{n}} = \frac{2.4664}{\sqrt{5}} = 1.103$, $\mu = 38$

$P(30.78 < \overline{X} < 37.404)$

$= P\left(\frac{30.78 - 38}{1.103} < \frac{\overline{X} - \mu}{\hat{\sigma}_{\bar{x}}} < \frac{37.404 - 38}{1.103}\right)$

$= P(-6.55 < Z < -0.54)$

$= P(Z < -0.54) - P(Z < -6.55)$

$= [1 - P(Z < 0.54)] - [1 - P(Z < 6.55)]$

$= 0.7054 - (1 - 1)$

$= 0.7054$

The probability that this shift will be detected on the next sample is p = 1–0.7054 = 0.2946.

b) $\text{ARL} = \frac{1}{p} = \frac{1}{0.2946} = 1.42$

15-30. a) $\hat{\sigma} = \frac{\overline{R}}{d_2} = \frac{0.344}{2.326} = 0.148$ $\hat{\sigma}_{\overline{x}} = \frac{\hat{\sigma}}{\sqrt{n}} = \frac{0.148}{\sqrt{5}} = 0.066$, $\mu = 14.6$

$P(14.312 < X < 14.708)$

$= P\left(\frac{14.312 - 14.6}{0.066} < \frac{X - \mu}{\sigma_x} < \frac{14.708 - 14.6}{0.066}\right)$

$= P(-4.36 < Z < 1.64)$

$= P(Z < 1.64) - P(Z < -4.36)$

$= P(Z < 1.64) - [1 - P(Z < 4.36)]$

$= 0.94950 - (1 - 1)$

$= 0.94950$

The probability that this shift will be detected on the next sample is p = 1−0.94950 = 0.0505.

b) $ARL = \frac{1}{p} = \frac{1}{0.0505} = 19.8$

15-31. a) $\hat{\sigma} = \frac{\overline{R}}{d_2} = \frac{6.75}{2.059} = 3.28$ $\hat{\sigma}_{\overline{x}} = \frac{\hat{\sigma}}{\sqrt{n}} = \frac{3.28}{\sqrt{4}} = 1.64$, $\mu = 14.5$

$P(5.795 < \overline{X} < 15.63)$

$= P\left(\frac{5.795 - 14.5}{1.64} < \frac{\overline{X} - \mu}{\sigma_{\overline{x}}} < \frac{15.63 - 14.5}{1.64}\right)$

$= P(-5.31 < Z < 0.69)$

$= P(Z < 0.69) - P(Z < -5.31)$

$= P(Z < 0.69) - [1 - P(Z < 5.31)]$

$= 0.7549 - (1 - 1)$

$= 0.7549$

The probability that this shift will be detected on the next sample is p = 1−0.7549 = 0.2451.

b) $ARL = \frac{1}{p} = \frac{1}{0.2451} = 4.08$

15-32. a) $\hat{\sigma} = \frac{\overline{R}}{d_2} = 14.74$ $\hat{\sigma}_{\overline{x}} = \frac{\hat{\sigma}}{\sqrt{n}} = \frac{14.74}{\sqrt{5}} = 6.592$

$P(203.22 < \overline{X} < 242.78)$

$= P\left(\frac{203.22 - 215}{6.592} < \frac{\overline{X} - \mu}{\sigma_{\overline{x}}} < \frac{242.78 - 215}{6.592}\right)$

$= P(-1.79 < Z < 4.214)$

$= P(Z < 4.214) - P(Z < -1.79)$

$= 1 - 0.0368$

$= 0.96327$

The probability that this shift will be detected on the next sample is p = 1−0.44568 = 0.55432.

b) $ARL = \frac{1}{p} = \frac{1}{0.55432} = 1.80$

15-33. a) $\hat{\sigma} = \frac{\overline{R}}{d_2} = 1.4$ $\hat{\sigma}_{\overline{X}} = \frac{\hat{\sigma}}{\sqrt{n}} = \frac{1.4}{\sqrt{5}} = 0.626$, $\mu = 17$

$P(18.12 < X < 21.88)$

$= P\left(\frac{18.12-17}{0.626} < \frac{\overline{X}-\mu}{\sigma_{\overline{X}}} < \frac{21.88-17}{0.626}\right)$

$= P(1.79 < Z < 7.80)$

$= P(Z < 7.80) - P(Z < 1.79)$

$= P(Z < 7.80) - P(Z < 1.79)]$

$= 1 - 0.96327$

$= 0.03763$

The probability that this shift will be detected on the next sample is p = 1−0.03763 = 0.96327.

b) $ARL = \frac{1}{p} = \frac{1}{0.96327} = 1.04$

15-34. a) $\hat{\sigma} = \frac{\overline{R}}{d_2} = 0.000397$ $\hat{\sigma}_{\overline{X}} = \frac{\hat{\sigma}}{\sqrt{n}} = \frac{0.000397}{\sqrt{5}} = 0.0001775$, $\mu = 0.0625$

$P(0.0624 < \overline{X} < 0.0635)$

$= P\left(\frac{0.0624-0.0625}{0.0001775} < \frac{\overline{X}-\mu}{\sigma_{\overline{X}}} < \frac{0.0635-0.0625}{0.0001775}\right)$

$= P(-0.56 < Z < 5.63)$

$= P(Z < 5.63) - P(Z < -0.56)$

$= 1 - 0.28774$

$= 0.71226$

The probability that this shift will be detected on the next sample is p = 1 − 0.71226 = 0.28774.

b) $ARL = \frac{1}{p} = \frac{1}{0.28774} = 3.475$

15-35. a) $\hat{\sigma} = \frac{\overline{R}}{d_2} = 0.669$ $\hat{\sigma}_{\overline{X}} = \frac{\hat{\sigma}}{\sqrt{n}} = \frac{0.669}{\sqrt{3}} = 0.386$,$\mu = 5.5$

$P(5.125 < \overline{X} < 7.443 | \mu = 5.5)$

$= P\left(\frac{5.125-5.5}{0386} < \frac{\overline{X}-\mu}{\sigma_{\overline{X}}} < \frac{7.443-5.5}{0.386}\right)$

$= P(-0.97 < Z < 5.03)$

$= P(Z < 5.03) - P(Z < -0.97)$

$= P(Z < 5.03) - [1 - P(Z < 0.97)]$

$= 1 - (1 - 0.83397)$

$= 0.83397$

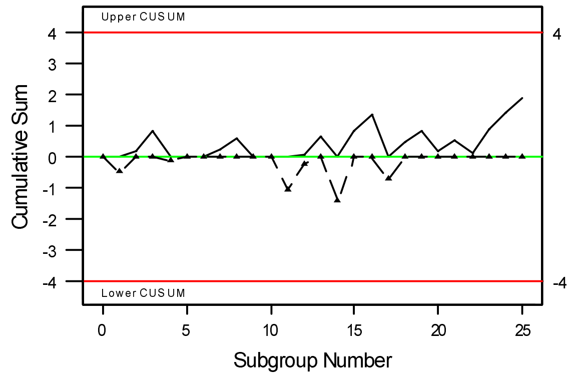
The probability that this shift will be detected on the next sample is p = 1−0.83397 = 0.16603.

b) $ARL = \frac{1}{p} = \frac{1}{0.16603} = 6.02$

Section 15-10

15-36. a)

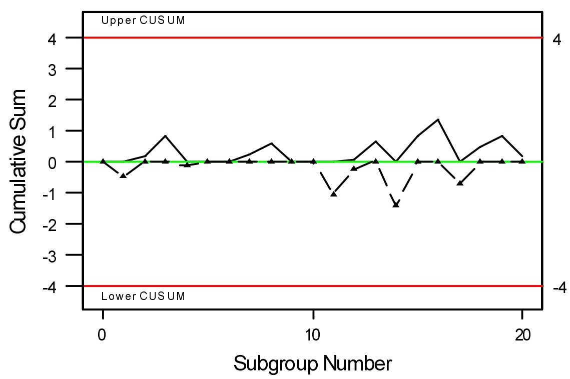
CUSUM Chart for purity



The CUSUM control chart for purity does not indicate an out-of-control situation. The S_H values do not plot beyond the values of –H and H.

b)

CUSUM Chart for purityad

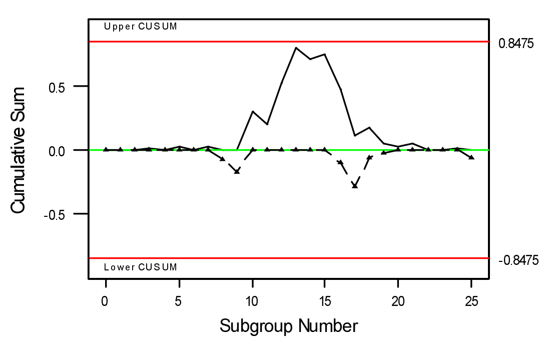


The process appears to be in statistical control.

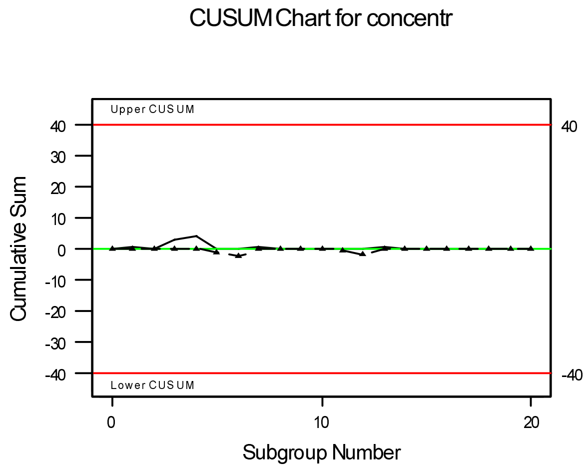
15-37. a) $\sigma = 0.1695$

b)

CUSUM Chart for diameter



15-38. a)



The process appears to be in statistical control.

b) With the target = 100 and a shift to 104 is a shift of $\frac{104-100}{4}$ = 0.5σ. From 15-9 with h = 5 and the row of 0.50, ARL = 38.0.

15-39. a) A shift to 51 is a shift of $\frac{\mu-\mu_0}{\sigma} = \frac{51-50}{2}$ = 0.5 standard deviations. From Table 15-9, ARL = 38.0

b) If n = 4, the shift to 51 is a shift of $\frac{\mu-\mu_0}{\sigma/\sqrt{n}} = \frac{51-50}{2/\sqrt{4}}$ = 1 standard deviation. From Table 15-9, ARL = 10.4

15-40. a) with target = 100 and a shift to 102 is a shift of $\frac{102-100}{4}$ = 0.5 standard deviations. From Table 15-9, ARL = 38. The hours of production are 2(38) = 76.

b) The ARL = 38. However, the time to obtain 38 samples is now 0.5(38) = 19.

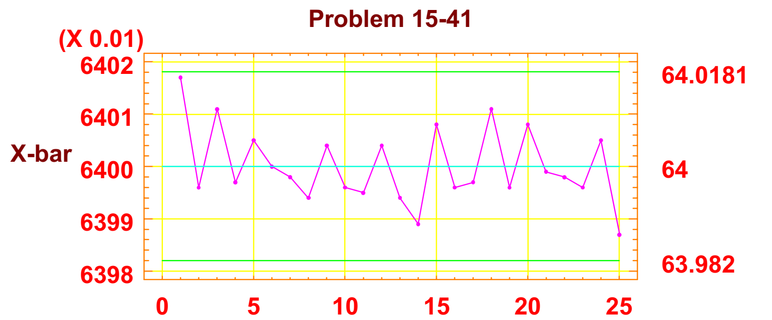
c) From Table 15-9, the ARL when there is no shift is 465. Consequently, the time between false alarms is 0.5(465) = 232.5 hours.

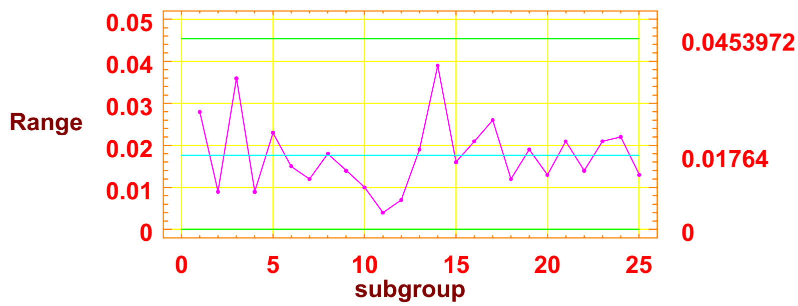
d) If the process shifts to 102, the shift is $\frac{\mu-\mu_0}{\sigma/\sqrt{n}} = \frac{102-100}{4/\sqrt{4}}$ 1 standard deviation. From Table 14-9, the ARL for this shift is 10.4. Therefore, the time to detect the shift is 2(10.4) = 20.8 hours. Although this time is slightly longer than the result in part (b), the time between false alarms is 2(465) = 930 hours, which is better than the result in part (c).

Supplementary Exercises

15-41. a)

```
                    X-bar and Range - Initial Study
------------------------------------------------------------------------------
X-bar                                      | Range
-----                                      | -----
UCL:  +    3.0 sigma = 64.0181             | UCL:  +    3.0 sigma = 0.0453972
Centerline           = 64                  | Centerline           = 0.01764
LCL:  -    3.0 sigma = 63.982              | LCL:  -    3.0 sigma = 0
                                           |
out of limits = 0                          | out of limits = 0
------------------------------------------------------------------------------
Chart: Both      Normalize: No
Estimated
process mean  = 64
process sigma =  0.0104194
mean Range    =  0.01764
```





The process is in control.

b) $\hat{\mu} = \overline{\overline{x}} = 64 \qquad \hat{\sigma} = \frac{\overline{R}}{d_2} = \frac{0.01764}{1.693} = 0.0104$

c) $PCR = \frac{USL - LSL}{6\hat{\sigma}} = \frac{64.02 - 63.98}{6(0.0104)} = 0.641$

The process does not meet the minimum capability level of $PCR \geq 1.33$.

d)

$$PCR_k = \min\left[\frac{USL - \overline{\overline{x}}}{3\hat{\sigma}}, \frac{\overline{\overline{x}} - LSL}{3\hat{\sigma}}\right]$$

$$= \min\left[\frac{64.02 - 64}{3(0.0104)}, \frac{64 - 63.98}{3(0.0104)}\right]$$

$$= \min[0.641, 0.641]$$

$$= 0.641$$

e) In order to make this process a "six-sigma process", the variance σ^2 would have to be decreased such that $PCR_k = 2.0$. The value of the variance is found by solving $PCR_k = \frac{\overline{\overline{x}} - LSL}{3\sigma} = 2.0$ for σ:

$$\frac{64 - 61}{3\sigma} = 2.0$$

$$6\sigma = 64. - 61$$

$$\sigma = \frac{64. - 61}{6}$$

$$\sigma = 0.50$$

Therefore, the process variance would have to be decreased to $\sigma^2 = (0.50)^2 = 0.025$.

f) $\hat{\sigma}_{\overline{X}} = 0.0104$

$P(63.98 < X < 64.02)$

$= P\left(\frac{63.98 - 64.005}{0.0104} < \frac{X - \mu}{\sigma_X} < \frac{64.02 - 64.005}{0.0104}\right)$

$= P(-2.40 < Z < 1.44)$

$= P(Z < 1.44) - P(Z < -2.40)$

$= 0.92506 - 0.00714$

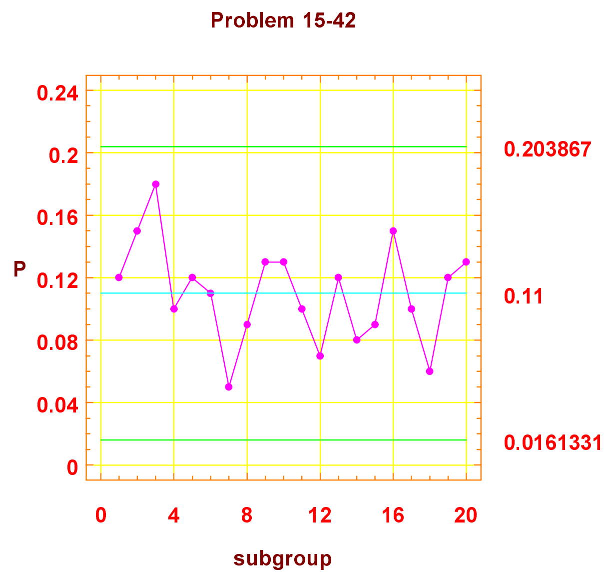
$= 0.91792$

The probability that this shift will be detected on the next sample is $p = 1 - 0.91792 = 0.08208$

$ARL = \frac{1}{p} = \frac{1}{0.08208} = 12.18$

15-42. a)

```
                                P Chart - Initial Study
P Chart
-----
UCL:  +    3.0 sigma = 0.203867
Centerline           = 0.11
LCL:  -    3.0 sigma = 0.0161331
out of limits = 0
Estimated
mean P  = 0.11
sigma   = 0.031289
```

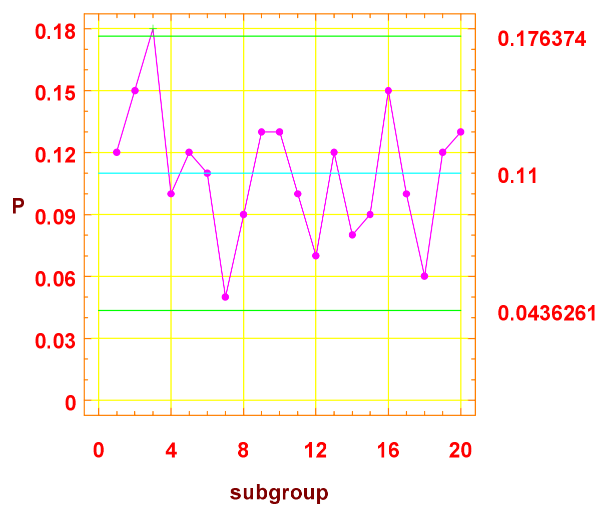


There are no points beyond the control limits. The process is in control.

b)

```
                                P Chart - Initial Study
                                Sample Size, n = 200
P Chart
-----
UCL:  +    3.0 sigma = 0.176374
Centerline           = 0.11
LCL:  -    3.0 sigma = 0.0436261
out of limits = 1
Estimated
mean P  = 0.11
sigma   = 0.0221246
```

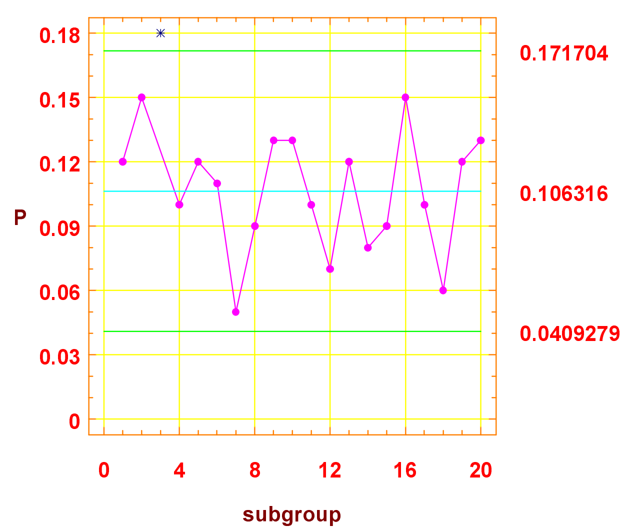
Problem 15-42b
Sample Size n = 200



There is one point beyond the upper control limit. The process is out of control.
The revised limits are:

```
                         P Chart - Revised Limits
                           Sample Size, n = 200
P Chart
-----
UCL:  +    3.0 sigma = 0.171704
Centerline           = 0.106316
LCL:  -    3.0 sigma = 0.0409279
out of limits = 0
Estimated
mean P   = 0.106316
sigma    = 0.021796
```

Problem 15-42b Revised Limits
Sample Size, n=200

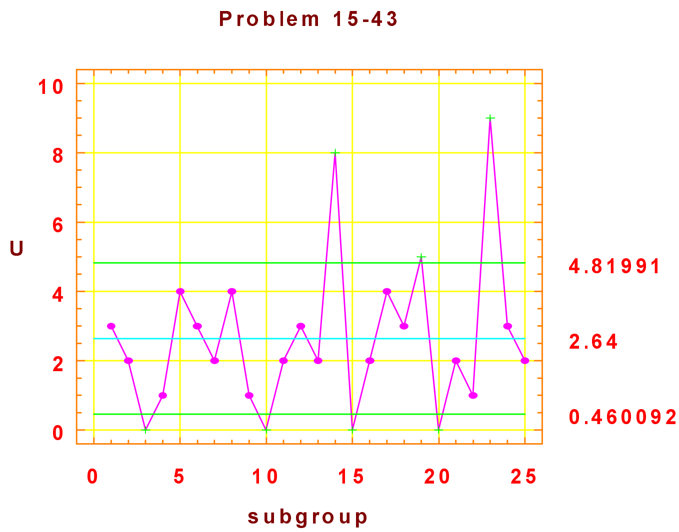


There are no points beyond the control limits. The process is now in control.

c) A larger sample size with the same number of defective items will result in more narrow control limits. The control limits corresponding to the larger sample size are more sensitive.

15-43. a)

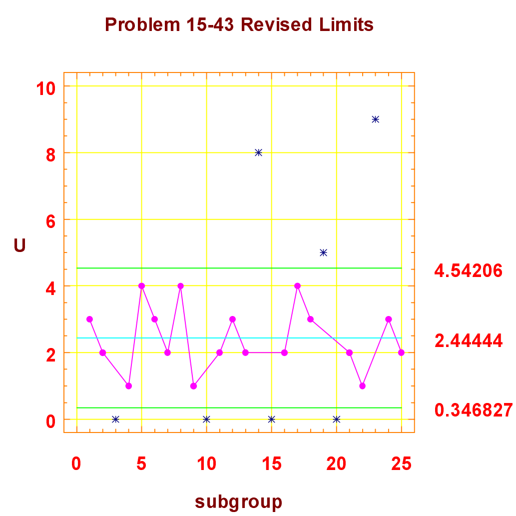
```
                                  U Chart - Initial Study
U Chart
-----
UCL:  +    3.0 sigma = 4.81991
Centerline           = 2.64
LCL:  -    3.0 sigma = 0.460092
out of limits = 7
Estimated
mean  U = 2.64
sigma   = 0.726636
```



There are points beyond the control limits. The process is out of control. The points beyond the control limits are 3, 10, 14, 15, 19, 20, and 23.

b)

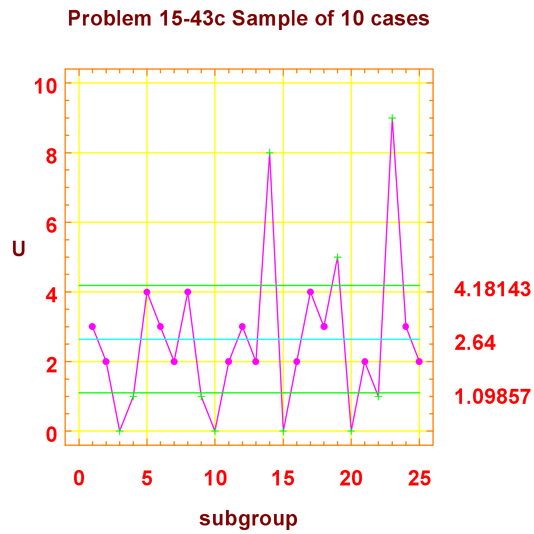
```
                                  U Chart - Revised Limits
U Chart
-----
UCL:  +    3.0 sigma = 4.54206
Centerline           = 2.44444
LCL:  -    3.0 sigma = 0.346827
out of limits = 0
Estimated
mean  U = 2.44444
sigma   = 0.699206
```



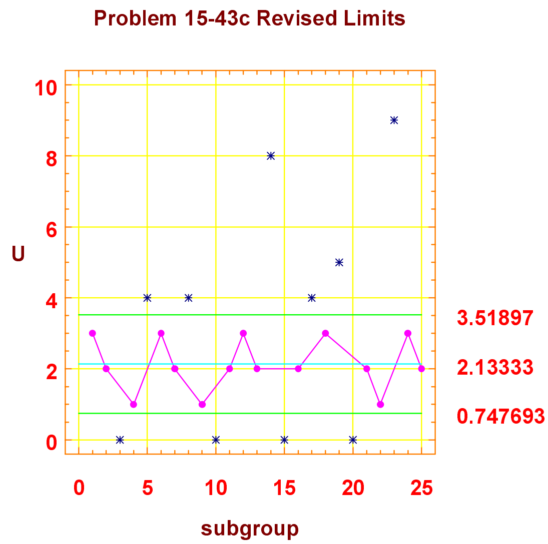
There are no points beyond the limits. The process is now in control.

c)

```
                            U Chart - Initial Study
                             Samples of 10 cases
U Chart
-----
UCL:  +    3.0 sigma = 4.18143
Centerline           = 2.64
LCL:  -    3.0 sigma = 1.09857
out of limits = 10
Estimated
mean  U = 2.64
sigma   = 0.513809
```



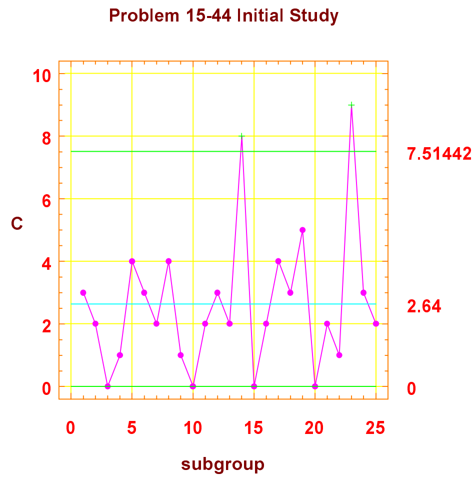

```
                            U Chart - Revised Limits
U Chart
-----
UCL:  +    3.0 sigma = 3.51897
Centerline           = 2.13333
LCL:  -    3.0 sigma = 0.747693
out of limits = 0
Estimated
mean  U = 2.13333
sigma   = 0.46188
```



There are no points beyond the control limits. The process is now in control.Larger sample size narrows the control limits and causes more sample observations to be deemed out-of-control.

15-44. a)

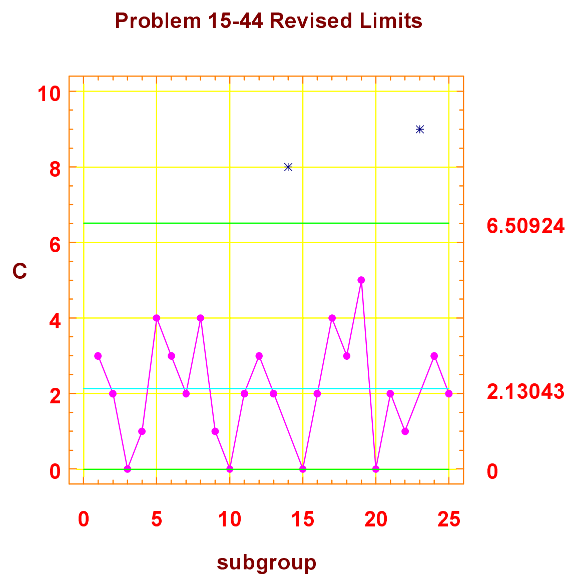
```
                                C Chart - Initial Study
C Chart
-----
UCL:  +    3.0 sigma = 7.51442
Centerline           = 2.64
LCL:  -    3.0 sigma = 0
out of limits = 2
Estimated
mean  C = 2.64
sigma   = 1.62481
```



There are points beyond the control limits. The process is out of control. The points are 14 and 23.

b)

```
                                C Chart - Revised Limits
C Chart
-----
UCL:  +    3.0 sigma = 6.50924
Centerline           = 2.13043
LCL:  -    3.0 sigma = 0
out of limits = 0
Estimated
mean  C = 2.13043
sigma   = 1.4596
```



There are no points beyond the control limits. The process is in control.

15-44. a) Let p denote the probability that a point plots outside of the control limits when the mean has shifted from μ_0 to $\mu = \mu_0 + 1.5\sigma$. Then,

$$P(LCL < \overline{X} < UCL)$$
$$= P\left(\mu_0 - \frac{3\sigma}{\sqrt{n}} < \overline{X} < \mu_0 + \frac{3\sigma}{\sqrt{n}}\right)$$
$$= P\left(\frac{-1.5\sigma}{\sigma/\sqrt{n}} - 3 < \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} < \frac{1.5\sigma}{\sigma/\sqrt{n}} + 3\right)$$
$$= P(-6 < Z < 0)$$
$$= P(Z < 0) - P(Z < -6)$$
$$= P(Z < 0) - [1 - P(Z < 6)]$$
$$= 0.5 - [1 - 1]$$
$$= 0.5$$

Therefore, the probability the shift is undetected for three consecutive samples is $(1-p)^3 = (0.5)^3 = 0.125$.

b) If 2-sigma control limits were used, then

$$1 - p = P(LCL < \overline{X} < UCL)$$
$$= P\left(\mu_0 - \frac{2\sigma}{\sqrt{n}} < \overline{X} < \mu_0 + \frac{2\sigma}{\sqrt{n}}\right)$$
$$= P\left(\frac{-1.5\sigma}{\sigma/\sqrt{n}} - 2 < \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} < \frac{1.5\sigma}{\sigma/\sqrt{n}} + 2\right)$$
$$= P(-5 < Z < -1)$$
$$= P(Z < -1) - P(Z < -5)$$
$$= 1 - P(Z < 1) - [1 - P(Z < 5)]$$
$$= 1 - 0.84134 - [1 - 1]$$
$$= 0.15866$$

Therefore, the probability the shift is undetected for three consecutive samples is $(1-p)^3 = (0.15866)^3 = 0.004$.

c) The 2-sigma limits are more narrow than the 3-sigma limits. Since the 2-sigma limits have a smaller probability of a shift being undetected, it would be better than the 3-sigma limits for a mean shift of 1.5σ.

15-46. ARL = 1/p where p is the probability a point falls outside the control limits.

a) $\mu = \mu_0 + \sigma$ and $n = 1$

$$p = P(\overline{X} > UCL) + P(\overline{X} < LCL)$$
$$= P\left(Z > \frac{\mu_0 + \frac{3\sigma}{\sqrt{n}} - \mu_0 - \sigma}{\sigma/\sqrt{n}}\right) + P\left(Z < \frac{\mu_0 - \frac{3\sigma}{\sqrt{n}} - \mu_0 - \sigma}{\sigma/\sqrt{n}}\right)$$
$$= P(Z > 3 - \sqrt{n}) + P(Z < -3 - \sqrt{n})$$
$$= P(Z > 2) + P(Z < -4) \qquad \text{when } n = 1$$
$$= 1 - P(Z < 2) + [1 - P(Z < 4)]$$
$$= 1 - 0.97725 + [1 - 1]$$
$$= 0.02275$$

Therefore, ARL = 1/p = 1/0.02275 = 43.9.

b) $\mu = \mu_0 + 2\sigma$

$P(\overline{X} > UCL) + P(\overline{X} < LCL)$

$$= P\left(Z > \frac{\mu_0 + \frac{3\sigma}{\sqrt{n}} - \mu_0 - 2\sigma}{\sigma/\sqrt{n}}\right) + P\left(Z < \frac{\mu_0 - \frac{3\sigma}{\sqrt{n}} - \mu_0 - 2\sigma}{\sigma/\sqrt{n}}\right)$$

$= P(Z > 3 - 2\sqrt{n}) + P(Z < -3 - 2\sqrt{n})$

$= P(Z > 1) + P(Z < -5)$ when $n = 1$

$= 1 - P(Z < 1) + [1 - P(Z < 5)]$

$= 1 - 0.84134 + [1 - 1]$

$= 0.15866$

Therefore, ARL = 1/p = 1/0.15866 = 6.30.

c) $\mu = \mu_0 + 3\sigma$

$P(\overline{X} > UCL) + P(\overline{X} < LCL)$

$$= P\left(Z > \frac{\mu_0 + \frac{3\sigma}{\sqrt{n}} - \mu_0 - 3\sigma}{\sigma/\sqrt{n}}\right) + P\left(Z < \frac{\mu_0 - \frac{3\sigma}{\sqrt{n}} - \mu_0 - 3\sigma}{\sigma/\sqrt{n}}\right)$$

$= P(Z > 3 - 3\sqrt{n}) + P(Z < -3 - 3\sqrt{n})$

$= P(Z > 0) + P(Z < -6)$ when $n = 1$

$= 1 - P(Z < 0) + [1 - P(Z < 6)]$

$= 1 - 0.50 + [1 - 1]$

$= 0.50$

Therefore, ARL = 1/p = 1/0.50 = 2.00.

d) The ARL is decreasing as the magnitude of the shift increases from σ to 2σ to 3σ. The ARL will decrease as the magnitude of the shift increases since a larger shift is more likely to be detected earlier than a smaller shift.

15-47. a) Because ARL = 370, on the average we expect there to be one false alarm every 370 hours. Each 30-day month contains 30 × 24 = 720 hours of operation. Consequently, we expect 720/370 = 1.9 false alarms each month.

b) The 2-sigma limits do reduce the ARL for detecting a shift in the mean of magnitude σ since the limits are more narrow. The number of false alarms has increased using 2-sigma limits.

c) With 2-sigma limits the probability of a point plotting out of control is determined as follows, when $\mu = \mu_0 + \sigma$

$P(X > UCL) + P(X < LCL)$

$$= P\left(\frac{X - \mu_0 - \sigma}{\sigma} > \frac{\mu_0 + 2\sigma - \mu_0 - \sigma}{\sigma}\right) + P\left(\frac{X - \mu_0 - \sigma}{\sigma} < \frac{\mu_0 - 2\sigma - \mu_0 - \sigma}{\sigma}\right)$$

$= P(Z > 1) + P(Z < -3)$

$= 1 - P(Z < 1) + [1 - P(Z < 3)]$

$= 1 - 0.84134 + 1 - 0.99865$

$= 0.160$

Therefore, ARL=1/p = 1/0.160 = 6.25. The 2-sigma limits do reduce the ARL for detecting a shift in the mean of magnitude σ. The number of false alarms has increased using 2-sigma limits.

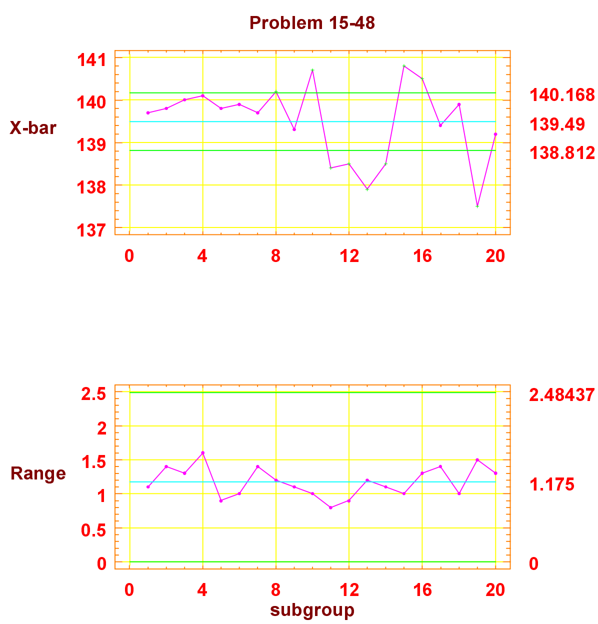
d) The in-control ARL = 1/p, where

$$p = P(X > UCL|\mu = \mu_0) + P(X < LCL|\mu = \mu_0)$$
$$= P\left(\frac{X-\mu_0}{\sigma} > \frac{\mu_0 + 2\sigma - \mu_0}{\sigma}\right) + P\left(\frac{X-\mu_0}{\sigma} < \frac{\mu_0 - 2\sigma - \mu_0}{\sigma}\right)$$
$$= P(Z > 2) + P(Z < -2)$$
$$= 1 - P(Z < 2) + [1 - P(Z < 2)]$$
$$= 2(1 - 0.97725)$$
$$= 0.0455$$

Therefore, ARL = 1/0.0455 = 21.98. The number of false alarms per month is 720/21.98 = 32.76. This is an excessive number of false alarms (more than one per day) and 2-sigma limits are not recommended for routine production. Thus, this in-control ARL performance is probably not satisfactory.

15-48. a)

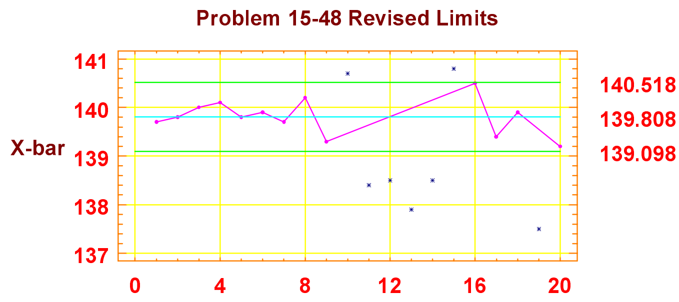
```
                              X-bar and Range - Initial Study
                                     Charting xbar
X-bar                                      | Range
-----                                      | -----
UCL:   +     3.0 sigma = 140.168           | UCL:   +     3.0 sigma = 2.48437
Centerline             = 139.49            | Centerline             = 1.175
LCL:   -     3.0 sigma = 138.812           | LCL:   -     3.0 sigma = 0
out of limits = 9                          | out of limits = 0
Estimated
process mean  = 139.49
process sigma =    0.505159
mean Range    =    1.175
```

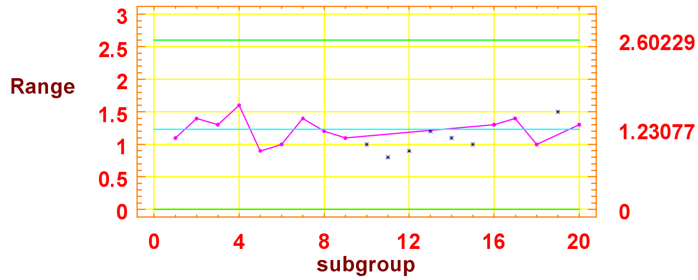


There are points beyond the control limits. The process is out of control. The points are 4, 8, 10, 13, 15, 16, and 19.

b) Revised control limits are given in the table below:

```
                              X-bar and Range - Initial Study
                                     Charting Xbar
X-bar                                      | Range
-----                                      | -----
UCL:   +     3.0 sigma = 140.518           | UCL:   +     3.0 sigma = 2.60229
Centerline             = 139.808           | Centerline             = 1.23077
LCL:   -     3.0 sigma = 139.098           | LCL:   -     3.0 sigma = 0
out of limits = 0                          | out of limits = 0
Estimated
process mean  = 139.808
process sigma =    0.529136
mean Range    =    1.23077
```





There are no points beyond the control limits the process is now in control.

The process standard deviation estimate is given by $\hat{\sigma} = \frac{\overline{R}}{d_2} = \frac{1.23077}{2.326} = 0.529$

c) $PCR = \frac{USL - LSL}{6\hat{\sigma}} = \frac{142 - 138}{6(0.529)} = 1.26$

$$PCR_k = \min\left[\frac{USL - \overline{\overline{x}}}{3\hat{\sigma}}, \frac{\overline{\overline{x}} - LSL}{3\hat{\sigma}}\right]$$

$$= \min\left[\frac{142 - 139.808}{3(0.529)}, \frac{139.808 - 138}{3(0.529)}\right]$$

$$= \min[1.38, 1.14]$$

$$= 1.14$$

Since PCR exceeds unity, the natural tolerance limits lie inside the specification limits and very few defective units will be produced.

PCR is slightly larger than PCR_k indicating that the process is somewhat off center.

d) In order to make this process a "six-sigma process", the variance σ^2 would have to be decreased such that $PCR_k = 2.0$. The value of the variance is found by solving $PCR_k = \frac{\overline{\overline{x}} - LSL}{3\sigma} = 2.0$ for σ:

$$\frac{139.808 - 138}{3\sigma} = 2.0$$

$$6\sigma = 139.808 - 138$$

$$\sigma = \frac{139.808 - 138}{6}$$

$$\sigma = 0.3013$$

Therefore, the process variance would have to be decreased to $\sigma^2 = (0.3013)^2 = 0.091$.

e) $\hat{\sigma}_{\overline{X}} = 0.529$

$p = P(139.098 < X < 140.518 | \mu = 139.7)$

$= P\left(\frac{139.098 - 139.7}{0.529} < \frac{X - \mu}{\sigma_X} < \frac{140.518 - 139.7}{0.529}\right)$

$= P(-1.14 < Z < 1.55)$

$= P(Z < 1.55) - P(Z < -1.14)$

$= P(Z < 1.55) - [1 - P(Z < 1.14)]$

$= 0.93943 - [1 - 0.87285]$

$= 0.8123$

The probability that this shift will be detected on the next sample is 1–p = 1–0.8123 = 0.1877.

$ARL = \frac{1}{1-p} = \frac{1}{0.1877} = 5.33$

15-49. a) The $P(LCL < \hat{P} < UCL)$, when p = 0.08, is needed.

$LCL = \overline{p} - 3\sqrt{\frac{\overline{p}(1-\overline{p})}{n}} = 0.05 - 3\sqrt{\frac{0.05(1-0.05)}{100}} = -0.015 \rightarrow 0$

$UCL = \overline{p} + 3\sqrt{\frac{\overline{p}(1-\overline{p})}{n}} = 0.05 + 3\sqrt{\frac{0.05(1-0.05)}{100}} = 0.115$

Therefore, when p = 0.08

$P(0 \le \hat{P} \le 0.115) = P(\hat{P} \le 0.115)$

$= P\left(\frac{\hat{P} - 0.08}{\sqrt{\frac{0.08(0.92)}{100}}} \le \frac{0.115 - 0.08}{\sqrt{\frac{0.08(0.92)}{100}}}\right)$

$= P(Z \le 1.29)$

$= 0.90$

Using the normal approximation to the distribution of $\hat{P}$. Therefore, the probability of detecting the shift on the first sample following the shift is 1 – 0.90 = 0.10.

b) The probability that the control chart detects a shift to 0.08 on the second shift is 0.10(0.90) = 0.09.

c) p = 0.10

$P(0 \le \hat{P} \le 0.115) = P(\hat{P} \le 0.115)$

$= P\left(\frac{\hat{P} - 0.10}{\sqrt{\frac{0.10(0.90)}{100}}} \le \frac{0.115 - 0.10}{\sqrt{\frac{0.10(0.90)}{100}}}\right)$

$= P(Z \le 0.5)$

$= 0.69146$

using the normal approximation to the distribution of $\hat{P}$. Therefore, the probability of detecting the shift on the first sample following the shift is 1 – 0.69146 = 0.30854.

The probability that the control chart detects a shift to 0.10 on the second shift is 0.69146(0.30854) = 0.2133.

d) A larger shift is generally easier to detect. Therefore, we should expect a shift to 0.10 to be detected quicker than a shift to 0.08.

15-50. $\bar{u} = 8$

a) n = 4

$$UCL = \bar{u} + 3\sqrt{\frac{\bar{u}}{n}} = 8 + 3\sqrt{\frac{8}{4}} = 12.24$$

$$LCL = \bar{u} - 3\sqrt{\frac{\bar{u}}{n}} = 8 - 3\sqrt{\frac{8}{4}} = 3.78$$

$$P(\overline{U} > 12.24 \text{ when } \lambda = 16) = P\left(Z > \frac{12.24 - 16}{\sqrt{16/4}}\right)$$

$= P(Z > -1.88)$
$= 1 - P(Z < -1.88)$
$= 1 - 0.03005$
$= 0.96995$

$$P(\overline{U} < 3.78) = P\left(Z < \frac{3.78 - 16}{\sqrt{16/4}}\right)$$

$= P(Z < -6.11)$
$= 0$

b) n = 10

$$UCL = \bar{u} + 3\sqrt{\frac{\bar{u}}{n}} = 8 + 3\sqrt{\frac{8}{10}} = 10.68$$

$$LCL = \bar{u} - 3\sqrt{\frac{\bar{u}}{n}} = 8 - 3\sqrt{\frac{8}{10}} = 5.32$$

$$P(U > 10.68 \text{ when } \lambda = 16) = P\left(Z > \frac{10.68 - 16}{\sqrt{\frac{16}{10}}}\right)$$

$= P(Z > -4.22)$
$= 1$

15-51. $\bar{u} = 10$

a) n = 1

$$UCL = \bar{u} + 3\sqrt{\frac{\bar{u}}{n}} = 10 + 3\sqrt{\frac{10}{1}} = 19.49$$

$$LCL = \bar{u} - 3\sqrt{\frac{\bar{u}}{n}} = 10 - 3\sqrt{\frac{10}{1}} = 0.51$$

$$P(\overline{U} > 19.94 \text{ when } \lambda = 14) = P\left(Z > \frac{19.94 - 14}{\sqrt{14}}\right)$$

$= P(Z > 1.37)$
$= 1 - P(Z < 1.37)$
$= 1 - 0.91465$
$= 0.085$

and

$$P(\overline{U} < 0.51) = P\left(Z < \frac{0.51 - 14}{\sqrt{14}}\right) = 0$$

b) n = 5

$$UCL = \bar{u} + 3\sqrt{\frac{\bar{u}}{n}} = 10 + 3\sqrt{\frac{10}{5}} = 14.24$$

$$LCL = \bar{u} - 3\sqrt{\frac{\bar{u}}{n}} = 10 - 3\sqrt{\frac{10}{5}} = 5.76$$

$$P(\overline{U} > 14.24 \text{ when } \lambda = 14) = P\left(Z > \frac{14.24 - 14}{\sqrt{\frac{14}{5}}}\right)$$

= P(Z >0.14)
= 1 – 0.55567
= 0.444

$$P(\overline{U} < 5.76 \text{ when } \lambda = 14) = P\left(Z < \frac{5.76 - 14}{\sqrt{\frac{14}{5}}}\right)$$

= P(Z < –4.93)
= 0

15-52.

$CL = \mu$

$$UCL = \mu + 2\frac{\sigma}{\sqrt{n}}$$

$$LCL = \mu - 2\frac{\sigma}{\sqrt{n}}$$

$$P\left(\overline{X} > \mu + 2\frac{\sigma}{\sqrt{n}}\right) = P\left(\frac{\overline{X} - \mu}{\sigma / \sqrt{n}} > 2\right)$$

$= P(Z > 2)$
$= 1 - P(Z < 2)$
$= 1 - 0.97725$
$= 0.02275$

and

$$P\left(\overline{X} < \mu - 2\frac{\sigma}{\sqrt{n}}\right) = P\left(\frac{\overline{X} - \mu}{\sigma / \sqrt{n}} < -2\right)$$

$= P(Z < -2)$
$= 1 - P(Z < 2)$
$= 1 - 0.97725$
$= 0.02275$

The answer is 0.02275 + 0.02275 = 0.0455.
The answer for 3-sigma control limits is 0.0027. The 3-sigma control limits result in much fewer false alarms.

15-53.

$CL = \mu$

$UCL = \mu + k\dfrac{\sigma}{\sqrt{n}}$

$LCL = \mu - k\dfrac{\sigma}{\sqrt{n}}$

$$P\left(\overline{X} > \mu + k\frac{\sigma}{\sqrt{n}}\right) = P\left(\frac{\overline{X} - \mu - \delta}{\sigma/\sqrt{n}} > k - \frac{\delta}{\sigma/\sqrt{n}}\right)$$
$$= P\left(Z > k - \frac{\delta}{\sigma/\sqrt{n}}\right)$$
$$= 1 - \Phi\left(k - \frac{\delta}{\sigma/\sqrt{n}}\right)$$

$$P\left(\overline{X} < \mu + k\frac{\sigma}{\sqrt{n}}\right) = P\left(\frac{\overline{X} - \mu - \delta}{\sigma/\sqrt{n}} < -k - \frac{\delta}{\sigma/\sqrt{n}}\right)$$
$$= P\left(Z < -k - \frac{\delta}{\sigma/\sqrt{n}}\right)$$
$$= \Phi\left(-k - \frac{\delta}{\sigma/\sqrt{n}}\right)$$

The answer is

$$1 - \Phi\left(k - \frac{\delta\sqrt{n}}{\sigma}\right) + \Phi\left(-k - \frac{\delta\sqrt{n}}{\sigma}\right).$$

15-54. From Exercise 15-52, p = 0.0455.

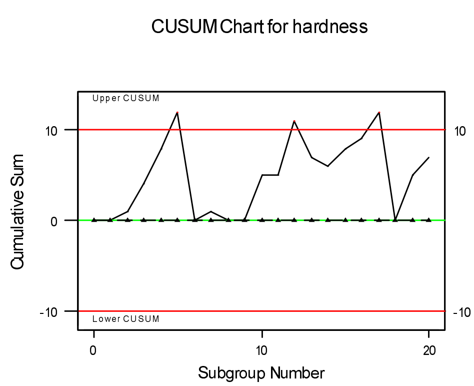
a) The probability of a false alarm is the probability that $\overline{X}$ is outside the control limits when there is no shift, i.e., 0.0455. The probability there is not a false alarm on the first sample is 1 – 0.0455 = 0.9545. Hence, the probability of a false alarm on the second sample but not on the first (assuming the two samples are independent) is (0.9545)(0.0455) = 0.04343.

b) Assuming all the samples are independent, the probability that there is not a false alarm in the first three samples is $(0.9545)^3 = 0.86962$.

15-55. PCR = 2 but $\mu = USL + 3\sigma$

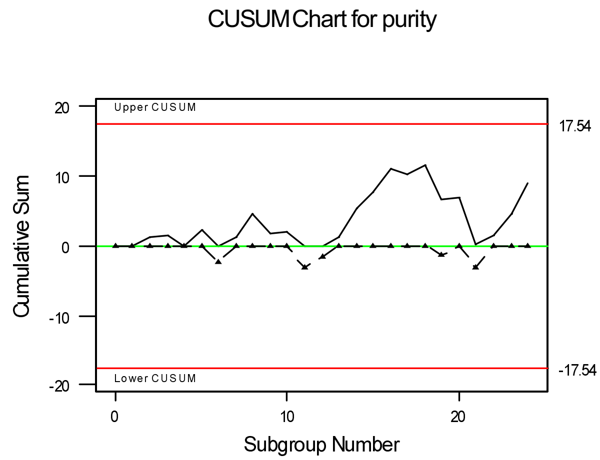
$$P(X < USL) = P\left(Z < \frac{(\mu - 3\sigma) - \mu}{\sigma}\right) = P(Z < -3) = 0.00135$$

15-56.



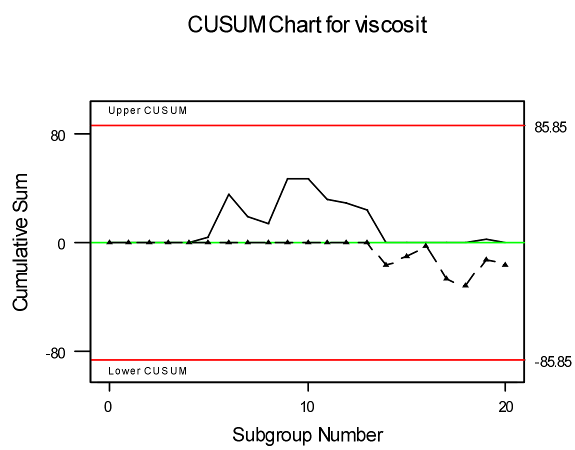
The process is not in control.

15-57.



Process standard deviation is estimated using the average moving range of size 2: MR/d_2, where d_2 = 1.128 for a moving range of 2. The estimate is 3.508.

15-58.



Process standard deviation is estimated using the average moving range of size 2: MR/d_2, where d_2 = 1.128 for a moving range of 2. The estimate is 17.17.

Mind-Expanding Exercises

15-59. Let p denote the probability that a point plots outside of the control limits when the mean has shifted from μ_0 to $\mu = \mu_0 + 1.5\sigma$. Then:

$$1 - p = P\left(LCL < \overline{X} < UCL \mid \mu = \mu_0 + 1.5\sigma\right) = P\left(\mu_0 - 3\frac{\sigma}{\sqrt{n}} < \overline{X} < \mu_0 - 3\frac{\sigma}{\sqrt{n}} \mid \mu = \mu_0 + 1.5\sigma\right)$$

$$= P\left(\frac{-1.5\sigma}{\sigma/\sqrt{n}} - 3 < \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} < \frac{-1.5\sigma}{\sigma/\sqrt{n}}\right)$$

$$= P(-6 < Z < 0) = 0.5$$

Therefore, the probability the shift is undetected for three consecutive samples is $(1\text{-}p)^3 = 0.5^3 = 1/8$.

If 2-sigma control limits were used, then

$$1-p = P\left(LCL < \overline{X} < UCL \mid \mu = \mu_0 + 1.5\sigma\right) = P\left(\mu_0 - 2\frac{\sigma}{\sqrt{n}} < \overline{X} < \mu_0 - 2\frac{\sigma}{\sqrt{n}} \mid \mu = \mu_0 + 1.5\sigma\right)$$

$$= P\left(\frac{-1.5\sigma}{\sigma/\sqrt{n}} - 2 < \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} < \frac{-1.5\sigma}{\sigma/\sqrt{n}}\right)$$

$$= P(-5 < Z < -1) = 1 - 0.84134 = 0.15866$$

Therefore, the probability the shift is undetected for three consecutive samples is $(1\text{-}p)^3 = 0.15866^3 = 0.004$.

15-60.

$$LCL = \mu_0 - k\sigma/\sqrt{n}$$

$$CL = \mu_0$$

$$UCL = \mu_0 + k\sigma/\sqrt{n}$$

a) ARL = 1/p where p is the probability a point plots outside of the control limits. Then,

$$1-p = P\left(LCL < \overline{X} < UCL \mid \mu_0\right) = P\left(-k < \frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}} < k \mid \mu_0\right)$$

$$= P(-k < Z < k) = \Phi(k) - \Phi(-k) = 2\Phi(k) - 1$$

where $\Phi(Z)$ is the standard normal cumulative distribution function. Therefore, $p = 2 - 2\Phi(k)$ and $ARL = 1/[2-2\Phi(k)]$. The mean time until a false alarm is k/p hours.

b) ARL=1/p where

$$1-p = P\left(LCL < \overline{X} < UCL \mid \mu_1 = \mu_0 + \delta\sigma\right) = P\left(-k - \frac{\delta\sigma}{\sigma/\sqrt{n}} < \overline{X} < k - \frac{\delta\sigma}{\sigma/\sqrt{n}}\right)$$

$$= P\left(-k - \sqrt{n}\delta < Z < k - \sqrt{n}\delta\right)$$

$$= \Phi(k - \sqrt{n}\delta) - \Phi(-k - \sqrt{n}\delta)$$

and

$$p = 1 - \Phi(k - \sqrt{n}\delta) + \Phi(-k - \sqrt{n}\delta)$$

c) ARL=1/p where 1-p = $P(-3<Z<3)= 0.9973$. Thus, ARL = 1/0.0027=370.4. If k = 2, 1-p = $P(-2<Z<2)=0.9545$ and ARL=1/p = 22.0.
The 2-sigma limits result in a false alarm for every 22 points on the average. This is a high number of false alarms for the routine use of a control chart.

d) From part b), ARL = 1/p, where 1-p = $P(-3-\sqrt{5} < Z < 3-\sqrt{5}) = 0.7764$. ARL = 4.47 assuming 3-sigma control limits.

15-61. Determine n such that

$$0.5 = P(LCL < \hat{P} < UCL \mid p = p_c)$$

$$= P\left(\frac{\overline{p} - k\sqrt{\frac{\overline{p}(1-\overline{p})}{n}} - p_c}{\sqrt{\frac{p_c(1-p_c)}{n}}} < \frac{\hat{P} - p_c}{\sqrt{\frac{p_c(1-p_c)}{n}}} < \frac{\overline{p} + k\sqrt{\frac{\overline{p}(1-\overline{p})}{n}} - p_c}{\sqrt{\frac{p_c(1-p_c)}{n}}} \mid p = p_c\right)$$

$$= P\left(\frac{\overline{p}-p_c}{\sqrt{\frac{p_c(1-p_c)}{n}}} - \frac{k\sqrt{\overline{p}(1-\overline{p})}}{\sqrt{p_c(1-p_c)}} < Z < \frac{\overline{p}-p_c}{\sqrt{\frac{p_c(1-p_c)}{n}}} - \frac{k\sqrt{\overline{p}(1-\overline{p})}}{\sqrt{p_c(1-p_c)}}\right)$$

Use the fact that if $p_c > \overline{p}$ then the probability is approximately equal to the probability that Z is less than the upper limit. Then, the probability above approximately equals 0.5 if

$$\frac{\overline{p}-p_c}{\sqrt{\frac{p_c(1-p_c)}{n}}} = \frac{k\sqrt{\overline{p}(1-\overline{p})}}{\sqrt{p_c(1-p_c)}}$$

Solving for n, $n = \dfrac{k^2\overline{p}(1-\overline{p})}{(\overline{p}-p_c)^2}$

15-62. The LCL is $\overline{p} - \frac{k\sqrt{\overline{p}(1-\overline{p})}}{n}$

$\overline{p} - \frac{k\sqrt{\overline{p}(1-\overline{p})}}{n}$ =0 or $n = \frac{k^2(1-\overline{p})}{\overline{p}}$

15-63. The $P(LCL < \hat{P} < UCL | p = 0.08)$ is desired. Now, using the normal approximation:

$$LCL = \overline{p} - 3\sqrt{\frac{\overline{p}(1-\overline{p})}{n}} = 0.05 - 3\sqrt{\frac{0.05(0.95)}{100}} = -0.015 \to 0$$

$$UCL = \overline{p} + 3\sqrt{\frac{\overline{p}(1-\overline{p})}{n}} = 0.05 + 3\sqrt{\frac{0.05(0.95)}{100}} = 0.115$$

$$P(0 < \hat{P} < 0.115 | p = 0.08) = P(\hat{P} < 0.115 | p = 0.08)$$

$$= P\left(\frac{\hat{P} - 0.08}{\sqrt{\frac{0.08(1-0.08)}{100}}} < \frac{0.115 - 0.08}{\sqrt{\frac{0.08(1-0.08)}{100}}}\right) = P(Z < 1.29) = 0.90$$

Therefore, the probability of detecting shift on the first sample following the shift is 1 − 0.90 = 0.10. The probability of detecting a shift by at least the third sample following the shift can be determined from the geometric distribution to be 0.10 + 0.90(0.10) +0.90^2(0.10) = 0.27

15-64. The process should be centered at the middle of the specifications; that is, at 100.
For an $\overline{x}$ chart:

$$LCL = \mu_0 - k\sigma / \sqrt{n} = 100 + 3(5)/2 = 107.5$$

$$CL = \mu_0 = 100$$

$$UCL = \mu_0 + k\sigma / \sqrt{n} = 100 - 3(5)/2 = 92.5$$

$$P(LCL < \overline{X} < UCL | \mu = 105) = P\left(\frac{92.5 - 105}{5/2} < \frac{\overline{X} - 105}{5/2} < \frac{107.5 - 105}{5/2}\right)$$

$$= P(-5 < Z < 1) = 0.84$$

The requested probability is then 1 – 0.84 = 0.16. The ARL = 1/0.16 = 6.25. With μ = 105, the specifications at 100±15 and σ = 5, the probability of a defective item is

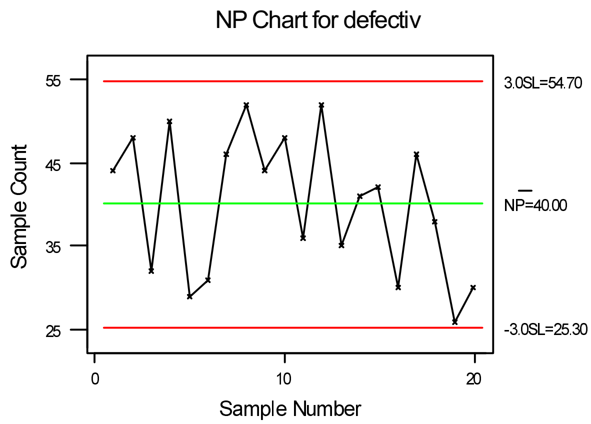
$$P(X < 85) + P(X > 115) = P\left(\frac{X - 105}{5} < \frac{85 - 105}{5}\right) + P\left(\frac{X - 105}{5} > \frac{115 - 105}{5}\right)$$

$$= P(Z < -3) + P(Z > 2) = 0.0241$$

Therefore, the average number of observations until a defective occurs, follows from the geometric distribution as 1/0.0241 = 41.49. However, the $\overline{x}$ chart only requires 6.25 samples of 4 observations each = 6.25(4) = 25 observations, on average, to detect the shift.

15-65. Let X denote the number of defectives in a sample of n. Then X has a binomial distribution with E(X) = np and V(X) = np(1-p). Therefore, the estimates of the mean and standard deviation of X are $n\overline{p}$ and $\sqrt{n\overline{p}(1-\overline{p})}$, respectively. Using these estimates results in the given control limits.

b)



c) Note that

$$n\bar{p} - 3\sqrt{n\bar{p}(1-\bar{p})} < n\hat{P} < n\bar{p} + 3\sqrt{n\bar{p}(1-\bar{p})}$$

therefore,

$$\bar{p} - 3\sqrt{\frac{\bar{p}(1-\bar{p})}{n}} < \hat{P} < \bar{p} + 3\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$$

Therefore, the nP control chart will always provide results equivalent to the P chart.

15-66. a)

$$z_0 = \mu_0$$
$$z_1 = \lambda\bar{x}_1 + (1-\lambda)\mu_0$$
$$z_2 = \lambda\bar{x}_2 + (1-\lambda)z_0 = \lambda\bar{x}_2 + (1-\lambda)\lambda\bar{x}_1 + (1-\lambda)^2\mu_0$$

in general,

$$z_t = \lambda\sum_{j=0}^{t-1}(1-\lambda)^j\bar{x}_{t-j} + (1-\lambda)^t\mu_0$$

Therefore,

$$E(Z_t) = \lambda\sum_{j=0}^{t-1}(1-\lambda)^j\bar{x}_{t-j} + (1-\lambda)^t\mu_0$$
$$= \frac{\lambda[(1-\lambda)^t - 1]}{(1-\lambda)-1}\mu_0 + (1-\lambda)^t\mu_0$$
$$= \mu_0$$

b)

$$V(Z_t) = \lambda^2\sum_{j=0}^{t-1}(1-\lambda)^{2j}V(\bar{x}_{t-j}) = \lambda^2\sum_{j=0}^{t-1}(1-\lambda)^{2j}\frac{\sigma^2}{n}$$
$$= \frac{\lambda^2[(1-\lambda)^{2t} - 1]}{(1-\lambda)^2 - 1}\frac{\sigma^2}{n}$$
$$= \frac{\lambda}{2-\lambda}[1-(1-\lambda)^{2t}]\frac{\sigma^2}{n}$$

c) The control limits are

$$LCL = \mu_0 - 3\frac{\sigma}{\sqrt{n}}\sqrt{\frac{\lambda}{2-\lambda}}\sqrt{1-(1-\lambda)^{2t}}$$

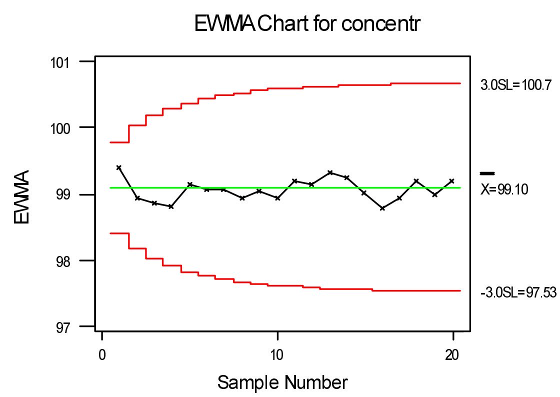
$$CL = \mu_0$$

$$UCL = \mu_0 + 3\frac{\sigma}{\sqrt{n}}\sqrt{\frac{\lambda}{2-\lambda}}\sqrt{1-(1-\lambda)^{2t}}$$

d) As $\lambda \to 1$, $z_t \to \overline{x}_t$, therefore, the EWMA performs as a Shewhart control chart.

e) $\lambda \to 0$, the weight given to past observations decreases very slowly. Consequently, the EWMA applies approximately equal weight to previous observations and the control chart performance is similar to that of a CUSUM chart.

f) The centerline for the chart is chosen to be $\overline{x} = 99.1$ from Example 15-2. also, σ is estimated in Example 15-2 as 2.59/1.128 = 2.296. Finally, λ is chosen to be 0.1 in the EWMA control chart.

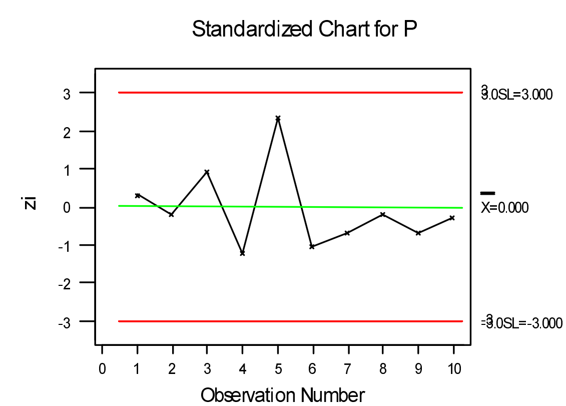


15-67. $-3 < Z_t < 3$ if and only if

$$-3 < \frac{\hat{P} - \overline{p}}{\sqrt{\frac{\overline{p}(1-\overline{p})}{n}}} < 3 \quad \text{or} \quad \overline{p} - 3\sqrt{\frac{\overline{p}(1-\overline{p})}{n}} < \hat{P}_i < \overline{p} + 3\sqrt{\frac{\overline{p}(1-\overline{p})}{n}}$$

Therefore, a point is in control on this chart if and only if the point is in control on the original p chart.

15-68. For unequal sample sizes, the standard control chart can be used with the value of n equal to the size of each sample. That is, $Z_i = \frac{\hat{P} - \overline{p}}{\sqrt{\frac{\overline{p}(1-\overline{p})}{n_i}}}$ where n_i is the size of the ith sample.



Appendix I

I-1. From the multiplication rule, the answer is $5 \times 3 \times 4 \times 2 = 120$

I-2. From the multiplication rule, $3 \times 4 \times 3 = 36$

I-3. From the multiplication rule, $3 \times 4 \times 3 \times 4 = 144$

I-4. From equation I-1, the answer is 10! = 3628800

I-5. From the multiplication rule and equation I-1, the answer is 5!5! = 14400

I-6. From equation I-3, $\frac{7!}{3!4!} = 35$ sequences are possible

I-7. a) From equation I-4, the number of samples of size five is $\binom{140}{5} = \frac{140!}{5!135!} = 416965528$

b) There are 10 ways of selecting one nonconforming chip and there are $\binom{130}{4} = \frac{130!}{4!126!} = 11358880$ ways of selecting four conforming chips. Therefore, the number of samples that contain exactly one nonconforming chip is $10 \times \binom{130}{4} = 113588800$

c) The number of samples that contain at least one nonconforming chip is the total number of samples $\binom{140}{5}$ minus the number of samples that contain no nonconforming chips $\binom{130}{5}$. That is $\binom{140}{5}$ - $\binom{130}{5}$

$= \frac{140!}{5!135!} - \frac{130!}{5!125!} = 130721752$

I-8. a) If the chips are of different types, then every arrangement of 5 locations selected from the 12 results in a different layout. Therefore, $P_5^{12} = \frac{12!}{7!} = 95040$ layouts are possible.

b) If the chips are of the same type, then every subset of 5 locations chosen from the 12 results in a different layout. Therefore, $\binom{12}{5} = \frac{12!}{5!7!} = 792$ layouts are possible.

I-9. a) $\frac{7!}{2!5!} = 21$ sequences are possible.

b) $\frac{7!}{1!1!1!1!1!2!} = 2520$ sequences are possible.

c) 6! = 720 sequences are possible.

I-10. a) Every arrangement of 7 locations selected from the 12 comprises a different design.
$P_7^{12} = \frac{12!}{5!} = 3991680$ designs are possible.

b) Every subset of 7 locations selected from the 12 comprises a new design. $\frac{12!}{5!7!} = 792$ designs are possible.

c) First the three locations for the first component are selected in $\binom{12}{3} = \frac{12!}{3!9!} = 220$ ways. Then, the four locations for the second component are selected from the nine remaining locations in $\binom{9}{4} = \frac{9!}{4!5!} = 126$ ways. From the multiplication rule, the number of designs is $220 \times 126 = 27720$

I-11. a) From the multiplication rule, $10^3 = 1000$ prefixes are possible

b) From the multiplication rule, $8 \times 2 \times 10 = 160$ are possible

c) Every arrangement of three digits selected from the 10 digits results in a possible prefix.

$P_3^{10} = \frac{10!}{7!} = 720$ prefixes are possible.

I-12. a) From the multiplication rule, $2^8 = 256$ bytes are possible

b) From the multiplication rule, $2^7 = 128$ bytes are possible

I-13. a) The total number of samples possible is $\binom{24}{4} = \frac{24!}{4!20!} = 10626$. The number of samples in which exactly one tank has high viscosity is $\binom{6}{1}\binom{18}{3} = \frac{6!}{1!5!} \times \frac{18!}{3!15!} = 4896$. Therefore, the probability is $\frac{4896}{10626} = 0.461$

b) The number of samples that contain no tank with high viscosity is $\binom{18}{4} = \frac{18!}{4!14!} = 3060$. Therefore, the requested probability is $1 - \frac{3060}{10626} = 0.712$.

c) The number of samples that meet the requirements is $\binom{6}{1}\binom{4}{1}\binom{14}{2} = \frac{6!}{1!5!} \times \frac{4!}{1!3!} \times \frac{14!}{2!12!} = 2184$.

Therefore, the probability is $\frac{2184}{10626} = 0.206$

I-14. a) The total number of samples is $\binom{12}{3} = \frac{12!}{3!9!} = 220$. The number of samples that result in one nonconforming part is $\binom{2}{1}\binom{10}{2} = \frac{2!}{1!1!} \times \frac{10!}{2!8!} = 90$. Therefore, the requested probability is 90/220 = 0.409.

b) The number of samples with no nonconforming part is $\binom{10}{3} = \frac{10!}{3!7!} = 120$. The probability of at least one nonconforming part is $1 - \frac{120}{220} = 0.455$.

I-15. a) The probability that both parts are defective is $\frac{5}{50} \times \frac{4}{49} = 0.0082$

b) The total number of samples is $\binom{50}{2} = \frac{50!}{2!48!} = \frac{50 \times 49}{2}$. The number of samples with two defective parts is $\binom{5}{2} = \frac{5!}{2!3!} = \frac{5 \times 4}{2}$. Therefore, the probability is $\frac{\frac{5 \times 4}{2}}{\frac{50 \times 49}{2}} = \frac{5 \times 4}{50 \times 49} = 0.0082$.

Appendix II

II-1. a) $E(e^{tX}) = \sum_{x=1}^{m} \frac{e^{tx}}{m} = \frac{1}{m}\sum_{x=1}^{m}(e^t)^x = \frac{(e^t)^{m+1} - e^t}{m(e^t - 1)} = \frac{e^t(1-e^{tm})}{m(1-e^t)}$

b) $M(t) = \frac{1}{m}e^t(1-e^{tm})(1-e^t)^{-1}$ and

$$\frac{dM(t)}{dt} = \frac{1}{m}\left\{e^t(1-e^{tm})(1-e^t)^{-1} + e^t(-me^{tm})(1-e^t)^{-1} + e^t(1-e^{tm})(-1)(1-e^t)^{-2}(-e^t)\right\}$$

$$\frac{dM(t)}{dt} = \frac{e^t}{m(1-e^t)}\left\{1 - e^{tm} - me^{tm} + \frac{(1-e^{tm})e^t}{1-e^t}\right\}$$

$$= \frac{e^t}{m(1-e^t)^2}\left\{1 - e^{tm} - me^{tm} + me^{(m+1)t}\right\}$$

Using L'Hospital's rule,

$$\lim_{t\to 0}\frac{dM(t)}{dt} = \lim_{t\to 0}\frac{e^t}{m}\lim_{t\to 0}\frac{-me^{tm} - m^2e^{tm} + m(m+1)e^{(m+1)t}}{-2(1-e^t)e^t}$$

$$= \lim_{t\to 0}\frac{e^t}{m}\lim_{t\to 0}\frac{-m^2e^{tm} - m^3e^{tm} + m(m+1)^2e^{(m+1)t}}{-2(1-e^t)e^t - 2e^{-t}(-e^t)}$$

$$= \frac{1}{m}\times\frac{m(m+1)^2 - m^2 - m^3}{2} = \frac{m^2+m}{2m} = \frac{m+1}{2}$$

Therefore, $E(X) = \frac{m+1}{2}$.

$$\frac{d^2M(t)}{dt^2} = \frac{e^t}{m(1-e^t)^2}\left\{-me^{tm} - m^2e^{tm} + m(m+1)e^{(m+1)t}\right\}$$

$$+\frac{e^t}{m(1-e^t)^2}\left\{1 - e^{tm} - me^{tm} + me^{(m+1)t}\right\}$$

$$+\frac{2e^{2t}}{m(1-e^t)^3}\left\{1 - e^{tm} - me^{tm} + me^{(m+1)t}\right\}$$

Using L'Hospital's rule,

$$\lim_{t\to 0}\frac{d^2M(t)}{dt^2} = \lim_{t\to 0}\frac{e^t}{m}\lim_{t\to 0}\frac{-m^2e^{tm} - m^3e^{tm} + m(m+1)^2e^{(m+1)t}}{2(1-e^t)(-e^t)}$$

$$+\frac{m+1}{2}\text{(from the derivation of E(X))}$$

$$+\lim_{t\to 0}\frac{2e^{2t}}{m}\lim_{t\to 0}\frac{-me^{tm} - m^2e^{tm} + m(m+1)e^{(m+1)t}}{3(1-e^t)^2(-e^t)}$$

$$= \lim_{t\to 0}\frac{e^t}{m}\lim_{t\to 0}\frac{-m^3e^{tm} - m^4e^{tm} + m(m+1)^3e^{(m+1)t}}{2(1-e^t)(-e^t) + 2(-e^t)(-e^t)} + \frac{m+1}{2}$$

$$+\lim_{t\to 0}\frac{2e^{2t}}{m}\lim_{t\to 0}\frac{-m^2e^{tm} - m^3e^{tm} + m(m+1)^2e^{(m+1)t}}{3(1-e^t)^2(-e^t) + 6e^{2t}(1-e^t)}$$

$$= \frac{m(m+1)^3 - m^3 - m^4}{2m} + \frac{m+1}{2}$$

$$+ \lim_{t \to 0} \frac{2e^{2t}}{m} \lim_{t \to 0} \frac{-m^3 e^{tm} - m^4 e^{tm} + m(m+1)^3 e^{(m+1)t}}{6e^{2t}(1-e^t) + (\text{terms that tend zero})}$$

$$= \frac{(m+1)^3 - m^2 - m^3}{2} + \frac{m+1}{2} + \frac{-(m+1)^3 + m^2 + m^3}{3}$$

$$= \frac{-m^2 - m^3 + (m+1)^3 + 3m + 3}{6} = \frac{2m^2 + 6m + 6}{6}$$

Then,

$$V(X) = \frac{3m^2 + 4m + 4}{3} - \frac{(m+1)^2}{4} = \frac{12m^2 + 16m + 16 - 3m^2 - 6m - 3}{12}$$

$$= \frac{9m^2 + 10m - 13}{12}$$

II-2. a) $E(e^{tX}) = \sum_{x=0}^{\infty} e^{tx} \frac{e^{-\lambda}\lambda^x}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda e^t)^x}{x!} = e^{-\lambda} e^{\lambda e^t} = e^{\lambda(e^t - 1)}$

b) $\frac{dM(t)}{dt} = \lambda e^t e^{\lambda(e^t - 1)}$

$$\left.\frac{dM(t)}{dt}\right|_{t=0} = \lambda = E(X)$$

$$\frac{d^2M(t)}{dt^2} = \lambda^2 e^{2t} e^{\lambda(e^t-1)} + \lambda e^t e^{\lambda(e^t-1)}$$

$$\left.\frac{d^2M(t)}{dt^2}\right|_{t=0} = \lambda^2 + \lambda$$

$$V(X) = \lambda^2 + \lambda - \lambda^2 = \lambda$$

II-3. a) $E(e^{tX}) = \sum_{x=1}^{\infty} e^{tx}(1-p)^{x-1}p = \frac{p}{1-p} \sum_{x=1}^{\infty} [e^t(1-p)]^x$

$$= \frac{e^t(1-p)}{1-(1-p)e^t}\left(\frac{p}{1-p}\right) = \frac{pe^t}{1-(1-p)e^t}$$

b) $\frac{dM(t)}{dt} = pe^t(1-(1-p)e^t)^{-2}(1-p)e^t + pe^t(1-(1-p)e^t)^{-1}$

$$= p(1-p)e^{2t}(1-(1-p)e^t)^{-2} + pe^t(1-(1-p)e^t)^{-1}$$

$$\left.\frac{dM(t)}{dt}\right|_{t=0} = \frac{1-p}{p} + 1 = \frac{1}{p} = E(X)$$

$$\frac{d^2M(t)}{dt^2} = p(1-p)e^{2t}2(1-(1-p)e^t)^{-3}(1-p)e^t + p(1-p)(1-(1-p)e^t)^{-2}2e^{2t}$$

$$+ pe^t(1-(1-p)e^t)^{-2}(1-p)e^t + pe^t(1-(1-p)e^t)^{-1}$$

$$\left.\frac{d^2M(t)}{dt^2}\right|_{t=0} = \frac{2(1-p)^2}{p^2} + \frac{2(1-p)}{p} + \frac{1-p}{p} + 1 = \frac{2(1-p)^2 + 3p(1-p) + p^2}{p^2}$$

$$= \frac{2 + 2p^2 - 4p + 3p - 3p^2 + p^2}{p^2} = \frac{2-p}{p^2}$$

$$V(X) = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2}$$

II-4. $M_Y(t) = Ee^{tY} = Ee^{t(X_1+X_2)} = Ee^{tX_1}Ee^{tX_2}$

$$= (1-2t)^{-k_1/2}(1-2t)^{-k_2/2} = (1-2t)^{-(k_1+k_2)/2}$$

Therefore, Y has a chi-square distribution with $k_1 + k_2$ degrees of freedom.

II-5. a) $E(e^{tX}) = \int_0^{\infty} e^{tx} 4xe^{-2x}dx = 4\int_0^{\infty} xe^{(t-2)x}dx$

Using integration by parts with u = x and $dv = e^{(t-2)x}dx$ and du = 1,

$v = \dfrac{e^{(t-2)x}}{t-2}$ we obtain

$$\left.\frac{xe^{(t-2)x}}{t-2}\right|_0^{\infty} - \int_0^{\infty}\frac{e^{(t-2)x}}{t-2}dx = \left.\frac{xe^{(t-2)x}}{t-2}\right|_0^{\infty} - \left.\frac{e^{(t-2)x}}{(t-2)^2}\right|_0^{\infty}$$

This integral only exists for t < 2. In that case, $E(e^{tX}) = \dfrac{4}{(t-2)^2}$ for t < 2

b) $\dfrac{dM(t)}{dt} = -8(t-2)^{-3}$ and $\left.\dfrac{dM(t)}{dt}\right|_{t=0} = -8(-2)^{-3} = 1 = E(X)$

c) $\dfrac{d^2M(t)}{dt^2} = 24(t-2)^{-4}$ and $\left.\dfrac{d^2M(t)}{dt^2}\right|_{t=0} = \dfrac{24}{16} = \dfrac{3}{2}$. Therefore, $V(X) = \dfrac{3}{2} - (1)^2 = \dfrac{1}{2}$

II-6. a) $E(e^{tX}) = \int_{\alpha}^{\beta}\dfrac{e^{tx}}{\beta-\alpha}dx = \left.\dfrac{e^{tx}}{t(\beta-\alpha)}\right|_{\alpha}^{\beta} = \dfrac{e^{t\beta}-e^{t\alpha}}{t(\beta-\alpha)}$

b) $\dfrac{dM(t)}{dt} = \dfrac{e^{t\beta}-e^{t\alpha}}{-(\beta-\alpha)t^2} + \dfrac{\beta e^{t\beta}-e^{t\alpha}}{t(\beta-\alpha)}$

$$= \frac{(\beta t-1)e^{t\beta} - (\alpha t-1)e^{t\alpha}}{t^2(\beta-\alpha)}$$

Using L'Hospital's rule,

$$\lim_{t\to 0}\frac{dM(t)}{dt} = \frac{(\beta t-1)\beta e^{t\beta} + \beta e^{t\beta} - (\alpha t-1)\alpha e^{t\alpha} - \alpha e^{t\alpha}}{2t(\beta-\alpha)}$$

$$= \frac{\beta^2(\beta t-1)e^{t\beta} + \beta^2 e^{t\beta} + \beta^2 e^{t\beta} - \alpha^2(\alpha t-1)e^{t\alpha} - \alpha^2 e^{t\alpha} - \alpha^2 e^{t\alpha}}{2(\beta-\alpha)}$$

$$= \frac{\beta^2-\alpha^2}{2(\beta-\alpha)} = \frac{1}{2(\beta+\alpha)} = E(X)$$

II-7. a) $M(t) = \int_0^{\infty} e^{tx}\lambda e^{-\lambda x}dx = \lambda \int_0^{\infty} e^{(t-\lambda)x}dx$

$$= \lambda \frac{e^{(t-\lambda)x}}{t-\lambda}\Bigg|_0^{\infty} = \frac{-\lambda}{t-\lambda} = \frac{1}{1-\frac{t}{\lambda}} = \left(1-\frac{t}{\lambda}\right)^{-1} \text{ for } t < \lambda$$

b) $\frac{dM(t)}{dt} = (-1)\left(1-\frac{t}{\lambda}\right)^{-2}\left(\frac{-1}{\lambda}\right) = \frac{1}{\lambda\left(1-\frac{t}{\lambda}\right)^2}$

$$\left.\frac{dM(t)}{dt}\right|_{t=0} = \frac{1}{\lambda}$$

$$\frac{d^2M(t)}{dt^2} = \frac{2}{\lambda^2(1-\frac{t}{\lambda})^3}$$

$$\left.\frac{d^2M(t)}{dt^2}\right|_{t=0} = \frac{2}{\lambda^2}$$

$$V(X) = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2}$$

II-8. a) $M(t) = \int_0^{\infty} e^{tx}\frac{\lambda}{\Gamma(r)}(\lambda x)^{r-1}e^{-\lambda x}dx = \frac{\lambda^r}{\Gamma(r)}\int_0^{\infty} x^{r-1}e^{(t-\lambda)x}dx$

Let u = (λ-t)x. Then,

$M(t) = \frac{\lambda^r}{\Gamma(r)}\int_0^{\infty}\left(\frac{u}{\lambda-t}\right)^{r-1} e^{-u}\frac{du}{\lambda-t} = \frac{\lambda^r\Gamma(r)}{\Gamma(r)(\lambda-t)^r} = \frac{1}{(1-\frac{t}{\lambda})^r} = (1-\frac{t}{\lambda})^{-r}$ from the definition of the gamma function for $t < \lambda$.

b) $M'(t) = -r(1-\frac{t}{\lambda})^{-r-1}(-\frac{1}{\lambda})$

$$M'(t)\big|_{t=0} = \frac{r}{\lambda} = E(X)$$

$$M''(t) = \frac{r(r+1)}{\lambda^2}(1-\frac{t}{\lambda})^{r-2}$$

$$M''(t)\big|_{t=0} = \frac{r(r+1)}{\lambda^2}$$

$$V(X) = \frac{r(r+1)}{\lambda^2} - \left(\frac{r}{\lambda}\right)^2 = \frac{r}{\lambda^2}$$

II-9. a) $E(e^{tY}) = \prod_{i=1}^{n} E(e^{tX_i}) = \left(1-\frac{t}{\lambda}\right)^{-n}$

b) From Exercise II-8, Y has a gamma distribution with parameter λ and n.

II-10. a) $M_Y(t) = e^{\mu_1 t+\sigma_1^2\frac{t^2}{2}+\mu_2 t+\sigma_2^2\frac{t^2}{2}} = e^{(\mu_1+\mu_2)t+(\sigma_1^2+\sigma_2^2)\frac{t^2}{2}}$

b) Y has a normal distribution with mean $\mu_1 + \mu_2$ and variances $\sigma_1^2 + \sigma_2^2$

II-11. Because a chi-square distribution is a special case of the gamma distribution with $\lambda = \frac{1}{2}$ and $r = \frac{k}{2}$, from Exercise II-8.

$$M(t) = (1-2t)^{-k/2}$$

$$M'(t) = -\frac{k}{2}(1-2t)^{-\frac{k}{2}-1}(-2) = k(1-2t)^{-\frac{k}{2}-1}$$

$$M'(t)\big|_{t=0} = k = E(X)$$

$$M''(t) = 2k(\frac{k}{2}+1)(1-2t)^{-\frac{k}{2}-2}$$

$$M''(t)\big|_{t=0} = 2k(\frac{k}{2}+1) = k^2 + 2k$$

$$V(X) = k^2 + 2k - k^2 = 2k$$

II-12. a) $M(t) = M(0) + M'(0)t + M'(0)\frac{t^2}{2!} + ... + M^{(r)}(0)\frac{t^r}{r!} + ...$ by Taylor's expansion. Now, M(0) = 1 and $M^{(r)}(0) = \mu_r'$ and the result is obtained.

b) From Exercise II-8, $M(t) = 1 + \frac{r}{\lambda}t + \frac{r(r+1)}{\lambda^2}\frac{t^2}{2!} + ...$

c) $\mu_1' = \frac{r}{\lambda}$ and $\mu_2' = \frac{r(r+1)}{\lambda^2}$ which agrees with Exercise II-8.

Appendix III

III-1. $f_Y(y) = \frac{1}{4}$ at y = 3, 5, 7, 9 from Theorem III-1.

III-2. Because $X \geq 0$, the transformation is one-to-one; that is $y = x^2$ and $x = \sqrt{y}$. From Theorem III-1,

$$f_Y(y) = f_X(\sqrt{y}) = \binom{3}{\sqrt{y}} p^{\sqrt{y}} (1-p)^{3-\sqrt{y}} \text{ for y = 0, 1, 4, 9.}$$

If p = 0.25, $f_Y(y) = \binom{3}{\sqrt{y}} (0.25)^{\sqrt{y}} (0.75)^{3-\sqrt{y}}$ for y = 0, 1, 4, 9.

III-3. a) $f_Y(y) = f_X\left(\frac{y-10}{2}\right)\left(\frac{1}{2}\right) = \frac{y-10}{72}$ for $10 \leq y \leq 22$

b) $E(Y) = \int_{10}^{22} \frac{y^2 - 10y}{72} dy = \frac{1}{72}\left(\frac{y^3}{3} - \frac{10y^2}{2}\right)\Bigg|_{10}^{22} = 18$

III-4. Because y = -2 ln x, $e^{-\frac{y}{2}} = x$. Then, $f_Y(y) = f_X(e^{\frac{-y}{2}})\left|-\frac{1}{2}e^{\frac{-y}{2}}\right| = \frac{1}{2}e^{\frac{-y}{2}}$ for $0 \leq e^{\frac{-y}{2}} \leq 1$ or $y \geq 0$,

which is an exponential distribution (which equals a chi-square distribution with k = 2 degrees of freedom).

III-5. a) Let Q = R. Then,

$$p = i^2 r \qquad q = r \qquad \text{and} \qquad i = \sqrt{\frac{p}{q}} \qquad r = q$$

$$J = \begin{vmatrix} \frac{\partial i}{\partial p} & \frac{\partial i}{\partial q} \\ \frac{\partial r}{\partial p} & \frac{\partial r}{\partial q} \end{vmatrix} = \begin{vmatrix} \frac{1}{2}(pq)^{-1/2} & -\frac{1}{2}p^{1/2}q^{-3/2} \\ 0 & 1 \end{vmatrix} = \frac{1}{2}(pq)^{-1/2}$$

$$f_{PQ}(p,q) = f_{IR}\left(\sqrt{\frac{p}{q}}, q\right)\frac{1}{2}(pq)^{-1/2} = 2\left(\sqrt{\frac{p}{q}}\right)\frac{1}{2}(pq)^{-1/2} = q^{-1}$$

for $0 \leq \sqrt{\frac{p}{q}} \leq 1$, $0 \leq q \leq 1$

That is, for $0 \leq p \leq q$, $0 < q \leq 1$.

$$f_P(p) = \int_p^1 q^{-1} dq = -\ln p \quad \text{for } 0 < p \leq 1.$$

b) $E(P) = -\int_0^1 p \ln p \; dp$. Let *u* = ln p and *dv* = p dp. Then, *du* = 1/p and

v = $\frac{p^2}{2}$. Therefore, $E(P) = -(\ln p)\frac{p^2}{2}\Big|_0^1 + \int_0^1 \frac{p}{2} dp = \frac{p^2}{4}\Big|_0^1 = \frac{1}{4}$

III-6. a) If $y = x^2$, then $x = \sqrt{y}$ for $x \geq 0$ and $y \geq 0$. Thus, $f_Y(y) = f_X(\sqrt{y})\frac{1}{2}y^{-\frac{1}{2}} = \frac{e^{-\sqrt{y}}}{2\sqrt{y}}$ for y > 0.

b) If $y = x^{1/2}$, then $x = y^2$ for $x \geq 0$ and $y \geq 0$. Thus, $f_Y(y) = f_X(y^2)2y = 2ye^{-y^2}$ for y > 0.

c) If y = ln x, then $x = e^y$ for $x \geq 0$. Thus, $f_Y(y) = f_X(e^y)e^y = e^y e^{-e^y} = e^{y-e^y}$ for $-\infty < y < \infty$.

III-7. a) Now, $\int_0^\infty av^2e^{-bv}dv$ must equal one. Let u = bv, then 1 = $a\int_0^\infty (\frac{u}{b})^2 e^{-u}\frac{du}{b} = \frac{a}{b^3}\int_0^\infty u^2e^{-u}du$. From the definition of the gamma function the last expression is $\frac{a}{b^3}\Gamma(3) = \frac{2a}{b^3}$. Therefore, $a = \frac{b^3}{2}$.

b) If $w = \frac{mv^2}{2}$, then $v = \sqrt{\frac{2w}{m}}$ for $v \geq 0,\ \ w \geq 0$.

$$f_W(w) = f_V\left(\sqrt{\frac{2w}{m}}\right)\frac{dv}{dw} = \frac{b^3 2w}{2m}e^{-b\sqrt{\frac{2w}{m}}}(2mw)^{-1/2}$$

$$= b^3m^{-3/2}w^{1/2}e^{-b\sqrt{\frac{2w}{m}}}$$

for $w \geq 0$.

III-8. If $y = e^x$, then x = ln y for $1 \leq x \leq 2$ and $e^1 \leq y \leq e^2$. Thus, $f_Y(y) = f_X(\ln y)\frac{1}{y} = \frac{1}{y}$ for $1 \leq \ln y \leq 2$.

That is, $f_Y(y) = \frac{1}{y}$ for $e \leq y \leq e^2$.

III-9. Now $P(Y \leq a) = P(X \geq u(a)) = \int_{u(a)}^{\infty} f_X(x)dx$. By changing the variable of integration from x to y by using x = u(y), we obtain

$P(Y \leq a) = \int_a^{-\infty} f_X(u(y))u'(y)dy$ because as x tends to , y = h(x) tends to - . Then,

$P(Y \leq a) = \int_{-\infty}^{a} f_X(u(y))(-u'(y))dy$. Because h(x) is decreasing, $u'(y)$ is negative. Therefore, $|u'(y)|$ = -$u'(y)$ and Theorem III-1 holds in this case also.

III-10. If y = $(x-2)^2$, then x = $2-\sqrt{y}$ for $0 \leq x \leq 2$ and x = $2+\sqrt{y}$ for $2 \leq x \leq 4$. Thus,

$$f_Y(y) = f_X(2-\sqrt{y})|-\tfrac{1}{2}y^{-1/2}| + f_X(2+\sqrt{y})|\tfrac{1}{2}y^{-1/2}|$$

$$= \frac{2-\sqrt{y}}{16\sqrt{y}} + \frac{2+\sqrt{y}}{16\sqrt{y}}$$

$$= \left(\frac{1}{4}\right)y^{-1/2} \text{ for } 0 \leq y \leq 4$$

III-11. a) Let $a = s_1 s_2$ and y = s_1. Then, s_1 = y, $s_2 = \frac{a}{y}$ and

$$J = \begin{vmatrix} \frac{\partial s_1}{\partial a} & \frac{\partial s_1}{\partial y} \\ \frac{\partial s_2}{\partial a} & \frac{\partial s_2}{\partial y} \end{vmatrix} = \begin{vmatrix} 0 & 1 \\ \frac{1}{y} & -\frac{a}{2} y^{-3/2} \end{vmatrix} = \frac{1}{y}$$. Then, $f_{AY}(a, y) = f_{S_1 S_2}(y, \frac{a}{y})(\frac{1}{y}) = 2y(\frac{a}{8y})(\frac{1}{y}) = \frac{a}{4y}$

for $0 \le y \le 1$ and $0 \le \frac{a}{y} \le 4$. That is, for $0 \le y \le 1$ and $0 \le a \le 4y$.

b) $f_A(a) = \int_{a/4}^{1} \frac{a}{4y} dy = -\frac{a}{4} \ln(\frac{a}{4})$ for $0 < a \le 4$.

III-12. Let r = v/i and s = i. Then, $i = s$ and v = rs

$$J = \begin{vmatrix} \frac{\partial i}{\partial r} & \frac{\partial i}{\partial s} \\ \frac{\partial v}{\partial r} & \frac{\partial v}{\partial s} \end{vmatrix} = \begin{vmatrix} 0 & 1 \\ s & r \end{vmatrix} = s$$

$f_{RS}(r, s) = f_{IV}(s, rs)s = e^{-rs}s$ for $rs \ge 0$ and $1 \le s \le 2$. That is,

$f_{RS}(r, s) = se^{-rs}$ for $1 \le s \le 2$ and $r \ge 0$.

Then, $f_R(r) = \int_1^2 se^{-rs} ds$. Let *u* = s and *dv* = e^{-rs} ds. Then, *du* = ds and *v* = $\frac{-e^{-rs}}{r}$

Then,

$$f_R(r) = -s\frac{-e^{-rs}}{r}\Bigg|_1^2 + \int_1^2 \frac{-e^{-rs}}{r} ds = \frac{e^{-r} - 2e^{-2r}}{r} - \frac{-e^{-rs}}{r^2}\Bigg|_1^2$$

$$= \frac{e^{-r} - 2e^{-2r}}{r} + \frac{e^{-r} - e^{-2r}}{r^2}$$

$$= \frac{e^{-r}(r+1) - e^{-2r}(2r+1)}{r^2}$$

for r > 0.

Appendix V

V-1. a) From Example V-1 the posterior distribution for μ is normal with mean $\dfrac{(\sigma^2\mu_0 + \sigma_0^2 x)}{\sigma_0^2 + \sigma^2}$ and variance

$$\frac{\sigma_0^2\sigma^2}{\sigma_0^2 + \sigma^2}.$$

b) The Bayes estimator for μ is obtained from Example V-1 to be $\dfrac{(\sigma^2\mu_0 + \sigma_0^2 x)}{\sigma_0^2 + \sigma^2}$

V-2. a) Because $f(x|\mu) = \dfrac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ and $f(\mu) = \dfrac{1}{b-a}$ for $a \le \mu \le b$, the joint distribution is

$$f(x,\mu) = \frac{1}{(b-a)\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$ for $-\infty < x < \infty$ and $a \le \mu \le b$. Then,

$$f(x) = \frac{1}{b-a}\int_a^b \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} d\mu$$ and this integral is recognized as a normal probability. Therefore,

$$f(x) = \frac{1}{b-a}\left[\Phi\left(\frac{b-x}{\sigma}\right) - \Phi\left(\frac{a-x}{\sigma}\right)\right]$$ where $\Phi(x)$ is the standard normal cumulative distribution function.

Then, $$f(\mu|x) = \frac{f(x,\mu)}{f(x)} = \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma\left[\Phi\left(\frac{b-x}{\sigma}\right) - \Phi\left(\frac{a-x}{\sigma}\right)\right]}$$

b) The Bayes estimator is $$\tilde{\mu} = \int_a^b \frac{\mu e^{-\frac{(x-\mu)^2}{2\sigma^2}} d\mu}{\sqrt{2\pi}\sigma\left[\Phi\left(\frac{b-x}{\sigma}\right) - \Phi\left(\frac{a-x}{\sigma}\right)\right]}.$$

Let $v = (x - \mu)$. Then, $dv = - d\mu$ and

$$\tilde{\mu} = \int_{x-b}^{x-a} \frac{(x-v)e^{-\frac{v^2}{2\sigma^2}} dv}{\sqrt{2\pi}\sigma\left[\Phi\left(\frac{b-x}{\sigma}\right) - \Phi\left(\frac{a-x}{\sigma}\right)\right]}$$

$$= \frac{x\left[\Phi\left(\frac{x-a}{\sigma}\right) - \Phi\left(\frac{x-b}{\sigma}\right)\right]}{\left[\Phi\left(\frac{b-x}{\sigma}\right) - \Phi\left(\frac{a-x}{\sigma}\right)\right]} - \int_{x-b}^{x-a} \frac{v e^{-\frac{v^2}{2\sigma^2}} dv}{\sqrt{2\pi}\sigma\left[\Phi\left(\frac{b-x}{\sigma}\right) - \Phi\left(\frac{a-x}{\sigma}\right)\right]}$$

Let $w = \dfrac{v^2}{2\sigma^2}$. Then, $dw = [\frac{2v}{2\sigma^2}]dv = [\frac{v}{\sigma^2}]dv$ and

$$\tilde{\mu} = x - \int_{\frac{(x-b)^2}{2\sigma^2}}^{\frac{(x-a)^2}{2\sigma^2}} \frac{\sigma e^{-w} dw}{\sqrt{2\pi}\left[\Phi\left(\frac{b-x}{\sigma}\right) - \Phi\left(\frac{a-x}{\sigma}\right)\right]}$$

$$= x + \frac{\sigma}{\sqrt{2\pi}} \left[\frac{e^{-\frac{(x-a)^2}{2\sigma^2}} - e^{-\frac{(x-b)^2}{2\sigma^2}}}{\Phi\left(\frac{b-x}{\sigma}\right) - \Phi\left(\frac{a-x}{\sigma}\right)} \right]$$

V-3. a) $f(x) = \frac{e^{-\lambda}\lambda^x}{x!}$ for x = 0, 1, 2, and $f(\lambda) = \left(\frac{m+1}{\lambda_0}\right)^{m+1} \frac{\lambda^m e^{-(m+1)\frac{\lambda}{\lambda_0}}}{\Gamma(m+1)}$ for $\lambda > 0$. Then,

$$f(x,\lambda) = \frac{(m+1)^{m+1}\lambda^{m+x} e^{-\lambda-(m+1)\frac{\lambda}{\lambda_0}}}{\lambda_0^{m+1}\Gamma(m+1)x!}.$$

This last density is recognized to be a gamma density as a function of λ. Therefore, the posterior distribution of λ is a gamma distribution with parameters m + x + 1 and 1 + $\frac{m+1}{\lambda_0}$.

b) The mean of the posterior distribution can be obtained from the results for the gamma distribution to be

$$\frac{m+x+1}{\left[1+\frac{m+1}{\lambda_0}\right]} = \lambda_0\left(\frac{m+x+1}{m+\lambda_0+1}\right)$$

V-4. a) From Example V-1, the Bayes estimate is $\tilde{\mu} = \frac{\frac{9}{25}(4) + 1(4.85)}{\frac{9}{25}+1} = 4.625$

b) The posterior variance is (from Example V-1) $\frac{1(\frac{9}{25})}{1+\frac{9}{25}} = 0.2647$. Therefore, as in Example V-2 a 95%

Bayes interval estimate for μ is $4.625 \pm 1.96\sqrt{0.2647} = 4.625 \pm 1.008 = (3.62, 5.63)$

c) The classical 95% confidence interval for μ is

$$\bar{x} \pm 1.96\frac{\sigma}{\sqrt{n}} = 4.85 \pm 1.96(\tfrac{3}{5}) = 4.85 \pm 1.176 = (3.67, 6.03).$$

The classical interval is longer because it does not take into account the prior information.

V-5. a) From Example V-1, $\tilde{\mu} = \frac{(0.01)(5.03) + (\frac{1}{25})(5.05)}{0.01+\frac{1}{25}} = 5.046$

b) The posterior variance is $\frac{(0.01)(0.04)}{0.01+0.04} = 0.008$. The 95% Bayes interval estimate is

$5.046 \pm 1.96\sqrt{0.008} = 5.046 \pm 0.1753 = (4.87, 5.22)$

c) Let X denote the weight of a box. Then, P(X < 4.95) = $P(\frac{X-\mu}{\sigma} < \frac{4.95-\mu}{\sigma})$

V-6. $f(x|\lambda) = \lambda e^{-\lambda x}$, $x \geq 0$ and $f(\lambda) = 0.01e^{-0.01\lambda}$. Then,

$f(x_1, x_2, \lambda) = \lambda^2 e^{-\lambda(x_1+x_2)} 0.01e^{-0.01\lambda} = 0.01\lambda^2 e^{-\lambda(x_1+x_2+0.01)}$. As a function of λ, this is recognized as a gamma density with parameters 3 and $x_1 + x_2 + 0.01$. Therefore, the posterior mean for λ is

$$\tilde{\lambda} = \frac{3}{x_1+x_2+0.01} = \frac{3}{2\bar{x}+0.01} = 0.00133.$$

Appendix VII

VII-1. a) Analysis of Variance for Force

Source	DF	SS	MS	F	P
Primer	2	4.5811	2.2906	19.00	0.050
Appl	1	4.9089	4.9089	40.72	0.024
Primer*Appl	2	0.2411	0.1206	1.47	0.269
Error	12	0.9867	0.0822		
Total	17	10.7178			

b) $\sigma^2 = 0.082$

$$\sigma^2_{\tau\beta} = \frac{0.1206 - 0.0822}{3} = 0.013$$

$$\sigma^2_{\beta} = \frac{4.9089 - 0.1206}{9} = 0.532$$

$$\sigma^2_{\tau} = \frac{2.2906 - 0.1206}{6} = 0.362$$

VII-2.

Source	DF	SS	MS	F	P
Machine	3	24.333	8.111	1.36	0.340
Operator	2	26.333	13.167	2.21	0.190
Machine*Operator	6	35.667	5.944	1.59	0.234
Error	12	45.000	3.750		
Total	23	131.333			

With operators random, there is no significant difference among operators or machines.

VII-3.

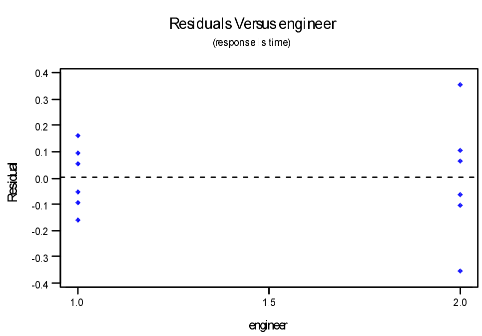
Source	DF	SS	MS	F	P
operator	2	0.01005	0.00503	0.07	0.937
engineer	1	0.04688	0.04688	0.62	0.512
operator*engineer	2	0.15005	0.07503	1.26	0.350
Error	6	0.35785	0.05964		
Total	11	0.56483			

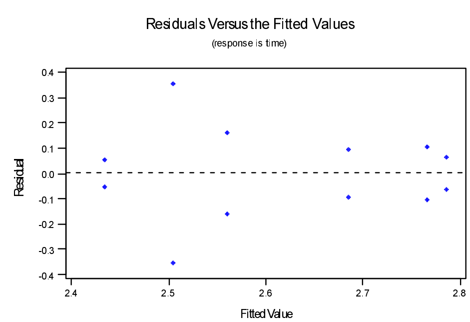
Source	Variance component	Error term	Expected Mean Square (using unrestricted model)
1 operator	-0.01750	3	(4) + 2(3) + 4(1)
2 engineer		3	(4) + 2(3) + Q[2]
3 operator*engineer	0.00769	4	(4) + 2(3)
4 Error	0.05964		(4)

a) $H_0{:}\sigma^2_{\tau\beta} = 0 \quad H_1{:}\sigma^2_{\tau\beta} \neq 0$

b) Because the significance level for operator*engineer is 0.350, we cannot reject H_0.

c) Possible violation of the constant variance assumption.





d) $\sigma^2 = 0.0596$

$$\sigma^2_{\tau\beta} = \frac{0.0750 - 0.0596}{2} = 0.0077$$

$$\sigma^2_{\beta} = \frac{0.0050 - 0.0596}{4} = -0.0137 \rightarrow 0$$

VII-4.

Source	DF	SS	MS	F	P
furnacep	1	7160	7160	17.50	0.053
temperat	2	945342	472671	1155.52	0.001
furnacep*temperat	2	818	409	0.91	0.427
Error	12	5371	448		
Total	17	958691			

Source	Variance component	Error term	Expected Mean Square (using unrestricted model)
1 furnacep	750.11	3	(4) + 3(3) + 9(1)
2 temperat		3	(4) + 3(3) + Q[2]
3 furnacep*temperat	-12.83	4	(4) + 3(3)
4 Error	447.56		(4)

a) $H_0: \sigma^2_{\tau} = 0$ $\quad H_1: \sigma^2_{\tau} \neq 0$

$H_0: \beta_1 = \beta_2 = \beta_3 = 0$ $\quad H_1$: not all $\beta_i = 0$

$H_0: \sigma^2_{\tau\beta} = 0$ $\quad H_1: \sigma^2_{\tau\beta} \neq 0$

b) The second null hypotheses can be rejected at = 0.05 from the ANOVA table. The first and third null hypothesis cannot be rejected.

c) $\sigma^2 = 447.56$

$$\sigma^2_{\tau\beta} = \frac{409.056 - 447.556}{3} = -12.83 \rightarrow 0$$

$$\sigma^2_{\tau} = \frac{7160.06 - 409}{9} = 745.83$$

VII-5.

Source	DF	SS	MS	F	P
operator	2	0.01005	0.00503	0.07	0.937
engineer	1	0.04688	0.04688	0.62	0.512
operator*engineer	2	0.15005	0.07503	1.26	0.350
Error	6	0.35785	0.05964		
Total	11	0.56483			

Source	Variance component	Error term	Expected Mean Square (using unrestricted model)
1 operator	-0.01750	3	(4) + 2(3) + 4(1)
2 engineer	-0.00469	3	(4) + 2(3) + 6(2)
3 operator*engineer	0.00769	4	(4) + 2(3)
4 Error	0.05964		(4)

The ANOVA table is shown above. The conclusions do not differ from Exercise VII-3. No effects are significant.