Applied Statistics and Probability for Engineers

Third Edition

Douglas C. Montgomery
Arizona State University

George C. Runger
Arizona State University

John Wiley & Sons, Inc.

ACQUISITIONS EDITOR *Wayne Anderson*
ASSISTANT EDITOR *Jenny Welter*
MARKETING MANAGER *Katherine Hepburn*
SENIOR PRODUCTION EDITOR *Norine M. Pigliucci*
DESIGN DIRECTOR *Maddy Lesure*
ILLUSTRATION EDITOR *Gene Aiello*
PRODUCTION MANAGEMENT SERVICES *TechBooks*

This book was set in Times Roman by TechBooks and printed and bound by Donnelley/Willard.
The cover was printed by Phoenix Color Corp.

This book is printed on acid-free paper. ∞

Library of Congress Cataloging-in-Publication Data

Montgomery, Douglas C.
Applied statistics and probability for engineers / Douglas C. Montgomery, George C. Runger.—3rd ed.
p. cm.
Includes bibliographical references and index.
ISBN 0-471-20454-4 (acid-free paper)
1. Statistics. 2. Probabilities. I. Runger, George C. II. Title.

QA276.12.M645 2002
519.5—dc21

2002016765

Printed in the United States of America.

10 9 8 7 6 5 4 3 2 1

Preface

The purpose of this *Student Solutions Manual* is to provide you with additional help in understanding the problem-solving processes presented in *Applied Statistics and Probability for Engineers*. The *Applied Statistics* text includes a section entitled "Answers to Selected Exercises," which contains the final answers to most odd-numbered exercises in the book. Within the text, problems with an answer available are indicated by the exercise number enclosed in a box.

This *Student Solutions Manual* provides complete worked-out solutions to a subset of the problems included in the "Answers to Selected Exercises." If you are having difficulty reaching the final answer provided in the text, the complete solution will help you determine the correct way to solve the problem.

Those problems with a complete solution available are indicated in the "Answers to Selected Exercises," again by a box around the exercise number. The complete solutions to this subset of problems may also be accessed by going directly to this *Student Solutions Manual*.

Chapter 2 Selected Problem Solutions

Section 2-2

2-43. 3 digits between 0 and 9, so the probability of any three numbers is 1/(10*10*10);
3 letters A to Z, so the probability of any three numbers is 1/(26*26*26); The probability your license plate is chosen is then $(1/10^3)*(1/26^3) = 5.7 \times 10^{-8}$

Section 2-3

2-49. 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) $P((A \cup B)') = 1 - P(A \cup B) = 1 - 0.4 = 0.6$
f) $P(A' \cup B) = P(A') + P(B) - P(A' \cap B) = 0.7 + 0.2 - 0.1 = 0.8$

Section 2-4

2-61. Need data from example
a) P(A) = 0.05 + 0.10 = 0.15
b) $P(A|B) = \dfrac{P(A \cap B)}{P(B)} = \dfrac{0.04 + 0.07}{0.72} = 0.153$
c) P(B) = 0.72
d) $P(B|A) = \dfrac{P(A \cap B)}{P(B)} = \dfrac{0.04 + 0.07}{0.15} = 0.733$
e) $P(A \cap B) = 0.04 + 0.07 = 0.11$
f) $P(A \cup B) = 0.15 + 0.72 – 0.11 = 0.76$

2-67. a) P(gas leak) = (55 + 32)/107 = 0.813
b) P(electric failure|gas leak) = (55/107)/(87/102) = 0.632
c) P(gas leak| electric failure) = (55/107)/(72/107) = 0.764

Section 2-5

2-73. Let F denote the event that a roll contains a flaw.
Let C denote the event that a roll is cotton.

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

2-79. 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.

$$
\begin{aligned}
P(C) &= P(C|A \cap B)P(A \cap B) + P(C|A \cap B')P(A \cap B') \\
&\quad + P(C|A' \cap B)P(A' \cap B) + P(C|A' \cap B')P(A' \cap B') \\
&= \frac{3}{23}\left(\frac{2}{24}\right)\left(\frac{5}{25}\right) + \frac{4}{23}\left(\frac{20}{24}\right)\left(\frac{5}{25}\right) + \frac{4}{23}\left(\frac{5}{24}\right)\left(\frac{20}{25}\right) + \frac{5}{23}\left(\frac{19}{24}\right)\left(\frac{20}{25}\right) \\
&= 0.20
\end{aligned}
$$

Section 2-6

2-87. 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.

2-89. 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, $P(B_1 \cup B_2 \cup ... \cup B_8) = P(B_1) + P(B_2) + ... + P(B_8)$

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 $5(\frac{1}{8})^4(\frac{7}{8}) = 0.00107$.

Section 2-7

2-97. 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)

$$\begin{aligned} P(G) &= P(G|H)P(H) + P(G|M)P(M) + P(G|P)P(P) \\ &= 0.95(0.40) + 0.60(0.35) + 0.10(0.25) \\ &= 0.615 \end{aligned}$$

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

Supplemental

2-105. a) No, $P(E_1 \cap E_2 \cap E_3) \neq 0$

b) No, $E_1' \cap E_2'$ is not $\varnothing$

c) $P(E_1' \cup E_2' \cup E_3') = P(E_1') + P(E_2') + P(E_3') - P(E_1' \cap E_2') - P(E_1' \cap E_3') - P(E_2' \cap E_3') + P(E_1' \cap E_2' \cap E_3')$
$= 40/240$

d) $P(E_1 \cap E_2 \cap E_3) = 200/240$
e) $P(E_1 \cup E_3) = P(E_1) + P(E_3) - P(E_1 \cap E_3) = 234/240$
f) $P(E_1 \cup E_2 \cup E_3) = 1 - P(E_1' \cap E_2' \cap E_3') = 1 - 0 = 1$

2-107. Let A_i denote the event that the ith bolt selected is not torqued to the proper limit.
a) Then,

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

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,

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

2-113. D = defective copy

a) $$P(D = 1) = \left(\frac{2}{75}\right)\left(\frac{73}{74}\right)\left(\frac{72}{73}\right) + \left(\frac{73}{75}\right)\left(\frac{2}{74}\right)\left(\frac{72}{73}\right) + \left(\frac{73}{75}\right)\left(\frac{72}{74}\right)\left(\frac{2}{73}\right) = 0.0778$$

b) $$P(D = 2) = \left(\frac{2}{75}\right)\left(\frac{1}{74}\right)\left(\frac{73}{73}\right) + \left(\frac{2}{75}\right)\left(\frac{73}{74}\right)\left(\frac{1}{73}\right) + \left(\frac{73}{75}\right)\left(\frac{2}{74}\right)\left(\frac{1}{73}\right) = 0.00108$$

2-117. Let A_i denote the event that the ith washer selected is thicker than target.

a) $$\left(\frac{30}{50}\right)\left(\frac{29}{49}\right)\left(\frac{28}{8}\right) = 0.207$$

b) $30/48 = 0.625$

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,

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

2-121. 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 does not operate is

$$P(A_1')P(A_2')P(A_3')P(A_4') = (0.02)(0.0199)(0.0199)(0.02) = 1.58 \times 10^{-7}$$

Chapter 3 Selected Problem Solutions

Section 3-2

3-13.

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

3-21. $P(X = 0) = 0.02^3 = 8 \times 10^{-6}$
$P(X = 1) = 3[0.98(0.02)(0.02)]=0.0012$
$P(X = 2) = 3[0.98(0.98)(0.02)]=0.0576$
$P(X = 3) = 0.98^3 = 0.9412$

3-25. X = number of components that meet specifications
P(X=0) = (0.05)(0.02)(0.01) = 0.00001
P(X=1) = (0.95)(0.02)(0.01) + (0.05)(0.98)(0.01)+(0.05)(0.02)(0.99) = 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

Section 3-3

3-27.

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

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) $P(X > 0) = 1 - P(X \le 0) = 1 - 5/8 = 3/8$

3-31.

$$F(x) = \begin{cases} 0 & x < 0 \\ 0.008 & 0 \le x < 1 \\ 0.104 & 1 \le x < 2 \\ 0.488 & 2 \le x < 3 \\ 1 & 3 \le x \end{cases} \text{ where } \begin{aligned} f(0) &= 0.2^3 = 0.008, \\ f(1) &= 3(0.2)(0.2)(0.8) = 0.096, \\ f(2) &= 3(0.2)(0.8)(0.8) = 0.384, \\ f(3) &= (0.8)^3 = 0.512, \end{aligned}$$

3-33. 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) $P(X>2) = 1 - P(X\le 2) = 0.5$

Section 3-4

3-37 Mean and Variance

$$\mu = E(X) = 0f(0) + 1f(1) + 2f(2) + 3f(3) + 4f(4)$$
$$= 0(0.2) + 1(0.2) + 2(0.2) + 3(0.2) + 4(0.2) = 2$$
$$V(X) = 0^2 f(0) + 1^2 f(1) + 2^2 f(2) + 3^2 f(3) + 4^2 f(4) - \mu^2$$
$$= 0(0.2) + 1(0.2) + 4(0.2) + 9(0.2) + 16(0.2) - 2^2 = 2$$

3-41. Mean and variance for exercise 3-19

$$\mu = E(X) = 10f(10) + 5f(5) + 1f(1)$$
$$= 10(0.3) + 5(0.6) + 1(0.1)$$
$$= 6.1 \text{ million}$$
$$V(X) = 10^2 f(10) + 5^2 f(5) + 1^2 f(1) - \mu^2$$
$$= 10^2(0.3) + 5^2(0.6) + 1^2(0.1) - 6.1^2$$
$$= 7.89 \text{ million}^2$$

3-45. Determine x where range is [0,1,2,3,x] and mean is 6.

$$\mu = E(X) = 6 = 0f(0) + 1f(1) + 2f(2) + 3f(3) + xf(x)$$
$$6 = 0(0.2) + 1(0.2) + 2(0.2) + 3(0.2) + x(0.2)$$
$$6 = 1.2 + 0.2x$$
$$4.8 = 0.2x$$
$$x = 24$$

Section 3-5

3-47. E(X) = (3+1)/2 = 2, V(X) = $[(3-1+1)^2 -1]/12$ = 0.667

3-49. X=(1/100)Y, Y = 15, 16, 17, 18, 19.

E(X) = (1/100) E(Y) = $\frac{1}{100}\left(\frac{15+19}{2}\right) = 0.17$ mm

$$V(X) = \left(\frac{1}{100}\right)^2 \left[\frac{(19-15+1)^2 - 1}{12}\right] = 0.0002 \text{ mm}^2$$

Section 3-6

3-57. a) $P(X = 5) = \binom{10}{5} 0.5^5 (0.5)^5 = 0.2461$

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

$$= 0.5^{10} + 10(0.5)^{10} + 45(0.5)^{10} = 0.0547$$

c) $P(X \ge 9) = \binom{10}{9} 0.5^9 (0.5)^1 + \binom{10}{10} 0.5^{10} (0.5)^0 = 0.0107$

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

$= 120(0.5)^{10} + 210(0.5)^{10} = 0.3223$

3-61. n=3 and p=0.25

$$F(x) = \begin{cases} 0 & x < 0 \\ 0.4219 & 0 \le x < 1 \\ 0.8438 & 1 \le x < 2 \\ 0.9844 & 2 \le x < 3 \\ 1 & 3 \le x \end{cases} \text{ where } \begin{aligned} f(0) &= \left(\frac{3}{4}\right)^3 = \frac{27}{64} \\ f(1) &= 3\left(\frac{1}{4}\right)\left(\frac{3}{4}\right)^2 = \frac{27}{64} \\ f(2) &= 3\left(\frac{1}{4}\right)^2\left(\frac{3}{4}\right) = \frac{9}{64} \\ f(3) &= \left(\frac{1}{4}\right)^3 = \frac{1}{64} \end{aligned}$$

3-63. a) $P(X = 1) = \binom{1000}{1} 0.001^1 (0.999)^{999} = 0.3681$

$b) P(X \ge 1) = 1 - P(X = 0) = 1 - \binom{1000}{1} 0.001^1 (0.999)^{999} = 0.6323$

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

$= 0.9198$

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

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

3-67. 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.

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

$= 1 - \left[\binom{125}{0} 0.1^0 (0.9)^{125} + \binom{125}{1} 0.1^1 (0.9)^{124} + \binom{125}{2} 0.1^2 (0.9)^{123} \right.$

$\left. + \binom{125}{3} 0.1^3 (0.9)^{122} + \binom{125}{4} 0.1^4 (0.9)^{121} \right] = 0.9961$

$b) P(X > 5) = 1 - P(X \le 5) = 0.9886$

3-69. Let X denote the number of questions answered correctly. Then, X is binomial with n = 25 and p = 0.25.

$$a)P(X \geq 20) = \binom{25}{20}0.25^{20}(0.75)^5 + \binom{25}{21}0.25^{21}(0.75)^4 + \binom{25}{22}0.25^{22}(0.75)^3$$
$$+\binom{25}{23}0.25^{23}(0.75)^2 + \binom{25}{24}0.25^{24}(0.75)^1 + \binom{25}{25}0.25^{25}(0.75)^0 \cong 0$$

$$b)P(X < 5) = \binom{25}{0}0.25^0(0.75)^{25} + \binom{25}{1}0.25^1(0.75)^{24} + \binom{25}{2}0.25^2(0.75)^{23}$$
$$+\binom{25}{3}0.25^3(0.75)^{22} + \binom{25}{4}0.25^4(0.75)^{21} = 0.2137$$

Section 3-7

3-71. a. $P(X = 1) = (1 - 0.5)^0 0.5 = 0.5$

b. $P(X = 4) = (1 - 0.5)^3 0.5 = 0.5^4 = 0.0625$

c. $P(X = 8) = (1 - 0.5)^7 0.5 = 0.5^8 = 0.0039$

d. $P(X \leq 2) = P(X = 1) + P(X = 2) = (1 - 0.5)^0 0.5 + (1 - 0.5)^1 0.5$
$= 0.5 + 0.5^2 = 0.75$

e. $P(X > 2) = 1 - P(X \leq 2) = 1 - 0.75 = 0.25$

3-75. Let X denote the number of calls needed to obtain a connection. Then, X is a geometric random variable with p = 0.02

a) $P(X = 10) = (1 - 0.02)^9 0.02 = 0.98^9 0.02 = 0.0167$

b) $P(X > 5) = 1 - P(X \leq 4) = 1 - [P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4)]$
$= 1 - [0.02 + 0.98(0.02) + 0.98^2(0.02) + 0.98^3(0.02)]$
$= 1 - 0.0776 = 0.9224$

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

3-77 p = 0.005 , r = 8

a. $P(X = 8) = 0.0005^8 = 3.91x10^{-19}$

b. $\mu = E(X) = \dfrac{1}{0.005} = 200$ days

c Mean number of days until all 8 computers fail. Now we use p=$3.91x10^{-19}$

$\mu = E(Y) = \dfrac{1}{3.91x10^{-91}} = 2.56x10^{18}$ days or $7.01\ x10^{15}$ years

3-81. 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.

3-83. 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, σ_X = 1731.18

Section 3-8

3-87. a) $$P(X=1)=\frac{\binom{4}{1}\binom{16}{3}}{\binom{20}{4}}=\frac{(4\times16\times15\times14)/6}{(20\times19\times18\times17)/24}=0.4623$$

b) $$P(X=4)=\frac{\binom{4}{4}\binom{16}{0}}{\binom{20}{4}}=\frac{1}{(20\times19\times18\times17)/24}=0.00021$$

c)
$$P(X\le 2)=P(X=0)+P(X=1)+P(X=2)$$
$$=\frac{\binom{4}{0}\binom{16}{4}}{\binom{20}{4}}+\frac{\binom{4}{1}\binom{16}{3}}{\binom{20}{4}}+\frac{\binom{4}{2}\binom{16}{2}}{\binom{20}{4}}$$
$$=\frac{\left(\frac{16\times15\times14\times13}{24}+\frac{4\times16\times15\times14}{6}+\frac{6\times16\times15}{2}\right)}{\left(\frac{20\times19\times18\times17}{24}\right)}=0.9866$$

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

3-91. Let X denote the number of men who carry the marker on the male chromosome for an increased risk for high blood pressure. N=800, K=240 n=10

a) n=10
$$P(X=1)=\frac{\binom{240}{1}\binom{560}{9}}{\binom{800}{10}}=\frac{\left(\frac{240!}{1!239!}\right)\left(\frac{560!}{9!551!}\right)}{\frac{800!}{10!790!}}=0.1201$$

b) n=10
$$P(X>1)=1-P(X\le 1)=1-[P(X=0)+P(X=1)]$$
$$P(X=0)=\frac{\binom{240}{0}\binom{560}{10}}{\binom{800}{10}}=\frac{\left(\frac{240!}{0!240!}\right)\left(\frac{560!}{10!560!}\right)}{\frac{800!}{10!790!}}=0.0276$$
$$P(X>1)=1-P(X\le 1)=1-[0.0276+0.1201]=0.8523$$

Section 3-9

3-97. 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) $P(X = 8) = \dfrac{e^{-4}4^8}{8!} = 0.0298$

3-99. $P(X = 0) = e^{-\lambda} = 0.05$. Therefore, $\lambda = -\ln(0.05) = 2.996$.
Consequently, $E(X) = V(X) = 2.996$.

3-101. a) Let X denote the number of flaws in one square meter of cloth. Then, X is a Poisson random variable with $\lambda = 0.1$. $P(X = 2) = \dfrac{e^{-0.1}(0.1)^2}{2!} = 0.0045$

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

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) $P(Y \geq 2) = 1 - P(Y \leq 1) = 1 - P(Y = 0) - P(Y = 1)$

$$= 1 - e^{-1} - e^{-1}$$

$$= 0.2642$$

3-105. 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,

$$P(Y = 10) = \binom{10}{10}(0.3935)^{10}(0.6065)^0 = 8.9x10^{-5}$$

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.

$$P(W = 0) = \binom{10}{0}(0.6065)^0(0.3935)^{10} = 8.9x10^{-5}$$

$$P(W = 1) = \binom{10}{1}(0.6065)^1(0.3935)^9 = 0.001372$$

$$P(W \leq 1) = 0.000089 + 0.001372 = 0.00146$$

Supplemental Exercises

3-107. 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.

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

3-109. Let Y denote the number of calls needed to obtain an answer in less than 30 seconds.

a) $P(Y = 4) = (1-0.75)^3 0.75 = 0.25^3 0.75 = 0.0117$

b) E(Y) = 1/p = 1/0.75 = 1.3333

3-111. 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 λ =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 λ = 2.5. $P(W < 2) = P(W = 0) + P(W = 1) = 0.2873$

3-119. 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) = $\binom{500}{0}(0.02)^0(0.98)^{500} = 4.1 \times 10^{-5}$

b) E(X) = 500(0.02) = 10

c) P(X >2) = 1 – P(X ≤ 1) = 0.9995

3-121. a) P(X ≤ 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) = $2^2(0.2) + 3^2(0.4) + 5^2(0.3) + 8^2(0.1) - (3.9)^2 = 3.09$

3-125. 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 λ = 0.25(8) = 2.

a) P(X ≥ 3) = 1 – P(X ≤ 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 λ = 4, and $P(Y<2) = P(Y \le 1) = e^{-4} + (e^{-4}4^1)/1! = 0.0916$.

3-127. 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 ≥ 1) = 0.9, then P(X = 0) = 0.1. Then $\binom{30}{0}(p)^0(1-p)^{30} = 0.1$, giving 30ln(1–p)=ln(0.1), which results in p = 0.0738.

3-129. a) Let X denote the number of flaws in 50 panels. Then, X is a Poisson random variable with λ = 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 panels.

c.) $P(Y \ge 1) = 1 - P(Y = 0) = 1 - e^{-0.02} = 0.0198$

Let V denote the number of panels with 2 or more flaws. Then V is a binomial random variable with n=50 and p=0.0198

$$P(V \le 2) = \binom{50}{0} 0.0198^{0} (.9802)^{50} + \binom{50}{1} 0.0198^{1} (0.9802)^{49}$$

$$+ \binom{50}{2} 0.0198^{2} (0.9802)^{48} = 0.9234$$

Chapter 4 Selected Problem Solutions

Section 4-2

4-1. 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) $P(3 \le X) = \int_3^{\infty} e^{-x} dx = (-e^{-x})\Big|_3^{\infty} = e^{-3} = 0.0498$

4-3 a) $P(X < 4) = \int_3^4 \frac{x}{8} dx = \frac{x^2}{16}\Big|_3^4 = \frac{4^2 - 3^2}{16} = 0.4375$, because $f_X(x) = 0$ for x < 3.

b) , $P(X > 3.5) = \int_{3.5}^5 \frac{x}{8} dx = \frac{x^2}{16}\Big|_{3.5}^5 = \frac{5^2 - 3.5^2}{16} = 0.7969$ because $f_X(x) = 0$ for x > 5.

c) $P(4 < X < 5) = \int_4^5 \frac{x}{8} dx = \frac{x^2}{16}\Big|_4^5 = \frac{5^2 - 4^2}{16} = 0.5625$

d) $P(X < 4.5) = \int_3^{4.5} \frac{x}{8} dx = \frac{x^2}{16}\Big|_3^{4.5} = \frac{4.5^2 - 3^2}{16} = 0.7031$

e) $P(X > 4.5) + P(X < 3.5) = \int_{4.5}^5 \frac{x}{8} dx + \int_3^{3.5} \frac{x}{8} dx = \frac{x^2}{16}\Big|_{4.5}^5 + \frac{x^2}{16}\Big|_3^{3.5} = \frac{5^2 - 4.5^2}{16} + \frac{3.5^2 - 3^2}{16} = 0.5.$

4-9 a) P(X < 2.25 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

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

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.

Section 4-3

4-11. a) P(X<2.8) = P(X ≤ 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) $P(X > 6) = 1 - F_X(6) = 0$

4-13. 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, $F_X(x) = \begin{cases} 0, x \le 0 \\ 1 - e^{-x}, x > 0 \end{cases}$

4-21. $F(x) = \int_0^x 0.5x dx = \frac{0.5x^2}{2}\Big|_0^x = 0.25x^2$ 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 4-4

4-25. $E(X) = \int_3^5 x\frac{x}{8}dx = \frac{x^3}{24}\Big|_3^5 = \frac{5^3 - 3^3}{24} = 4.083$

$$V(X) = \int_3^5 (x - 4.083)^2 \frac{x}{8}dx = \int_3^5 x^2 \tfrac{x}{8}dx - 4.083^2$$
$$= \frac{x^4}{32}\Big|_3^5 - 4.083^2 = 0.3291$$

4-27. a.) $E(X) = \int_{100}^{120} x\frac{600}{x^2}dx = 600\ln x\Big|_{100}^{120} = 109.39$

$$V(X) = \int_{100}^{120} (x - 109.39)^2 \frac{600}{x^2}dx = 600\int_{100}^{120} 1 - \tfrac{2(109.39)}{x} + \tfrac{(109.39)^2}{x^2}dx$$
$$= 600(x - 218.78\ln x - 109.39^2 x^{-1})\Big|_{100}^{120} = 33.19$$

b.) Average cost per part = \$0.50*109.39 = \$54.70

Section 4-5

4-33. a) f(x)= 2.0 for 49.75 < x < 50.25.
E(X) = (50.25 + 49.75)/2 = 50.0,

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

b) $F(x) = \int_{49.75}^{x} 2.0dx$ 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) $P(X < 50.1) = F(50.1) = 2(50.1) - 99.5 = 0.7$

4-35 $E(X) = \frac{(1.5 + 2.2)}{2} = 1.85\,\text{min}$

$$V(X) = \frac{(2.2 - 1.5)^2}{12} = 0.0408\,\text{min}^2$$

b) $P(X < 2) = \int_{1.5}^{2} \frac{1}{(2.2 - 1.5)} dx = \int_{1.5}^{2} 0.7dx = 0.7x\Big|_{1.5}^{2} = 0.7(.5) = 0.7143$

c.) $F(X) = \int_{1.5}^{x} \frac{1}{(2.2 - 1.5)} dx = \int_{1.5}^{x} 0.7dx = 0.7x\Big|_{1.5}^{x}$ for 1.5 < x < 2.2. Therefore,

$$F(x) = \begin{cases} 0, & x < 1.5 \\ 0.7x - 2.14, & 1.5 \le x < 2.2 \\ 1, & 2.2 \le x \end{cases}$$

Section 4-6

4-41 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

4-43. 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) $P(-2 < X < 8) = P(X < 8) - P(X < -2)$
$= P\left(Z < \frac{8-10}{2}\right) - P\left(Z < \frac{-2-10}{2}\right)$
$= P(Z < -1) - P(Z < -6)$
$= 0.15866.$

4-51. a) $P(X < 45) = P\left(Z < \frac{45-65}{5}\right)$
$= P(Z < -3)$
$= 0.00135$

b) $P(X > 65) = P\left(Z > \frac{65-60}{5}\right)$
$= PZ > 1)$
$= 1 - P(Z < 1)$
$= 1 - 0.841345$
$= 0.158655$

c) $P(X < x) = P\left(Z < \frac{x-60}{5}\right) = 0.99.$

Therefore, $\frac{x-60}{5} = 2.33$ and $x = 71.6$

4-55. a) $P(X > 90.3) + P(X < 89.7)$
$= P\left(Z > \frac{90.3-90.2}{0.1}\right) + P\left(Z < \frac{89.7-90.2}{0.1}\right)$
$= P(Z > 1) + P(Z < -5)$
$= 1 - P(Z < 1) + P(Z < -5)$

$=1 - 0.84134 + 0$
$= 0.15866.$

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) $P(89.7 < X < 90.3) = P\left(\frac{89.7-90}{0.1} < Z < \frac{90.3-90}{0.1}\right)$

$= P(-3 < Z < 3) = 0.9973.$

The yield is 100*0.9973 = 99.73%

4-59. a) $P(X > 0.0026) = P\left(Z > \frac{0.0026-0.002}{0.0004}\right)$

$= P(Z > 1.5)$
$= 1-P(Z < 1.5)$
$= 0.06681.$

b) $P(0.0014 < X < 0.0026) = P\left(\frac{0.0014-0.002}{0.0004} < Z < \frac{0.0026-0.002}{0.0004}\right)$

$= P(-1.5 < Z < 1.5)$
$= 0.86638.$

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, $P\left(Z < \frac{0.0006}{\sigma}\right) = 0.9975$. Therefore, $\frac{0.0006}{\sigma} = 2.81$ and $\sigma = 0.000214$.

<u>Section 4-7</u>

4-67 Let X denote the number of errors on a web site. Then, X is a binomial random variable with p = 0.05 and n = 100. Also, E(X) = 100 (0.05) = 5 and V(X) = 100(0.05)(0.95) = 4.75

$$P(X \ge 1) \cong P\left(Z \ge \frac{1-5}{\sqrt{4.75}}\right) = P(Z \ge -1.84) = 1 - P(Z < -1.84) = 1 - 0.03288 = 0.96712$$

4-69 Let X denote the number of hits to a web site. Then, X is a Poisson random variable with a of mean 10,000 per day. $E(X) = \lambda = 10,000$ and $V(X) = 10,000$

a)

$$P(X \ge 10{,}200) \cong P\left(Z \ge \frac{10{,}200 - 10{,}000}{\sqrt{10{,}000}}\right) = P(Z \ge 2) = 1 - P(Z < 2)$$
$$= 1 - 0.9772 = 0.0228$$

Expected value of hits days with more than 10,200 hits per day is
(0.0228)*365=8.32 days per year

b.) Let Y denote the number of days per year with over 10,200 hits to a web site.
Then, Y is a binomial random variable with n=365 and p=0.0228.
E(Y) = 8.32 and V(Y) = 365(0.0228)(0.9772)=8.13

$$P(Y > 15) \cong P\left(Z \geq \frac{15-8.32}{\sqrt{8.13}}\right) = P(Z \geq 2.34) = 1 - P(Z < 2.34)$$
$$= 1 - 0.9904 = 0.0096$$

Section 4-9

4-77. 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 is equal to 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.

4-79. Let X denote the time to failure (in hours) of fans in a personal computer. Then, X is an exponential random variable and $\lambda = 1/E(X) = 0.0003$.

a) P(X > 10,000) = $\int_{10,000}^{\infty} 0.0003 e^{-x0.0003} dx = -e^{-x0.0003} \Big|_{10,000}^{\infty} = e^{-3} = 0.0498$

b) P(X < 7,000) = $\int_{0}^{7,000} 0.0003 e^{-x0.0003} dx = -e^{-x0.0003} \Big|_{0}^{7,000} = 1 - e^{-2.1} = 0.8775$

4-81. 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) = $\int_{0}^{10} 0.1e^{-0.1x}dx = -e^{-0.1x}\Big|_{0}^{10} = 1 - e^{-1} = 0.6321$

4-83. 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.

4-87. 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.

$$P(X \geq 4) = 1 - P(X \leq 3) = 1 - \left[\frac{e^{-6}6^0}{0!} + \frac{e^{-6}6^1}{1!} + \frac{e^{-6}6^2}{2!} + \frac{e^{-6}6^3}{3!}\right] = 0.8488$$

Section 4-10

4-97. 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 - $P(Y \leq 2) = 1 - \frac{e^{-5}5^0}{0!} - \frac{e^{-5}5^1}{1!} - \frac{e^{-5}5^2}{2!} = 0.8754$.

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) = $\binom{10}{10}0.8754^{10}(1 - 0.8754)^0 = 0.2643$.

4-101. 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) = $\dfrac{r}{\lambda} = 5\times10^5$ bits.

b) V(X) = $\dfrac{r}{\lambda^2} = 5\times10^{10}$ and $\sigma_X = \sqrt{5\times10^{10}} = 223607$ 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.

$$P(Y\geq 3) = 1 - P(Y\leq 2) = 1 - \left[\tfrac{e^{-1}1^0}{0!} + \tfrac{e^{-1}1^1}{1!} + \tfrac{e^{-1}1^2}{2!}\right] = 0.0803$$

4-105. a) $\Gamma(6) = 5! = 120$

b) $\Gamma(\frac{5}{2}) = \frac{3}{2}\Gamma(\frac{3}{2}) = \frac{3}{2}\frac{1}{2}\Gamma(\frac{1}{2}) = \frac{3}{4}\pi^{1/2} = 1.32934$

c) $\Gamma(\frac{9}{2}) = \frac{7}{2}\Gamma(\frac{7}{2}) = \frac{7}{2}\frac{5}{2}\frac{3}{2}\frac{1}{2}\Gamma(\frac{1}{2}) = \frac{105}{16}\pi^{1/2} = 11.6317$

Section 4-11

4-109. β=0.2 and δ=100 hours

$$E(X) = 100\Gamma(1+\tfrac{1}{0.2}) = 100\times 5! = 12{,}000$$

$$V(X) = 100^2\Gamma(1+\tfrac{2}{0.2}) - 100^2[\Gamma(1+\tfrac{1}{0.2})]^2 = 3.61\times10^{10}$$

4-111. Let X denote lifetime of a bearing. β=2 and δ=10000 hours

a) $P(X>8000) = 1 - F_X(8000) = e^{-\left(\frac{8000}{10000}\right)^2} = e^{-0.8^2} = 0.5273$

b)

$$E(X) = 10000\Gamma(1+\tfrac{1}{2}) = 10000\Gamma(1.5)$$

$$= 10000(0.5)\Gamma(0.5) = 5000\sqrt{\pi} = 8862.3$$

= 8862.3 hours

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.

$$P(Y=10) = \binom{10}{10}0.5273^{10}(1-0.5273)^0 = 0.00166.$$

Section 4-12

4-117 X is a lognormal distribution with θ=5 and ω^2=9

a.)

$$P(X<13300) = P(e^W < 13300) = P(W < \ln(13300)) = \Phi\left(\frac{\ln(13330)-5}{3}\right)$$

$$= \Phi(1.50) = 0.9332$$

b.) Find the value for which P(X≤x)=0.95

$$P(X \le x) = P(e^W \le x) = P(W < \ln(x)) = \Phi\left(\frac{\ln(x) - 5}{3}\right) = 0.95$$

$$\frac{\ln(x) - 5}{3} = 1.65 \quad x = e^{1.65(3)+5} = 20952.2$$

c.) $\mu = E(X) = e^{\theta + \omega^2/2} = e^{5+9/2} = e^{9.5} = 13359.7$

$$V(X) = e^{2\theta + \omega^2}(e^{\omega^2} - 1) = e^{10+9}(e^9 - 1) = e^{19}(e^9 - 1) = 1.45 \times 10^{12}$$

4-119 a.) X is a lognormal distribution with θ=2 and ω^2=4

$$P(X < 500) = P(e^W < 500) = P(W < \ln(500)) = \Phi\left(\frac{\ln(500) - 2}{2}\right)$$

$$= \Phi(2.11) = 0.9826$$

b.)

$$P(X < 15000 \mid X > 1000) = \frac{P(1000 < X < 1500)}{P(X > 1000)}$$

$$= \frac{\left[\Phi\left(\frac{\ln(1500) - 2}{2}\right) - \Phi\left(\frac{\ln(1000) - 2}{2}\right)\right]}{\left[1 - \Phi\left(\frac{\ln(1000) - 2}{2}\right)\right]}$$

$$= \frac{\Phi(2.66) - \Phi(2.45)}{(1 - \Phi(2.45))} = \frac{0.9961 - 0.9929}{(1 - 0.9929)} = 0.0032 / 0.007 = 0.45$$

c.) The product has degraded over the first 1000 hours, so the probability of it lasting another 500 hours is very low.

4-121 Find the values of θand ω^2 given that E(X) = 100 and V(X) = 85,000

$x = \frac{100}{\sqrt{y}}$ $\qquad 85000 = e^{2\theta+\omega^2}(e^{\omega^2}-1)$

let $x = e^{\theta}$ and $y = e^{\omega^2}$ then (1)$100 = x\sqrt{y}$ and (2) $85000 = x^2y(y-1) = x^2y^2 - x^2y$

Square (1) $10000 = x^2y$ and substitute into (2)

$85000 = 10000(y-1)$

$y = 9.5$

Substitute y into (1) and solve for x $x = \frac{100}{\sqrt{9.5}} = 32.444$

$\theta = \ln(32.444) = 3.45$ and $\omega^2 = \ln(9.5) = 2.25$

Supplemental Exercises

4-127. 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.

4-129. a) Let Y denote the number of calls in 30 minutes. Then, Y is a Poisson random variable with $x = e^{\theta}$. $\qquad P(Y \le 2) = \frac{e^{-3}3^0}{0!} + \frac{e^{-3}3^1}{1!} + \frac{e^{-3}3^2}{2!} = 0.423$.

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

4-137. Let X denote the thickness.

a) P(X > 5.5) = $P\left(Z > \frac{5.5-5}{0.2}\right)$ = P(Z > 2.5) = 0. 0062

b) P(4.5 < X < 5.5) = $P\left(\frac{4.5-5}{0.2} < Z < \frac{5.5-5}{0.2}\right)$ = P (-2.5 < Z < 2.5) = 0.9876

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 $P\left(Z > \dfrac{x-5}{0.2}\right)$ = 0.9. Therefore, $\dfrac{x-5}{0.2}$ = 1.65 and x = 5.33.

4-139. 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.

4-141. If P(X > 10,000) = 0.99, then P(Z > $\frac{10{,}000-\mu}{600}$) = 0.99. Therefore, $\frac{10{,}000-\mu}{600}$ = -2.33 and $\mu = 11{,}398$.

4-143 X is an exponential distribution with E(X) = 7000 hours

a.) $P(X < 5800) = \int_0^{5800} \frac{1}{7000} e^{-\frac{x}{7000}} dx = 1 - e^{-\left(\frac{5800}{7000}\right)} = 0.5633$

b.) $P(X > x) = \int_x^{\infty} \frac{1}{7000} e^{-\frac{x}{7000}} dx = 0.9$ Therefore, $e^{-\frac{x}{7000}} = 0.9$

and $x = -7000\ln(0.9) = 737.5$ hours

Chapter 5 Selected Problem Solutions

Section 5-1

5-7.

$$E(X) = 1[f_{XY}(1,1) + f_{XY}(1,2) + f_{XY}(1,3)] + 2[f_{XY}(2,1) + f_{XY}(2,2) + f_{XY}(2,3)]$$
$$+ \ 3[f_{XY}(3,1) + f_{XY}(3,2) + f_{XY}(3,3)]$$
$$= \left(1\times\tfrac{9}{36}\right) + \left(2\times\tfrac{12}{36}\right) + \left(3\times\tfrac{15}{36}\right) = 13/6 = 2.167$$
$$V(X) = (1-\tfrac{13}{6})^2\ \tfrac{9}{36} + (2-\tfrac{13}{6})^2\ \tfrac{12}{36} + (3-\tfrac{13}{6})^2\ \tfrac{15}{36} = 0.639$$
$$E(Y) = 2.167$$
$$V(Y) = 0.639$$

5-11.

$$E(X) = -1(\tfrac{1}{8}) - 0.5(\tfrac{1}{4}) + 0.5(\tfrac{1}{2}) + 1(\tfrac{1}{8}) = \tfrac{1}{8}$$
$$E(Y) = -2(\tfrac{1}{8}) - 1(\tfrac{1}{4}) + 1(\tfrac{1}{2}) + 2(\tfrac{1}{8}) = \tfrac{1}{4}$$

5-15 a) The range of (X,Y) is $X \geq 0,\ \ Y \geq 0\ \ and\ \ X + Y \leq 4$. X is the number of pages with moderate graphic content and Y is the number of pages with high graphic output out of 4.

	x=0	x=1	x=2	x=3	x=4
y=4	$5.35\text{x}10^{-05}$	0	0	0	0
y=3	0.00183	0.00092	0	0	0
y=2	0.02033	0.02066	0.00499	0	0
y=1	0.08727	0.13542	0.06656	0.01035	0
y=0	0.12436	0.26181	0.19635	0.06212	0.00699

b.)

	x=0	x=1	x=2	x=3	x=4
f(x)	0.2338	0.4188	0.2679	0.0725	0.0070

c.)
E(X)=

$$\sum_{0}^{4} x_i f(x_i) = 0(0.2338) + 1(0.4188) + 2(0.2679) + 3(0.7248) = 4(0.0070) = 1.2$$

d.) $f_{Y|3}(y) = \dfrac{f_{XY}(3,y)}{f_X(3)}$, $f_x(3) = 0.0725$

y	$f_{Y\|3}(y)$
0	0.857
1	0.143
2	0
3	0
4	0

e) E(Y|X=3) = 0(0.857)+1(0.143) = 0.143

Section 5-2

5-17. a) $P(X=2) = f_{XYZ}(2,1,1) + f_{XYZ}(2,1,2) + f_{XYZ}(2,2,1) + f_{XYZ}(2,2,2) = 0.5$

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)

$$P(X=1 \text{ or } Z=1) = P(X=1) + P(Z=1) - P(X=1, Z=1) = 0.5 + 0.5 - 0.2 = 0.8$$

e) E(X) = 1(0.5) + 2(0.5) = 1.5

5-25. 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 on the CD material on counting techniques, 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) $P(X=1, Y=2, Z=1) = \dfrac{\binom{4}{1}\binom{5}{2}\binom{6}{1}}{\binom{15}{4}} = 0.1758$

b) $P(X=1, Y=1) = P(X=1, Y=1, Z=2) = \dfrac{\binom{4}{1}\binom{5}{1}\binom{6}{2}}{\binom{15}{4}} = 0.2198$

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.

5-29 a.) $P(X=2 \mid Y=2) = \dfrac{P(X=2, Y=2)}{P(Y=2)} = \dfrac{0.1944}{0.2646} = 0.7347$

$$P(X=2, Y=2) = 0.1922$$

$$P(Y=2) = \binom{4}{2} 0.3^2 0.7^4 = 0.2646 \qquad \text{from the binomial marginal distribution of } Y$$

b) Not possible, $x+y+z=4$, the probability is zero.

c.) $P(X \mid Y=2) = P(X=0 \mid Y=2), P(X=1 \mid Y=2), P(X=2 \mid Y=2)$

$$P(X=0 \mid Y=2) = \frac{P(X=0, Y=2)}{P(Y=2)} = \left(\frac{4!}{0!2!2!} 0.6^0 0.3^2 0.1^2\right) \Big/ 0.2646 = 0.0204$$

$$P(X=1 \mid Y=2) = \frac{P(X=1, Y=2)}{P(Y=2)} = \left(\frac{4!}{1!2!1!} 0.6^1 0.3^2 0.1^1\right) \Big/ 0.2646 = 0.2449$$

$$P(X=2 \mid Y=2) = \frac{P(X=2, Y=2)}{P(Y=2)} = \left(\frac{4!}{2!2!0!} 0.6^2 0.3^2 0.1^0\right) \Big/ 0.2646 = 0.7347$$

d.) E(X|Y=2)=0(0.0204)+1(0.2449)+2(0.7347) = 1.7142

5-31 a.), 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.

b.) first find $P(X \mid Y=2)$

$$P(Y=2) = P(X=1, Y=2, Z=0) + P(X=0, Y=2, Z=1)$$

$$= \frac{3!}{1!2!0!} 0.01(0.04)^2 0.95^0 + \frac{3!}{0!2!1!} 0.01^0 (0.04)^2 0.95^1 = 0.0046$$

$$P(X=0 \mid Y=2) = \frac{P(X=0, Y=2)}{P(Y=2)} = \left(\frac{3!}{0!2!1!} 0.01^0 0.04^2 0.95^1 \right) \Big/ 0.004608 = 0.98958$$

$$P(X=1 \mid Y=2) = \frac{P(X=1, Y=2)}{P(Y=2)} = \left(\frac{3!}{1!2!1!} 0.01^1 0.04^2 0.95^0 \right) \Big/ 0.004608 = 0.01042$$

$$E(X \mid Y=2) = 0(0.98958) + 1(0.01042) = 0.01042$$

$$V(X \mid Y=2) = E(X^2) - (E(X))^2 = 0.01042 - (0.01042)^2 = 0.01031$$

Section 5-3

5-35. a) $P(X<2, Y<3) = \frac{4}{81} \int_0^3 \int_0^2 xydxdy = \frac{4}{81}(2) \int_0^3 ydy = \frac{4}{81}(2)(\frac{9}{2}) = 0.4444$

b) P(X < 2.5) = P(X < 2.5, Y < 3) because the range of Y is from 0 to 3.

$$P(X<2.5, Y<3) = \frac{4}{81} \int_0^3 \int_0^{2.5} xydxdy = \frac{4}{81}(3.125) \int_0^3 ydy = \frac{4}{81}(3.125)\frac{9}{2} = 0.6944$$

c) $P(1<Y<2.5) = \frac{4}{81} \int_1^{2.5} \int_0^3 xydxdy = \frac{4}{81}(4.5) \int_1^{2.5} ydy = \frac{18}{81} \frac{y^2}{2} \Big|_1^{2.5} = 0.5833$

d) $P(X>1.8, 1<Y<2.5) = \frac{4}{81} \int_1^{2.5} \int_{1.8}^3 xydxdy = \frac{4}{81}(2.88) \int_1^{2.5} ydy = \frac{4}{81}(2.88)\frac{(2.5^2-1)}{2} = 0.3733$

e) $E(X) = \frac{4}{81} \int_0^3 \int_0^3 x^2 ydxdy = \frac{4}{81} \int_0^3 9ydy = \frac{4}{9} \frac{y^2}{2} \Big|_0^3 = 2$

f) $P(X<0, Y<4) = \frac{4}{81} \int_0^4 \int_0^0 xydxdy = 0 \int_0^4 ydy = 0$

5-37.

$$c\int_0^3\int_x^{x+2}(x+y)\,dy\,dx = \int_0^3 xy + \frac{y^2}{2}\Big|_x^{x+2}\,dx$$

$$= \int_0^3 \left[x(x+2) + \frac{(x+2)^2}{2} - x^2 - \frac{x^2}{2}\right]dx$$

$$= c\int_0^3 (4x+2)\,dx = \left[2x^2 + 2x\right]_0^3 = 24c$$

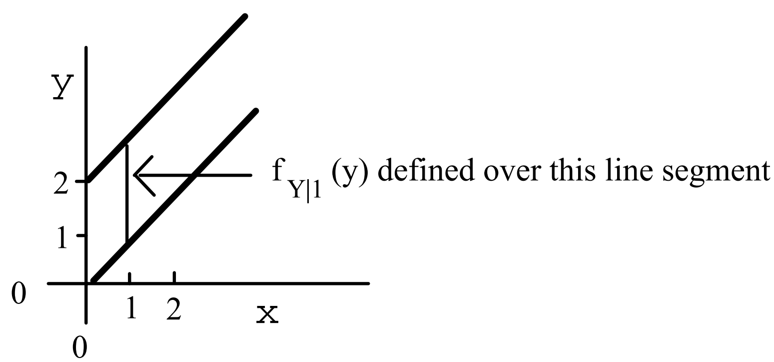
Therefore, c = 1/24.

5-39. a) $f_X(x)$ is the integral of $f_{XY}(x,y)$ over the interval from x to x+2. That is,

$$f_X(x) = \frac{1}{24}\int_x^{x+2}(x+y)\,dy = \frac{1}{24}\left[xy + \frac{y^2}{2}\Big|_x^{x+2}\right] = \frac{x}{6} + \frac{1}{12} \quad \text{for } 0 < x < 3.$$

b) $f_{Y|1}(y) = \frac{f_{XY}(1,y)}{f_X(1)} = \dfrac{\frac{1}{24}(1+y)}{\frac{1}{6}+\frac{1}{12}} = \dfrac{1+y}{6}$ for 1 < y < 3.

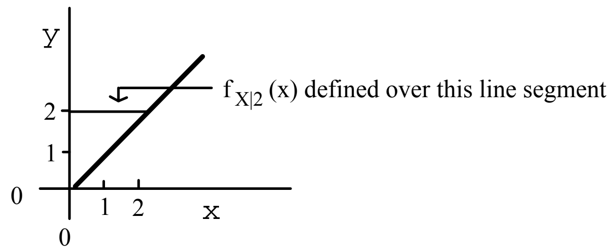
See the following graph,



c) E(Y|X=1) = $\displaystyle\int_1^3 y\left(\frac{1+y}{6}\right)dy = \frac{1}{6}\int_1^3 (y+y^2)\,dy = \frac{1}{6}\left(\frac{y^2}{2}+\frac{y^3}{3}\right)\Bigg|_1^3 = 2.111$

d.) $\displaystyle P(Y > 2 \mid X = 1) = \int_2^3 \left(\frac{1+y}{6}\right)dy = \frac{1}{6}\int_1^3 (1+y)\,dy = \frac{1}{6}\left(y + \frac{y^2}{2}\right)\Bigg|_1^2 = 0.4167$

e.) $f_{X|2}(x) = \frac{f_{XY}(x,2)}{f_Y(2)}$. Here $f_Y(y)$ is determined by integrating over *x*. There are three regions of integration. For $0 < y \le 2$ the integration is from 0 to *y*. For $2 < y \le 3$ the integration is from *y*-2 to *y*. For $3 < y < 5$ the integration is from *y* to 3. Because the condition is *x*=2, only the first integration is needed. $\displaystyle f_Y(y) = \frac{1}{24}\int_0^y (x+y)\,dx = \frac{1}{24}\left[\frac{x^2}{2} + xy\Big|_0^y\right] = \frac{y^2}{16}$ for $0 < y \le 2$.



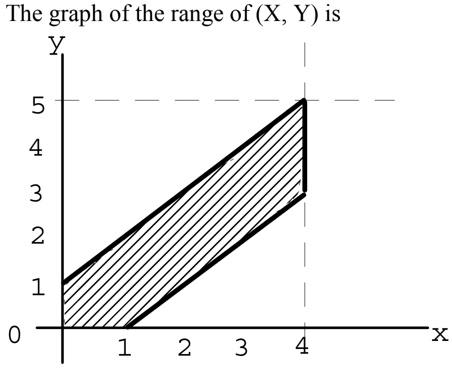
Therefore, $f_Y(2) = 1/4$ and $f_{X|2}(x) = \dfrac{\frac{1}{24}(x+2)}{1/4} = \dfrac{x+2}{6}$ for 0 < x < 2

5-43. Solve for c

$$c\int_0^{\infty}\int_0^{x} e^{-2x-3y}\,dydx = \frac{c}{3}\int_0^{\infty} e^{-2x}\left(1-e^{-3x}\right)dx = \frac{c}{3}\int_0^{\infty} e^{-2x} - e^{-5x}\,dx$$

$$= \frac{c}{3}\left(\frac{1}{2}-\frac{1}{5}\right) = \frac{1}{10}c. \quad c = 10$$

5-49. The graph of the range of (X, Y) is



$$\int_0^{1}\int_0^{x+1} c\,dydx + \int_1^{4}\int_{x-1}^{x+1} c\,dydx = 1$$

$$= c\int_0^{1}(x+1)dx + 2c\int_1^{4} dx$$

$$= \tfrac{3}{2}c + 6c = 7.5c = 1$$

Therefore, c = 1/7.5=2/15

5-51. a.)

$$f(x) = \int_0^{x+1} \frac{1}{7.5}\,dy = \left(\frac{x+1}{7.5}\right) \quad \text{for} \quad 0 < x < 1,$$

$$f(x) = \int_{x-1}^{x+1} \frac{1}{7.5}\,dy = \left(\frac{x+1-(x-1)}{7.5}\right) = \frac{2}{7.5} \quad \text{for} \quad 1 < x < 4$$

b.)

$$f_{Y|X=1}(y) = \frac{f_{XY}(1,y)}{f_X(1)} = \frac{1/7.5}{2/7.5} = 0.5$$

$$f_{Y|X=1}(y) = 0.5 \quad \text{for } 0 < y < 2$$

c.) $E(Y \mid X=1) = \int_0^2 \frac{y}{2}\,dy = \left.\frac{y^2}{4}\right|_0^2 = 1$

d.) $P(Y < 0.5 \mid X=1) = \int_0^{0.5} 0.5\,dy = 0.5y\Big|_0^{0.5} = 0.25$

5-53 a.) μ=3.2 λ=1/3.2

$$P(X > 5, Y > 5) = 10.24\int_5^\infty\int_5^\infty e^{-\frac{x}{3.2}-\frac{y}{3.2}}\,dydx = 3.2\int_5^\infty e^{-\frac{x}{3.2}}\left(e^{-\frac{5}{3.2}}\right)dx$$

$$= \left(e^{-\frac{5}{3.2}}\right)\left(e^{-\frac{5}{3.2}}\right) = 0.0439$$

$$P(X > 10, Y > 10) = 10.24\int_{10}^\infty\int_{10}^\infty e^{-\frac{x}{3.2}-\frac{y}{3.2}}\,dydx = 3.2\int_{10}^\infty e^{-\frac{x}{3.2}}\left(e^{-\frac{10}{3.2}}\right)dx$$

$$= \left(e^{-\frac{10}{3.2}}\right)\left(e^{-\frac{10}{3.2}}\right) = 0.0019$$

b.) Let X denote the number of orders in a 5-minute interval. Then X is a Poisson random variable with λ=5/3.2 = 1.5625.

$$P(X = 2) = \frac{e^{-1.5625}(1.5625)^2}{21} = 0.256$$

For both systems, $P(X=2)P(X=2) = 0.256^2 = 0.0655$

c.) The joint probability distribution is not necessary because the two processes are independent and we can just multiply the probabilities.

Section 5-4

5-55. 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)

$$P(X<0.5, Y<0.5)=\int_0^{0.5}\int_0^{0.5}\int_0^{1}(8xyz)dzdydx$$

$$=\int_0^{0.5}\int_0^{0.5}(4xy)dydx=\int_0^{0.5}(0.5x)dx=\left.\tfrac{x^2}{4}\right|_0^{0.5}=0.0625$$

c) P(Z < 2) = 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) $E(X)=\int_0^1\int_0^1\int_0^1(8x^2yz)dzdydx=\int_0^1(2x^2)dx=\tfrac{2x^3}{3}=2/3$

5-57. a) $f_{YZ}(y,z)=\int_0^1(8xyz)dx=4yz$ for 0 < y < 1 and 0 < z < 1.

Then, $f_{X|YZ}(x)=\dfrac{f_{XYZ}(x,y,z)}{f_{YZ}(y,z)}=\dfrac{8x(0.5)(0.8)}{4(0.5)(0.8)}=2x$ for 0 < x < 1.

b) Therefore, $P(X<0.5|Y=0.5, Z=0.8)=\int_0^{0.5}2xdx=0.25$

5-61 Determine c such that $f(xyz)=c$ is a joint density probability over the region x>0, y>0 and z>0 with x+y+z<1

$$f(xyz)=c\int_0^1\int_0^{1-x}\int_0^{1-x-y}dzdydx=\int_0^1\int_0^{1-x}c(1-x-y)dydx=\int_0^1\left(\left.c(y-xy-\frac{y^2}{2})\right|_0^{1-x}\right)dx$$

$$=\int_0^1 c\left((1-x)-x(1-x)-\frac{(1-x)^2}{2}\right)dx=\int_0^1 c\left(\frac{(1-x)^2}{2}\right)dx=c\left.\left(\frac{1}{2}x-\frac{x^2}{2}+\frac{x^3}{6}\right)\right|_0^1$$

$=c\dfrac{1}{6}$. Therefore, c = 6.

5-63 a.)

$$f(x)=6\int_0^{1-x}\int_0^{1-x-y}dzdy=\int_0^{1-x}6(1-x-y)dy=\left.\left(y-xy-\frac{y^2}{2}\right)\right|_0^{1-x}$$

$$=6(\frac{x^2}{2}-x+\frac{1}{2})=3(x-1)^2 \text{ for } 0<x<1$$

b.)

$$f(x,y)=6\int_0^{1-x-y}dz=6(1-x-y)$$

for $x>0,\ y>0$ *and* $x+y<1$

c.)

$$f(x \mid y=0.5, z=0.5)=\frac{f(x, y=0.5, z=0,5)}{f(y=0.5, z=0.5)}=\frac{6}{6}=1 \text{ For, } x=0$$

d.) The marginal $f_Y(y)$ is similar to $f_X(x)$ and $f_Y(y)=3(1-y)^2$ for $0<y<1$.

$$f_{X|Y}(x \mid 0.5)=\frac{f(x,0.5)}{f_Y(0.5)}=\frac{6(0.5-x)}{3(0.25)}=4(1-2x) \text{ for } x<0.5$$

5-65. 5-65. 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

$P(A)=1-[P(Z<3)]^{20}=1-0.99865^{20}=0.0267$.

Section 5-5

5-67. E(X) = 1(3/8)+2(1/2)+4(1/8)=15/8 = 1.875

E(Y) = 3(1/8)+4(1/4)+5(1/2)+6(1/8)=37/8 = 4.625

$$\begin{aligned}\mathrm{E(XY)} &= [1\times 3\times(1/8)]+[1\times 4\times(1/4)]+[2\times 5\times(1/2)]+[4\times 6\times(1/8)] \\ &= 75/8 = 9.375\end{aligned}$$

$$\sigma_{XY}=E(XY)-E(X)E(Y)=9.375-(1.875)(4.625)=0.703125$$

$V(X) = 1^2(3/8)+2^2(1/2)+4^2(1/8)-(15/8)^2 = 0.8594$

$V(Y) = 3^2(1/8)+4^2(1/4)+5^2(1/2)+6^2(1/8)-(15/8)^2 = 0.7344$

$$\rho_{XY}=\frac{\sigma_{XY}}{\sigma_X\sigma_Y}=\frac{0.703125}{\sqrt{(0.8594)(0.7344)}}=0.8851$$

5-69.

$$\sum_{x=1}^{3}\sum_{y=1}^{3}c(x+y)=36c, \qquad c=1/36$$

$$E(X)=\frac{13}{6} \qquad E(Y)=\frac{13}{6} \qquad E(XY)=\frac{14}{3} \qquad \sigma_{xy}=\frac{14}{3}-\left(\frac{13}{6}\right)^2=\frac{-1}{36}$$

$$E(X^2)=\frac{16}{3} \qquad E(Y^2)=\frac{16}{3} \qquad V(X)=V(Y)=\frac{23}{36}$$

$$\rho=\frac{\frac{-1}{36}}{\sqrt{\frac{23}{36}}\sqrt{\frac{23}{36}}}=-0.0435$$

5-73. $E(X)=\frac{2}{19}\int_0^1\int_0^{x+1} xdydx+\frac{2}{19}\int_1^5\int_{x-1}^{x+1} xdydx=2.614$

$$E(Y)=\frac{2}{19}\int_0^1\int_0^{x+1} ydydx+\frac{2}{19}\int_1^5\int_{x-1}^{x+1} ydydx=2.649$$

Now, $E(XY)=\frac{2}{19}\int_0^1\int_0^{x+1} xydydx+\frac{2}{19}\int_1^5\int_{x-1}^{x+1} xydydx=8.7763$

$$\sigma_{xy}=8.7763-(2.614)(2.649)=1.85181$$

$$E(X^2)=8.7632 \qquad E(Y^2)=9.07895$$

$$V(x)=1.930, \qquad V(Y)=2.062$$

$$\rho=\frac{1.852}{\sqrt{1.930}\sqrt{2.062}}=0.9279$$

Section 5-6

5-81. Because $\rho = 0$ and X and Y are normally distributed, X and Y are independent. Therefore,

$\mu_X = 0.1$mm σ_X=0.00031mm $\mu_Y = 0.23$mm σ_Y=0.00017mm

Probability X is within specification limits is

$$P(0.099535<X<0.100465)=P\left(\frac{0.099535-0.1}{0.00031}<Z<\frac{0.100465-0.1}{0.00031}\right)$$
$$=P(-1.5<Z<1.5)=P(Z<1.5)-P(Z<-1.5)$$
$$=0.8664$$

Probability that Y is within specification limits is

$$P(0.22966<X<0.23034)=P\left(\frac{0.22966-0.23}{0.00017}<Z<\frac{0.23034-0.23}{0.00017}\right)$$
$$=P(-2<Z<2)=P(Z<2)-P(Z<-2)$$
$$=0.9545$$

Probability that a randomly selected lamp is within specification limits is (0.8664)(.9594)=0.8270

Section 5-7

5-87. 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,

$$P(2X+3Y<30)=P(Z<\tfrac{30-30}{\sqrt{97}})=P(Z<0)=0.5$$

d) $P(2X+3Y<40)=P(Z<\tfrac{40-30}{\sqrt{97}})=P(Z<1.02)=0.8461$

5-89 a) Let T denote the total thickness. Then, T = X + Y and E(T) = 4 mm,

V(T) = $0.1^2+0.1^2=0.02\text{mm}^2$, and $\sigma_T = 0.1414$ mm.

b)

$$P(T > 4.3) = P\left(Z > \frac{4.3-4}{0.1414}\right) = P(Z > 2.12)$$

2.12) = 1 - 0.983 = 0.017

$$= 1 - P(Z < 2.12) = 1 - 0.983 = 0.0170$$

5-93. 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 $P(\overline{X} < 12) = P\left(Z < \frac{12-12.1}{0.05}\right) = P(Z < -2) = 0.023$

c) P($\overline{X}$ < 12) = 0.005 implies that $P\left(Z < \frac{12-\mu}{0.05}\right) = 0.005.$

Then $\frac{12-\mu}{0.05}$ = -2.58 and $\mu = 12.129$.

d.) P($\overline{X}$ < 12) = 0.005 implies that $P\left(Z < \frac{12-12.1}{\sigma/\sqrt{100}}\right) = 0.005.$

Then $\frac{12-12.1}{\sigma/\sqrt{100}}$ = -2.58 and $\sigma = 0.388$.

e.) P($\overline{X}$ < 12) = 0.01 implies that $P\left(Z < \frac{12-12.1}{0.5/\sqrt{n}}\right) = 0.01.$

Then $\frac{12-12.1}{0.5/\sqrt{n}}$ = -2.33 and $n = 135.72 \cong 136$.

Supplemental Exercises

5-97. a) $P(X < 0.5, Y < 1.5) = f_{XY}(0,1) + f_{XY}(0,0) = 1/8 + 1/4 = 3/8$.

b) $P(X \le 1) = f_{XY}(0,0) + f_{XY}(0,1) + f_{XY}(1,0) + f_{XY}(1,1) = 3/4$

c) $P(Y < 1.5) = f_{XY}(0,0) + f_{XY}(0,1) + f_{XY}(1,0) + f_{XY}(1,1) = 3/4$

d) $P(X > 0.5, Y < 1.5) = f_{XY}(1,0) + f_{XY}(1,1) = 3/8$

e) E(X) = 0(3/8) + 1(3/8) + 2(1/4) = 7/8.
V(X) = 0^2(3/8) + 1^2(3/8) + 2^2(1/4) - 7/8^2 =39/64
E(Y) = 1(3/8) + 0(3/8) + 2(1/4) = 7/8.
. V(Y) = 1^2(3/8) + 0^2(3/8) + 2^2(1/4) - 7/8^2 =39/64

5-105. a) $P(X < 1, Y < 1) = \int_0^1\int_0^1 \frac{1}{18}x^2 y\,dy\,dx = \int_0^1 \frac{1}{18}x^2 \left.\frac{y^2}{2}\right|_0^1 dx = \left.\frac{1}{36}\frac{x^3}{3}\right|_0^1 = \frac{1}{108}$

b) $P(X < 2.5) = \int_0^{2.5}\int_0^2 \frac{1}{18}x^2 y\,dy\,dx = \int_0^{2.5} \frac{1}{18}x^2 \left.\frac{y^2}{2}\right|_0^2 dx = \left.\frac{1}{9}\frac{x^3}{3}\right|_0^{2.5} = 0.5787$

c) $P(1 < Y < 2.5) = \int_0^3\int_1^2 \frac{1}{18}x^2 y\,dy\,dx = \int_0^3 \frac{1}{18}x^2 \left.\frac{y^2}{2}\right|_1^2 dx = \left.\frac{1}{12}\frac{x^3}{3}\right|_0^3 = \frac{3}{4}$

d)

$$P(X>2,1<Y<1.5)=\int_2^3\int_1^{1.5}\tfrac{1}{18}x^2y\,dy\,dx=\int_2^3\tfrac{1}{18}x^2\left.\tfrac{y^2}{2}\right|_1^{1.5}dx=\tfrac{5}{144}\left.\tfrac{x^3}{3}\right|_2^3$$

$$=\tfrac{95}{432}=0.2199$$

e) $E(X)=\int_0^3\int_0^2\tfrac{1}{18}x^3y\,dy\,dx=\int_0^3\tfrac{1}{18}x^3\,2dx=\tfrac{1}{9}\left.\tfrac{x^4}{4}\right|_0^3=\tfrac{9}{4}$

f) $E(Y)=\int_0^3\int_0^2\tfrac{1}{18}x^2y^2\,dy\,dx=\int_0^3\tfrac{1}{18}x^2\,\tfrac{8}{3}\,dx=\tfrac{4}{27}\left.\tfrac{x^3}{3}\right|_0^3=\tfrac{4}{3}$

5-107. 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 $f_{XYZ}(x,y,z)=\dfrac{1}{4\pi}$ 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) $f_{XY|1}(x,y)=\dfrac{f_{XYZ}(x,y,1)}{f_Z(1)}$ and $f_Z(z)=\iint\limits_{x^2+y^2\le 1}\tfrac{1}{4\pi}dy\,dx=1/4$

for $0<z<4$. Then, $f_{XY|1}(x,y)=\dfrac{1/4\pi}{1/4}=\tfrac{1}{\pi}\quad for\ x^2+y^2\le 1$.

d) $f_X(x)=\int_0^4\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}}\tfrac{1}{4\pi}dy\,dz=\int_0^4\tfrac{1}{2\pi}\sqrt{1-x^2}\,dz=\tfrac{2}{\pi}\sqrt{1-x^2}$ for $-1<x<1$

5-111. 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.

5-115. 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.

5-119 Let T denote the total thickness. Then, T = X_1+X_2 and

a.) E(T) = 0.5+1=1.5 mm

$V(T)=V(X_1)+V(X_2)+2Cov(X_1X_2)=0.01+0.04+2(0.14)=0.078\text{mm}^2$

where $Cov(XY)=\rho\sigma_X\sigma_Y=0.7(0.1)(0.2)=0.014$

b.) $P(T < 1) = P\left(Z < \frac{1-1.5}{0.078}\right) = P(Z < -6.41) \cong 0$

c.) Let P denote the total thickness. Then, $P = 2X_1 + 3X_2$ and
$E(P) = 2(0.5)+3(1)=4$ mm
$V(P)=4V(X_1)+9V(X_2) + 2(2)(3)Cov(X_1X_2)=4(0.01)+9(0.04)+2(2)(3)(0.014)=0.568mm^2$
where $Cov(XY)=\rho\sigma_X\sigma_Y=0.7(0.1)(0.2)=0.014$

5-121 Let X and Y denote the percentage returns for security one and two respectively.
If ½ of the total dollars is invested in each then ½X+ ½Y is the percentage return.
E(½X+ ½Y)=5 million
V(½X+ ½Y)=1/4 V(X)+1/4V(Y)-2(1/2)(1/2)Cov(X,Y)
where $Cov(XY)=\rho\sigma_X\sigma_Y=-0.5(2)(4)=-4$
V(½X+ ½Y)=1/4(4)+1/4(6)-2=3
Also, E(X)=5 and V(X) = 4. Therefore, the strategy that splits between the securities has a lower standard deviation of percentage return.

Chapter 6 Selected Problem Solutions

Sections 6-1and 6-2

6-1. Sample average:

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

Sample variance:

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

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

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

$$= \frac{0.0001569}{7} = 0.000022414 \quad (\text{mm})^2$$

Sample standard deviation:

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

The sample standard deviation could also be found using

$$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}} \quad \text{where} \quad \sum_{i=1}^{8}(x_i - \bar{x})^2 = 0.0001569$$

Dot Diagram:

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

There appears to be a possible outlier in the data set.

6-11. a) $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{5747}{8} = 7.184$

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, during which the pH of the solution could change, and operator skill in drawing the sample or using the instrument.

6-13. 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
$\bar{x} = 66.86$
$s = 10.74$

Section 6-3

6-15 a.) Stem-and-leaf display for Problem 6-15 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 out of 70 coupons survived beyond 2000 cycles.

6-19. Descriptive Statistics

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

6-25 Stem-and-leaf display for Problem 6-25. Yard: unit = 1.0
Note: Minitab has dropped the value to the right of the decimal to make this display.

```
   4    23*|2334
   7    23o|677
  15    24*|00112444
  19    24o|5578
  32    25*|0111122334444
  45    25o|5555556677899
 (15)   26*|000011123334444
  40    26o|566677888
  31    27*|0000112222233333444
  12    27o|66788999
   4    28*|003
```

```
1   28o|5
```

Sample Mean $\bar{x} = \dfrac{\sum_{i=1}^{n} x_i}{n} = \dfrac{\sum_{i=1}^{100} x_i}{100} = \dfrac{26070}{100} = 260.7 \quad yards$

Sample Standard Deviation

$$\sum_{i=1}^{100} x_i = 26070 \qquad \text{and} \qquad \sum_{i=1}^{100} x_i^2 = 6813256$$

$$s^2 = \frac{\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}}{n-1} = \frac{6813256 - \frac{(26070)^2}{100}}{100-1} = \frac{16807}{99}$$

$$= 169.7677 \quad yards^2$$

and

$$s = \sqrt{169.7677} = 13.03 \quad yards$$

Sample Median

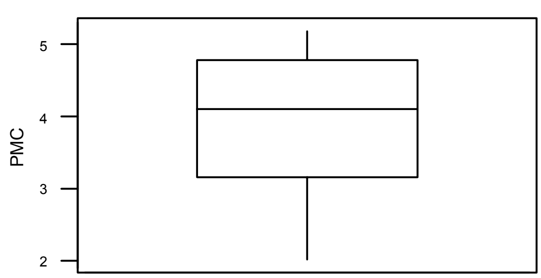
```
Variable              N     Median
yards               100     261.15
```

Section 6-5

6-43. 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)



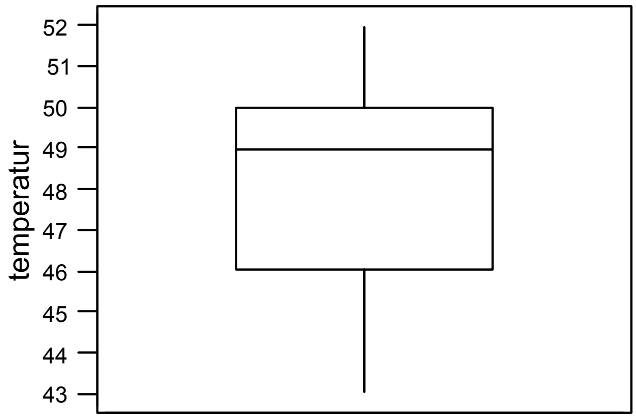
6-47.

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 slightly skewed.

Supplemental

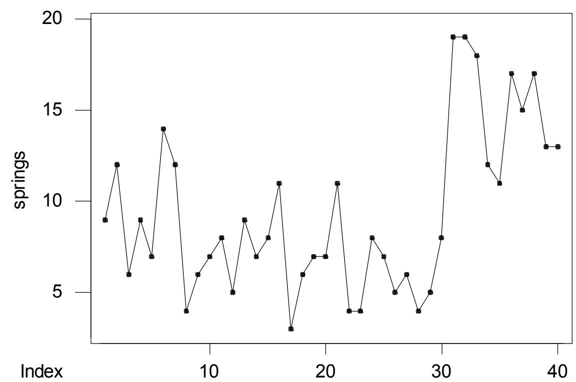
6-75 a) Sample 1 Range = 4
Sample 2 Range = 4
Yes, the two appear to exhibit the same variability
b) Sample 1 s = 1.604
Sample 2 s = 1.852
No, sample 2 has a larger standard deviation.
c) The sample range is a relatively crude measure of the sample variability as compared to the sample standard deviation 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.

6-79 a)Stem-and-leaf display for Problem 6-79: 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.

Chapter 7 Selected Problem Solutions

Section 7-2

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 $V(\hat{\Theta}_2) = 10 = MSE(\hat{\Theta}_2)$

$E(\hat{\Theta}_3) \neq \theta$ Bias $MSE(\hat{\Theta}_3) = 6$ [not that this 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, but it is biased. $\hat{\Theta}_2$ is the best "unbiased" estimator.

7-11

a.) The average of the 26 observations provided can be used as an estimator of the mean pull force since we know it is unbiased. This value is 75.427 pounds.

b.) The median of the sample can be used as an estimate of the point that divides the population into a "weak" and "strong" half. This estimate is 75.1 pounds.

c.) Our estimate of the population variance is the sample variance or 2.214 square pounds. Similarly, our estimate of the population standard deviation is the sample standard deviation or 1.488 pounds.

d.) The standard error of the mean pull force, estimated from the data provided is 0.292 pounds. This value is the standard deviation, not of the pull force, but of the mean pull force of the population.

e.) Only one connector in the sample has a pull force measurement under 73 pounds. Our point estimate for the proportion requested is then 1/26 = 0.0385

7-13 a.) To see if the estimator is unbiased, find:

$$E[(X_{\min} + X_{\max})/2] = \frac{1}{2}[E(X_{\min}) + E(X_{\max})] = \frac{1}{2}(\mu + \mu) = \mu$$

since the expected value of any observation arising from a normally distributed process is equal to the mean. So this is an unbiased estimator of the mean.

b.) The standard error of this estimator is:

$$\sqrt{V[(X_{\min} + X_{\max})/2]} = \frac{1}{2}\sqrt{[V(X_{\min}) + V(X_{\max}) + COV(X_{\min}, X_{\max})]} = \frac{1}{2}\sqrt{(\sigma^2 + \sigma^2)} = \frac{1}{\sqrt{2}}\sigma$$

c.) This estimator is not better than the sample mean because it has larger standard error for n > 2. This is due to the fact that this estimator uses only two observations from the available sample. The sample mean uses all the information available to compute the estimate.

7-17

a) $E(\hat{\mu}) = E(\alpha\overline{X}_1 + (1-\alpha)\overline{X}_2) = \alpha E(\overline{X}_1) + (1-\alpha)E(\overline{X}_2) = \alpha\mu + (1-\alpha)\mu = \mu$

b)

$$s.e.(\hat{\mu}) = \sqrt{V(\alpha\overline{X}_1 + (1-\alpha)\overline{X}_2)} = \sqrt{\alpha^2 V(\overline{X}_1) + (1-\alpha)^2 V(\overline{X}_2)}$$

$$= \sqrt{\alpha^2 \frac{\sigma_1^{\ 2}}{n_1} + (1-\alpha)^2 \frac{\sigma_2^{\ 2}}{n_2}} = \sqrt{\alpha^2 \frac{\sigma_1^{\ 2}}{n_1} + (1-\alpha)^2 a \frac{\sigma_1^{\ 2}}{n_2}}$$

$$= \sigma_1 \sqrt{\frac{\alpha^2 n_2 + (1-\alpha)^2 a n_1}{n_1 n_2}}$$

c) The value of alpha that minimizes the standard error is:

$$\alpha = \frac{an_1}{n_2 + an_1}$$

d) With $a = 4$ and $n_1=2n_2$, the value of alpha to choose is 8/9. The arbitrary value of α=0.5 is too small and will result in a larger standard error. With α=8/9 the standard error is

$$s.e.(\hat{\mu}) = \sigma_1 \sqrt{\frac{(8/9)^2 n_2 + (1/9)^2 8n_2}{2n_2^{\ 2}}} = \frac{0.667\sigma_1}{\sqrt{n_2}}$$

If α=0.05 the standard error is

$$s.e.(\hat{\mu}) = \sigma_1 \sqrt{\frac{(0.5)^2 n_2 + (0.5)^2 8n_2}{2n_2^{\ 2}}} = \frac{1.0607\sigma_1}{\sqrt{n_2}}$$

Section 7-5

7-33. $$P(1.009 \le \overline{X} \le 1.012) = P\left(\frac{1.009 - 1.01}{0.003/\sqrt{9}} \le \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \le \frac{1.012-1.01}{0.003/\sqrt{9}}\right)$$

$$= P(-1 \le Z \le 2) = P(Z \le 2) - P(Z \le -1)$$

$$= 0.9772 - 0.1587 = 0.8385$$

7-35. $$\mu_{\overline{X}} = 75.5\,psi\,,\ \sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}} = \frac{3.5}{\sqrt{6}} = 1.429$$

$$P(\overline{X} \ge 75.75) = P\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \ge \frac{75.75-75.5}{1.429}\right)$$

$$= P(Z \ge 0.175) = 1 - P(Z \le 1.75)$$

$$= 1 - 0.56945 = 0.43055$$

7-39 $\sigma^2 = 25$

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$$

$$n = \left(\frac{\sigma}{\sigma_{\bar{X}}}\right)^2 = \left(\frac{5}{1.5}\right)^2 = 11.11$$

$n \cong 12$

7-41 n = 36

$$\mu_X = \frac{a+b}{2} = \frac{(3+1)}{2} = 2$$

$$\sigma_X = \sqrt{\frac{(b-a+1)^2-1}{12}} = \sqrt{\frac{(3-1+1)^2-1}{12}} = \sqrt{\frac{8}{12}} = \sqrt{\frac{2}{3}}$$

$$\mu_{\bar{X}} = 2, \sigma_{\bar{X}} = \frac{\sqrt{2/3}}{\sqrt{36}} = \frac{\sqrt{2/3}}{6}$$

$$z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$$

Using the central limit theorem:

$$P(2.1 < \bar{X} < 2.5) = P\left(\frac{2.1-2}{\frac{\sqrt{2/3}}{6}} < Z < \frac{2.5-2}{\frac{\sqrt{2/3}}{6}}\right)$$

$$= P(0.7348 < Z < 3.6742)$$

$$= P(Z < 3.6742) - P(Z < 0.7348)$$

$$= 1 - 0.7688 = 0.2312$$

7-43.

$$\begin{array}{lll} n_1 = 16 & n_2 = 9 & \bar{X}_1 - \bar{X}_2 \sim N(\mu_{\bar{X}_1} - \mu_{\bar{X}_2}, \sigma^2_{\bar{X}_1} + \sigma^2_{\bar{X}_2}) \\ \mu_1 = 75 & \mu_2 = 70 & \\ \sigma_1 = 8 & \sigma_2 = 12 & \sim N(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}) \\ & & \sim N(75-70, \frac{8^2}{16} + \frac{12^2}{9}) \\ & & \sim N(5,20) \end{array}$$

a) $P(\bar{X}_1 - \bar{X}_2 > 4)$

$P(Z > \frac{4-5}{\sqrt{20}}) = P(Z > -0.2236) = 1 - P(Z \le -0.2236)$

$= 1 - 0.4115 = 0.5885$

b) $P(3.5 \le \overline{X}_1 - \overline{X}_2 \le 5.5)$

$$P(\tfrac{3.5-5}{\sqrt{20}} \le Z \le \tfrac{5.5-5}{\sqrt{20}}) = P(Z \le 0.1118) - P(Z \le -0.3354)$$

$$= 0.5445 - 0.3686 = 0.1759$$

Supplemental Exercises

7-49. $\overline{X}_1 - \overline{X}_2 \sim N(100-105, \dfrac{1.5^2}{25} + \dfrac{2^2}{25}) \sim N(-5, 0.2233)$

Chapter 8 Selected Problem Solutions

Section 8-2

8-1 a.) The confidence level for $\bar{x} - 2.14\sigma/\sqrt{n} \le \mu \le \bar{x} + 2.14\sigma/\sqrt{n}$ is determined by the by the value of z_0 which is 2.14. From Table II, we find $\Phi(2.14) = P(Z<2.14) = 0.9793$ and the confidence level is 97.93%.

b.) The confidence level for $\bar{x} - 2.49\sigma/\sqrt{n} \le \mu \le \bar{x} + 2.49\sigma/\sqrt{n}$ is determined by the by the value of z_0 which is 2.14. From Table II, we find $\Phi(2.49) = P(Z<2.49) = 0.9936$ and the confidence level is 99.36%.

c.) The confidence level for $\bar{x} - 1.85\sigma/\sqrt{n} \le \mu \le \bar{x} + 1.85\sigma/\sqrt{n}$ is determined by the by the value of z_0 which is 2.14. From Table II, we find $\Phi(1.85) = P(Z<1.85) = 0.9678$ and the confidence level is 96.78%.

8-7 a.) The 99% CI on the mean calcium concentration would be longer.

b). No, that is not the correct interpretation of a confidence interval. The probability that μ is between 0.49 and 0.82 is either 0 or 1.

c). Yes, this is the correct interpretation of a confidence interval. The upper and lower limits of the confidence limits are random variables.

8-13 a) 95% two sided CI on the mean compressive strength

$z_{\alpha/2} = z_{0.025} = 1.96$, and $\bar{x} = 3250$, $\sigma^2 = 1000$, n=12

$$\bar{x} - z_{0.025}\left(\frac{\sigma}{\sqrt{n}}\right) \le \mu \le \bar{x} + z_{0.025}\left(\frac{\sigma}{\sqrt{n}}\right)$$

$$3250 - 1.96\left(\frac{31.62}{\sqrt{12}}\right) \le \mu \le 3250 + 1.96\left(\frac{31.62}{\sqrt{12}}\right)$$

$$3232.11 \le \mu \le 3267.89$$

b.) 99% Two-sided CI on the true mean compressive strength

$z_{\alpha/2} = z_{0.005} = 2.58$

$$\bar{x} - z_{0.005}\left(\frac{\sigma}{\sqrt{n}}\right) \le \mu \le \bar{x} + z_{0.005}\left(\frac{\sigma}{\sqrt{n}}\right)$$

$$3250 - 2.58\left(\frac{31.62}{\sqrt{12}}\right) \le \mu \le 3250 + 2.58\left(\frac{31.62}{\sqrt{12}}\right)$$

$$3226.5 \le \mu \le 3273.5$$

8-15 Set the width to 6 hours with $\sigma = 25$, $z_{0.025} = 1.96$ solve for n.

$1/2 \text{ width} = (1.96)(25)/\sqrt{n} = 3$

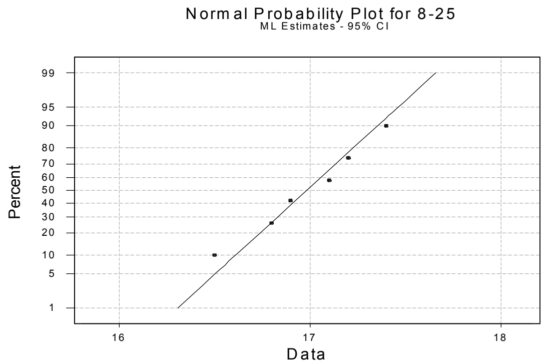
$49 = 3\sqrt{n}$

$n = \left(\frac{49}{3}\right)^2 = 266.78$

Therefore, n=267.

Section 8-3

8-25 a.) The data appear to be normally distributed based on examination of the normal probability plot below. Therefore, there is evidence to support that the level of polyunsaturated fatty acid is normally distributed.



b.) 99% CI on the mean level of polyunsaturated fatty acid.
For $\alpha = 0.01$, $t_{\alpha/2,n-1} = t_{0.005,5} = 4.032$

$$\bar{x} - t_{0.005,5}\left(\frac{s}{\sqrt{n}}\right) \le \mu \le \bar{x} + t_{0.005,5}\left(\frac{s}{\sqrt{n}}\right)$$

$$16.98 - 4.032\left(\frac{0.319}{\sqrt{6}}\right) \le \mu \le 16.98 + 4.032\left(\frac{0.319}{\sqrt{6}}\right)$$

$$16.455 \le \mu \le 17.505$$

8-29 95% lower bound confidence for the mean wall thickness
given $\bar{x} = 4.05$ $s = 0.08$ $n = 25$

$t_{\alpha,n-1} = t_{0.05,24} = 1.711$

$$\bar{x} - t_{0.05,24}\left(\frac{s}{\sqrt{n}}\right) \le \mu$$

$$4.05 - 1.711\left(\frac{0.08}{\sqrt{25}}\right) \le \mu$$

$$4.023 \le \mu$$

It may be assumed that the mean wall thickness will most likely be greater than 4.023 mm.

8-31 $\bar{x} = 1.10$ $s = 0.015$ $n = 25$

95% CI on the mean volume of syrup dispensed

For $\alpha = 0.05$ and $n = 25$, $t_{\alpha/2,n-1} = t_{0.025,24} = 2.064$

$$\bar{x} - t_{0.025,24}\left(\frac{s}{\sqrt{n}}\right) \le \mu \le \bar{x} + t_{0.025,24}\left(\frac{s}{\sqrt{n}}\right)$$

$$1.10 - 2.064\left(\frac{0.015}{\sqrt{25}}\right) \le \mu \le 1.10 + 2.064\left(\frac{0.015}{\sqrt{25}}\right)$$

$$1.093 \le \mu \le 1.106$$

Section 8-4

8-35 99% lower confidence bound for σ^2

For $\alpha = 0.01$ and $n = 15$, $\chi^2_{\alpha,n-1} = \chi^2_{0.01,14} = 29.14$

$$\frac{14(0.008)^2}{29.14} < \sigma^2$$

$$0.00003075 < \sigma^2$$

8-37 95% lower confidence bound for σ^2 given $n = 16$, $s^2 = (3645.94)^2$

For $\alpha = 0.05$ and $n = 16$, $\chi^2_{\alpha,n-1} = \chi^2_{0.05,15} = 25$

$$\frac{15(3645.94)^2}{25} < \sigma^2$$

$$7{,}975{,}727.09 < \sigma^2$$

8-39 95% confidence interval for σ: given $n = 51$, $s = 0.37$

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,

$$0.31 < \sigma < 0.46$$

8-41 90% lower confidence bound on σ (the standard deviation of the sugar content) given $n = 10$, $s^2 = 23.04$

For $\alpha = 0.1$ and $n = 10$, $\chi^2_{\alpha,n-1} = \chi^2_{0.1,9} = 19.02$

$$\frac{9(23.04)}{14.68} \le \sigma^2$$

$$14.13 \le \sigma^2$$

Take the square root of the endpoints of this interval to find the confidence interval for σ:

$$3.8 \le \sigma$$

Section 8-7

8-63 99% tolerance interval on the polyunsaturated fatty acid in this type of margarine that has a confidence level of 95%

$\bar{x}$ = 16.98 s = 0.319 n=6 and k = 5.775

$$\bar{x} - ks, \bar{x} + ks$$

$$16.98 - 5.775(0.319), 16.98 + 5.775(0.319)$$

$$(15.14, 18.82)$$

The 99% tolerance interval is much wider than the 99% confidence interval on the population mean ($16.46 \le \mu \le 17.51$).

8-67 90% lower tolerance bound on bottle wall thickness that has confidence level 90%.

given $\bar{x}$ = 4.05 s = 0.08 n = 25 and k = 1.702

$$\bar{x} - ks$$

$$4.05 - 1.702(0.08)$$

$$3.91$$

$$\text{The 90\% tolarance bound is } (3.91, \infty)$$

The lower tolerance bound is of interest if we want to make sure the wall thickness is at least a certain value so that the bottle will not break.

8-69 95% tolerance interval on the syrup volume that has 90% confidence level

$\bar{x}$ = 1.10 s = 0.015 n = 25 and k=2.474

$$\bar{x} - ks, \bar{x} + ks$$

$$1.10 - 2.474(0.015), 1.10 + 2.474(0.015)$$

$$(1.06, 1.14)$$

Supplemental Exercises

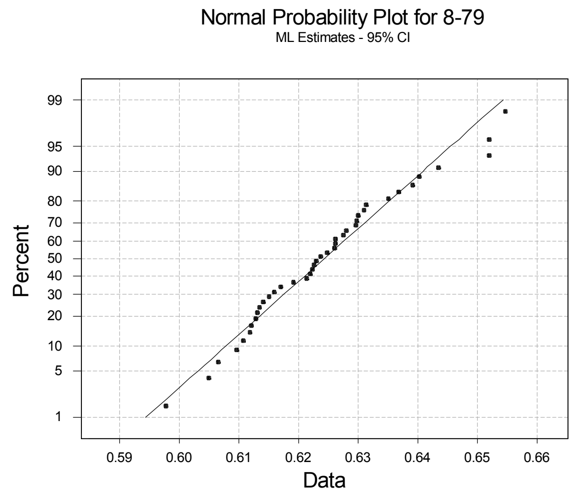
8-75 With $\sigma = 8$, the 95% confidence interval on the mean has length of at most 5; the error is then E = 2.5.

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 confidence and the length of the interval, decreases.

8-79 Normal probability plot for the coefficient of restitution



b.) 99% CI on the true mean coefficient of restitution

$\bar{x} = 0.624$, s = 0.013, n = 40 $t_{a/2,\, n-1} = t_{0.005,\, 39} = 2.7079$

$$\bar{x} - t_{0.005,39}\frac{s}{\sqrt{n}} \le \mu \le \bar{x} + t_{0.005,39}\frac{s}{\sqrt{n}}$$

$$0.624 - 2.7079\frac{0.013}{\sqrt{40}} \le \mu \le 0.624 + 2.7079\frac{0.013}{\sqrt{40}}$$

$$0.618 \le \mu \le 0.630$$

b.) 99% prediction interval on the coefficient of restitution for the next baseball that will be tested.

$$\bar{x} - t_{0.005,39}s\sqrt{1+\frac{1}{n}} \le x_{n+1} \le \bar{x} + t_{0.005,39}s\sqrt{1+\frac{1}{n}}$$

$$0.624 - 2.7079(0.013)\sqrt{1+\frac{1}{40}} \le x_{n+1} \le 0.624 + 2.7079(0.013)\sqrt{1+\frac{1}{40}}$$

$$0.588 \le x_{n+1} \le 0.660$$

c.) 99% tolerance interval on the coefficient of restitution with a 95% level of confidence

$$(\bar{x} - ks,\ \bar{x} + ks)$$

$$(0.624 - 3.213(0.013),\ 0.624 + 3.213(0.013))$$

$$(0.583,\ 0.665)$$

e.)The confidence interval in part (b) describes the confidence interval on the population mean and we may interpret this to mean that 99% of such intervals will cover the population mean. The prediction interval tells us that within that within a 99% probability that the next baseball will have a coefficient of restitution between 0.588 and 0.660. The tolerance interval captures 99% of the values of the normal distribution with a 95% level of confidence.

8-83 a.) 95% Confidence Interval on the population proportion

n=1200 x=8 $\hat{p} = 0.0067$ $z_{\alpha/2}=z_{0.025}=1.96$

$$\hat{p} - z_{a/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \le p \le \hat{p} + z_{a/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$0.0067 - 1.96\sqrt{\frac{0.0067(1-0.0067)}{1200}} \le p \le 0.0067 + 1.96\sqrt{\frac{0.0067(1-0.0067)}{1200}}$$

$$0.0021 \le p \le 0.0088$$

b.) Yes, there is evidence to support the claim that the fraction of defective units produced is one percent or less. This is true because the confidence interval does not include 0.01 and the upper limit of the control interval is lower than 0.01.

Chapter 9 Selected Problems Solutions

Section 9-1

9-1 a) $H_0: \mu = 25, H_1: \mu \neq 25$ Yes, because the hypothesis is stated in terms of the parameter of interest, inequality is in the alternative hypothesis, and the value in the null and alternative hypotheses matches.

b) $H_0: \sigma > 10, H_1: \sigma = 10$ No, because the inequality is in the null hypothesis.

c) $H_0: \bar{x} = 50, H_1: \bar{x} \neq 50$ No, because the hypothesis is stated in terms of the statistic rather than the parameter.

d) $H_0: p = 0.1, H_1: p = 0.3$ No, the values in the hull and alternative hypotheses do not match and both of the hypotheses are equality statements.

e) $H_0: s = 30, H_1: s > 30$ No, because the hypothesis is stated in terms of the statistic rather than the parameter.

9-3 a) $\alpha = P(\bar{X} \leq 11.5 \mid \mu = 12) = P\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq \frac{11.5 - 12}{0.5/\sqrt{16}}\right) = P(Z \leq -4) = 1 - P(Z \leq 4)$

$= 1 - 1 = 0.$

The probability of rejecting the null, when the null is true, is approximately 0 with a sample size of 16.

b) $\beta = P(\bar{X} > 11.5 \mid \mu = 11.25) = P\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} > \frac{11.5 - 11.25}{0.5/\sqrt{16}}\right) = P(Z > 2) = 1 - P(Z \leq 2)$

$= 1 - 0.97725 = 0.02275.$

The probability of accepting the null hypothesis when it is false is 0.02275.

9-9 a) $z = \frac{190 - 175}{20/\sqrt{10}} = 2.37$, Note that z is large, therefore **reject** the null hypothesis and conclude that the mean foam height is greater than 175 mm.

b) $P(\bar{X} > 190 \text{ when } \mu = 175)$

$= P\left(\frac{\bar{X} - 175}{20/\sqrt{10}} > \frac{190 - 175}{20/\sqrt{10}}\right)$

$= P(Z > 2.37) = 1 - P(Z \leq 2.37)$

$= 1 - 0.99111$

$= 0.00889.$

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 $\bar{x} = 190$ mm would be an unusual result.

9-17. The problem statement implies H_0: $p = 0.6$, H_1: $p > 0.6$ and defines an acceptance region as $\hat{p} \leq \frac{315}{500} = 0.63$ and rejection region as $\hat{p} > 0.63$

a) $\alpha = P(\hat{P} \geq 0.63 \mid p = 0.6) = P\left(Z \geq \frac{0.63 - 0.6}{\sqrt{\frac{0.6(0.4)}{500}}}\right).$

$= P(Z \geq 1.37) = 1 - P(Z < 1.37) = 0.08535$

b) $\beta = P(\hat{P} \leq 0.63 \text{ when } p = 0.75) = P(Z \leq -6.196) \cong 0.$

Section 9-2

9-21. a) 1) The parameter of interest is the true mean yield, μ.

2) $H_0 : \mu = 90$

3) $H_1 : \mu \neq 90$

4) $\alpha = 0.05$

5) $z_0 = \dfrac{\overline{x} - \mu}{\sigma / \sqrt{n}}$

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 $\alpha = 0.05$, $z_{\alpha/2} = z_{0.025} = 1.96$

$$\overline{x} - z_{0.025}\left(\frac{\sigma}{\sqrt{n}}\right) \le \mu \le \overline{x} + z_{0.025}\left(\frac{\sigma}{\sqrt{n}}\right)$$

$$90.48 - 1.96\left(\frac{3}{\sqrt{5}}\right) \le \mu \le 90.48 + 1.96\left(\frac{3}{\sqrt{5}}\right)$$

$87.85 \le \mu \le 93.11$

With 95% confidence, we believe the true mean yield of the chemical process is between 87.85% and 93.11%.

9-25. a) 1) The parameter of interest is the true mean tensile strength, μ.

2) $H_0 : \mu = 3500$

3) $H_1 : \mu \neq 3500$

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) $\overline{x} = 3250$, $\sigma = 60$

$$z_0 = \frac{3250 - 3500}{60/\sqrt{12}} = -14.43$$

8) Since $-14.43 < -2.58$, 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 $= 2[1 - \Phi(14.43)] = 2[1 - 1] = 0$

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 tensile strength is between 3232.11 psi and 3267.89 psi. We can test the hypotheses that the true mean strength is not equal to 3500 by noting that the value is not within the confidence interval.

9-27 a) 1) The parameter of interest is the true mean speed, μ.

2) $H_0 : \mu = 100$

3) $H_1 : \mu < 100$

4) $\alpha = 0.05$

5) $z_0 = \dfrac{\bar{x} - \mu}{\sigma / \sqrt{n}}$

6) Reject H_0 if $z_0 < -z_\alpha$ where $-z_{0.05} = -1.65$

7) $\bar{x} = 102.2$, $\sigma = 4$

$$z_0 = \frac{102.2 - 100}{4/\sqrt{8}} = 1.55$$

8) Since 1.55> −1.65, do not reject the null hypothesis and conclude the there is insufficient evidence to conclude that the true speed strength is less than 100 at $\alpha = 0.05$.

b) $\beta = \Phi\left(-z_{0.05} - \dfrac{(95-100)\sqrt{8}}{4}\right) = \Phi(-1.65 - -3.54) = \Phi(1.89) = 1$

Power = 1-β = 1-0.97062 = 0.02938

c) $n = \dfrac{(z_\alpha + z_\beta)^2 \sigma^2}{\delta^2} = \dfrac{(z_{0.05} + z_{0.15})^2 \sigma^2}{(95-100)^2} = \dfrac{(1.65 + 1.03)^2 (4)^2}{(5)^2} = 0.927,$

$n \cong 1$

d) $\bar{x} - z_{0.05}\left(\dfrac{\sigma}{\sqrt{n}}\right) \le \mu$

$$102.2 - 1.65\left(\frac{4}{\sqrt{8}}\right) \le \mu$$

$$99.866 \le \mu$$

Since the lower limit of the CI is just slightly below 100, we are confident that the mean speed is not less than 100 m/s.

9-29 a) 1) The parameter of interest is the true average battery life, μ.

2) $H_0 : \mu = 4$

3) $H_1 : \mu > 4$

4) $\alpha = 0.05$

5) $z_0 = \dfrac{\bar{x} - \mu}{\sigma / \sqrt{n}}$

6) Reject H_0 if $z_0 > z_\alpha$ where $z_{0.05} = 1.65$

7) $\bar{x} = 4.05$, $\sigma = 0.2$

$$z_0 = \frac{4.05-4}{0.2/\sqrt{50}} = 1.77$$

8) Since 1.77>1.65, reject the null hypothesis and conclude that the there is sufficient evidence to conclude that the true average battery life exceeds 4 hours at $\alpha = 0.05$.

b) $\beta = \Phi\left(z_{0.05} - \frac{(4.5-4)\sqrt{50}}{0.2}\right) = \Phi(1.65 - 17.68) = \Phi(-16.03) = 0$

Power = 1-β = 1-0 = 1

c) $n = \frac{(z_\alpha + z_\beta)^2\sigma^2}{\delta^2} = \frac{(z_{0.05} + z_{0.1})^2\sigma^2}{(4.5-4)^2} = \frac{(1.65+1.29)^2(0.2)^2}{(0.5)^2} = 34.7,$

$n \cong 35$

d) $\bar{x} - z_{0.05}\left(\frac{\sigma}{\sqrt{n}}\right) \le \mu$

$$4.05 - 1.65\left(\frac{0.2}{\sqrt{50}}\right) \le \mu$$

$$4.003 \le \mu$$

Since the lower limit of the CI is just slightly above 4, we conclude that average life is greater than 4 hours at α=0.05.

Section 9-3

9-31 a. 1) The parameter of interest is the true mean female body temperature, μ.

2) $H_0 : \mu = 98.6$

3) $H_1 : \mu \ne 98.6$

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 $t_{\alpha/2,n-1} = 2.064$

7) $\bar{x} = 98.264$, s = 0.4821 n=25

$$t_0 = \frac{98.264 - 98.6}{0.4821/\sqrt{25}} = -3.48$$

8) Since 3.48 > 2.064, reject the null hypothesis and conclude that the there is sufficient evidence to conclude that the true mean female body temperature is not equal to 98.6 °F at $\alpha = 0.05$.

P-value = 2* 0.001 = 0.002

b) $d = \frac{\delta}{\sigma} = \frac{|\mu - \mu_0|}{\sigma} = \frac{|98 - 98.6|}{0.4821} = 1.24$

Using the OC curve, Chart VI e) for $\alpha = 0.05$, d = 1.24, and n = 25, we get $\beta \cong 0$ and power of $1-0 \cong 1$.

c) $d = \frac{\delta}{\sigma} = \frac{|\mu - \mu_0|}{\sigma} = \frac{|98.2 - 98.6|}{0.4821} = 0.83$

Using the OC curve, Chart VI g) for $\alpha = 0.05$, d = 0.83, and $\beta \cong 0.1$ (Power=0.9),

$n^* = 20$. Therefore, $n = \frac{n^*+1}{2} = \frac{20+1}{2} = 10.5$ and n=11.

d) 95% two sided confidence interval

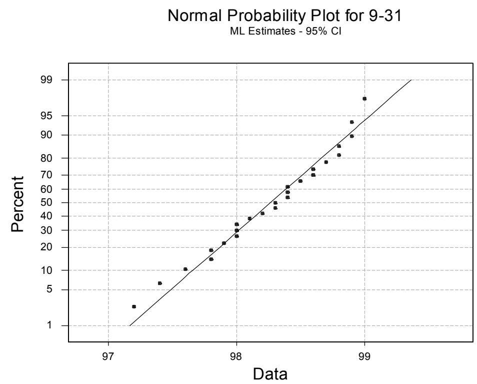
$$\bar{x} - t_{0.025,24}\left(\frac{s}{\sqrt{n}}\right) \le \mu \le \bar{x} + t_{0.025,24}\left(\frac{s}{\sqrt{n}}\right)$$

$$98.264 - 2.064\left(\frac{0.4821}{\sqrt{25}}\right) \le \mu \le 98.264 + 2.064\left(\frac{0.4821}{\sqrt{25}}\right)$$

$$98.065 \le \mu \le 98.463$$

We can conclude that the mean female body temperature is not equal to 98.6 since the value is not included inside the confidence interval.

e)



Data appear to be normally distributed.

9-37. a.) 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 mean coefficient of restitution, μ.

2) $H_0 : \mu = 0.635$

3) $H_1 : \mu > 0.635$

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 $t_{0.05,39} = 1.685$

7) $\bar{x} = 0.624$ $s = 0.013$ $n = 40$

$$t_0 = \frac{0.624 - 0.635}{0.013/\sqrt{40}} = -5.35$$

8) Since –5.25 < 1.685, do not reject the null hypothesis and conclude that there is not sufficient evidence to indicate that the true mean coefficient of restitution is greater than 0.635 at $\alpha = 0.05$.

b.)The P-value > 0.4, based on Table IV. Minitab gives *P-value* = 1.

c) $d = \dfrac{\delta}{\sigma} = \dfrac{|\mu - \mu_0|}{\sigma} = \dfrac{|0.64 - 0.635|}{0.013} = 0.38$

Using the OC curve, Chart VI g) for $\alpha = 0.05$, d = 0.38, and n = 40, we get $\beta \cong 0.25$ and power of 1–0.25 = 0.75.

d) $d = \dfrac{\delta}{\sigma} = \dfrac{|\mu - \mu_0|}{\sigma} = \dfrac{|0.638 - 0.635|}{0.013} = 0.23$

Using the OC curve, Chart VI g) for $\alpha = 0.05$, d = 0.23, and $\beta \cong 0.25$ (Power=0.75),

$n^* = 75$. Therefore, $n = \frac{n^* + 1}{2} = \frac{75+1}{2} = 38$ and n=38.

9-41 a) 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 mean concentration of suspended solids, μ.
2) $H_0 : \mu = 55$
3) $H_1 : \mu \neq 55$
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 $t_{0.025,59} = 2.000$
7) $\bar{x} = 59.87$ $s = 12.50$ $n = 60$

$$t_0 = \frac{59.87 - 55}{12.50/\sqrt{60}} = 3.018$$

8) Since 3.018 > 2.000, reject the null hypothesis and conclude that there is sufficient evidence to indicate that the true mean concentration of suspended solids is not equal to 55 at $\alpha = 0.05$.

b) From table IV the t0 value is found between the values of 0.001 and 0.0025 with 59 degrees of freedom, so 2*0.001<*P-value* = 2* 0.0025 Therefore, 0.002< P-value<0.005.
Minitab gives a p-value of 0.0038

c) $d = \frac{|50-55|}{12.50} = 0.4$, n=60 so, from the OC Chart VI e) for $\alpha = 0.05$, d= 0.4 and n=60 we find that $\beta \cong 0.2$. Therefore, the power = 1-0.2 = 0.8.

d) From the same OC chart, and for the specified power, we would need approximately 38 observations.

$d = \frac{|50-55|}{12.50} = 0.4$ Using the OC Chart VI e) for $\alpha = 0.05$, d = 0.4, and $\beta \cong 0.10$ (Power=0.90),

$n^* = 75$. Therefore, $n = \frac{n^* + 1}{2} = \frac{75+1}{2} = 38$ and n=38.

Section 9-4

9-43 a) In order to use the 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) $\chi_0^2 = \frac{(n-1)s^2}{\sigma^2}$
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 < P-value < 0.9

c) $\lambda = \dfrac{\sigma}{\sigma_0} = \dfrac{0.0125}{0.01} = 1.25$ power = 0.8, β=0.2

using chart VIk, the required sample size is 50

9-47. 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,

$$0.31 < \sigma < 0.46$$

Since 0.25 falls below the lower confidence bound we would conclude that the population standard deviation is not equal to 0.25.

9-49 Using the chart in the Appendix, with $\lambda = \sqrt{\dfrac{40}{18}} = 1.49$ and $\beta = 0.10$, we find n = 30.

Section 9-5

9-51 p= 0.15, p_0=0.10, n=85, and $z_{\alpha/2}$=1.96

$$\beta = \Phi\left(\frac{p_0 - p + z_{\alpha/2}\sqrt{p_0(1-p_0)/n}}{\sqrt{p(1-p)/n}}\right) - \Phi\left(\frac{p_0 - p - z_{\alpha/2}\sqrt{p_0(1-p_0)/n}}{\sqrt{p(1-p)/n}}\right)$$

$$= \Phi\left(\frac{0.10 - 0.15 + 1.96\sqrt{0.10(1-0.10)/85}}{\sqrt{0.15(1-0.15)/85}}\right) - \Phi\left(\frac{0.10 - 0.15 - 1.96\sqrt{0.10(1-0.10)/85}}{\sqrt{0.15(1-0.15)/85}}\right)$$

$$= \Phi(0.36) - \Phi(-2.94) = 0.6406 - 0.0016 = 0.639$$

$$n=\left(\frac{z_{\alpha/2}\sqrt{p_0(1-p_0)}-z_\beta\sqrt{p(1-p)}}{p-p_0}\right)^2$$
$$=\left(\frac{1.96\sqrt{0.10(1-0.10)}-1.28\sqrt{0.15(1-0.15)}}{0.15-0.10}\right)^2$$
$$=(10.85)^2=117.63\cong 118$$

9-53. 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) $z_0=\frac{x-np_0}{\sqrt{np_0(1-p_0)}}$ 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 $z_0 < -z_\alpha$ where $-z_\alpha = -z_{0.05} = -1.65$

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 $-0.53 > -1.65$, do not 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 $= 1 - \Phi(0.53) = 0.29806$

9-57. The problem statement implies that 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) The probability of a type 1 error is

$$\alpha = P(\hat{p}\ge 0.63 \mid p=0.6)=P\left(Z\ge \frac{0.63-0.6}{\sqrt{\frac{0.6(0.4)}{500}}}\right)=P(Z\ge 1.37)=1-P(Z<1.37)=0.08535$$

b) $\beta = P(\hat{P} \le 0.63 \mid p = 0.75) = P(Z \le -6.196) = 0.$

Section 9-7

9-59.

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) $$\chi_0^2 = \frac{(24-30.12)^2}{30.12} + \frac{(30-36.14)^2}{36.14} + \frac{(31-21.69)^2}{21.69} + \frac{(15-11.67)^2}{11.67} = 7.23$$

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) The P-value is between 0.05 and 0.1 using Table III. P-value = 0.0649 (found using Minitab)

9-63 The value of p must be estimated. Let the estimate be denoted by $\hat{p}_{sample}$

$$\text{sample mean} = \frac{0(39) + 1(23) + 2(12) + 3(1)}{75} = 0.6667$$

$$\hat{p}_{sample} = \frac{sample \quad mean}{n} = \frac{0.6667}{24} = 0.02778$$

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 under-filled 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 under-filled cartons is binomial at $\alpha = 0.05$.

b) The P-value is between 0.5 and 0.1 using Table III P-value = 0.3048 (found using Minitab)

Section 9-8

9-65. 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 $\chi^2_{.05,6} = 12.592$
7. The calculated test statistic is $\chi^2_0 = 11.65$
8. $\chi^2_0 \not> \chi^2_{0.05,6}$, 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)

9-69. 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 that the 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

Supplemental

9-75. $\sigma = 8$, $\delta = 204 - 200 = -4$, $\frac{\alpha}{2} = 0.025$, $z_{0.025} = 1.96$.

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.

9-77. a) Rejecting a null hypothesis provides a *stronger conclusion* than failing to reject a null hypothesis. Therefore, place what we are trying to demonstrate in the alternative hypothesis.

Assume that 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:

$$t_0 = \frac{\overline{x} - \mu_0}{s / \sqrt{n}}$$

6) Since no critical value is given, we will calculate the P-value

7) $\overline{x} = 153.7$, s= 11.3, n=20

$$t_0 = \frac{153.7 - 150}{11.3\sqrt{20}} = 1.46$$

P-value = $P(t \geq 1.46) = 0.05 < p - \text{value} < 0.10$

8) There is some modest 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.

9-79 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 = $P(\chi^2 < 5.546)$; $0.1 < P - \text{value} < 0.5$

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 = $P(\chi^2 < 30.81)$; $0.01 < P - \text{value} < 0.025$

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.

9-85 We can divide the real line under a standard normal distribution into eight intervals with equal probability. These intervals are [0,.32), [0.32, 0.675), [0.675, 1.15), [1.15, ∞) and their negative counterparts. The probability for each interval is p = 1/8 = .125 so the expected cell frequencies are E = np = (100) (0.125) = 12.5. The table of ranges and their corresponding frequencies is completed as follows.

Interval	Obs. Frequency.	Exp. Frequency.
$x \le 5332.5$	1	12.5
$5332.5 < x \le 5357.5$	4	12.5
$5357.5 < x \le 5382.5$	7	12.5
$5382.5 < x \le 5407.5$	24	12.5
$5407.5 < x \le 5432.5$	30	12.5
$5432.5 < x \le 5457.5$	20	12.5
$5457.5 < x \le 5482.5$	15	12.5
$x \ge 5482.5$	5	12.5

The test statistic is:

$$\chi_0^2 = \frac{(1-12.5)^2}{12.5} + \frac{(4-12.5)^2}{12.5} + \Lambda + \frac{(15-12.5)^2}{12.5} + \frac{(5-12.5)^2}{12.5} = 63.36$$

and we would reject if this value exceeds $\chi^2_{0.05,5} = 11.07$. Since $\chi_o^2 > \chi_{0.05,5}^2$, reject the hypothesis that the data are normally distributed

9-87 a) 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 mean overall distance for this brand of golf ball, μ.
2) $H_0 : \mu = 270$
3) $H_1 : \mu < 270$
4) $\alpha = 0.05$
5) Since n>>30 we can use the normal distribution

$$z_0 = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

6) Reject H_0 if $z_0 < -z_\alpha$ where $z_{0.05} = 1.65$
7) $\bar{x} = 1.25$ s = 0.25 n = 20

$$z_0 = \frac{260.30 - 270.0}{13.41/\sqrt{100}} = -7.23$$

8) Since $-7.23 < -1.65$, reject the null hypothesis and conclude there is sufficient evidence to indicate that the true mean distance is less than 270 yds at $\alpha = 0.05$.

b) The P-value $\cong 0$.

c) We can divide the real line under a standard normal distribution into eight intervals with equal probability. These intervals are [0,.32), [0.32, 0.675), [0.675, 1.15), [1.15, ∞) and their negative counterparts. The probability for each interval is p = 1/8 = .125 so the expected cell frequencies are E = np = (100) (0.125) = 12.5. The table of ranges and their corresponding frequencies is completed as follows.

Interval	Obs. Frequency.	Exp. Frequency.
$x \le 244.88$	16	12.5
$244.88 < x \le 251.25$	6	12.5
$251.25 < x \le 256.01$	17	12.5
$256.01 < x \le 260.30$	9	12.5
$260.30 < x \le 264.59$	13	12.5
$264.59 < x \le 269.35$	8	12.5
$269.35 < x \le 275.72$	19	12.5
$x \ge 275.72$	12	12.5

The test statistic is:

$$\chi^2{}_o = \frac{(16-12.5)^2}{12.5} + \frac{(6-12.5)^2}{12.5} + \Lambda + \frac{(19-12.5)^2}{12.5} + \frac{(12-12.5)^2}{12.5} = 12$$

and we would reject if this value exceeds $\chi^2{}_{0.05,5} = 11.07$. Since it does, we can reject the hypothesis that the data are normally distributed.

Chapter 10 Selected Problem Solutions

Section 10-2

10-1. a) 1) The parameter of interest is the difference in fill volume, $\mu_1 - \mu_2$ (note that Δ_0=0)

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) $\bar{x}_1 = 16.015$ $\quad \bar{x}_2 = 16.005$

$\sigma_1 = 0.02$ $\quad \sigma_2 = 0.025$

$n_1 = 10$ $\quad n_2 = 10$

$$z_0 = \frac{(16.015 - 16.005)}{\sqrt{\frac{(0.02)^2}{10} + \frac{(0.025)^2}{10}}} = 0.99$$

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

$$\beta = \Phi\left(z_{\alpha/2} - \frac{\Delta - \Delta_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \right) - \Phi\left(-z_{\alpha/2} - \frac{\Delta - \Delta_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \right)$$

$$= \Phi\left(1.96 - \frac{0.04}{\sqrt{\frac{(0.02)^2}{10} + \frac{(0.025)^2}{10}}} \right) - \Phi\left(-1.96 - \frac{0.04}{\sqrt{\frac{(0.02)^2}{10} + \frac{(0.025)^2}{10}}} \right)$$

$$= \Phi(1.96 - 3.95) - \Phi(-1.96 - 3.95) = \Phi(-1.99) - \Phi(-5.91)$$

$= 0.0233 - 0$

$= 0.0233$

Power = 1 $-$0.0233 = 0.9967

d) $$(\bar{x}_1 - \bar{x}_2) - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \le \mu_1 - \mu_2 \le (\bar{x}_1 - \bar{x}_2) + z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$$(16.015 - 16.005) - 1.96\sqrt{\frac{(0.02)^2}{10} + \frac{(0.025)^2}{10}} \le \mu_1 - \mu_2 \le (16.015 - 16.005) + 1.96\sqrt{\frac{(0.02)^2}{10} + \frac{(0.025)^2}{10}}$$

$$-0.0098 \le \mu_1 - \mu_2 \le 0.0298$$

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 $\Delta = 0.04$

$$n \cong \frac{\left(z_{\alpha/2}+z_\beta\right)^2\left(\sigma_1^2+\sigma_2^2\right)}{\delta^2}=\frac{(1.96+1.645)^2\left((0.02)^2+(0.025)^2\right)}{(0.04)^2}=8.33, \quad \text{n = 9,}$$

use $n_1 = n_2 = 9$

10-5. $\bar{x}_1 = 30.87 \quad \bar{x}_2 = 30.68$

$\sigma_1 = 0.10 \quad \sigma_2 = 0.15$

$n_1 = 12 \quad n_2 = 10$

a) 90% two-sided confidence interval:

$$\left(\bar{x}_1-\bar{x}_2\right)-z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}} \le \mu_1-\mu_2 \le \left(\bar{x}_1-\bar{x}_2\right)+z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$$

$$(30.87-30.68)-1.645\sqrt{\frac{(0.10)^2}{12}+\frac{(0.15)^2}{10}} \le \mu_1-\mu_2 \le (30.87-30.68)+1.645\sqrt{\frac{(0.10)^2}{12}+\frac{(0.15)^2}{10}}$$

$$0.0987 \le \mu_1-\mu_2 \le 0.2813$$

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:

$$\left(\bar{x}_1-\bar{x}_2\right)-z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}} \le \mu_1-\mu_2 \le \left(\bar{x}_1-\bar{x}_2\right)+z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$$

$$(30.87-30.68)-1.96\sqrt{\frac{(0.10)^2}{12}+\frac{(0.15)^2}{10}} \le \mu_1-\mu_2 \le (30.87-30.68)+1.96\sqrt{\frac{(0.10)^2}{12}+\frac{(0.15)^2}{10}}$$

$$0.0812 \le \mu_1-\mu_2 \le 0.299$$

We are 95% confident that the mean fill volume for machine 1 exceeds that of machine 2 by between 0.0812 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:

$$\mu_1-\mu_2 \le \left(\bar{x}_1-\bar{x}_2\right)+z_\alpha\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$$

$$\mu_1-\mu_2 \le (30.87-30.68)+1.645\sqrt{\frac{(0.10)^2}{12}+\frac{(0.15)^2}{10}}$$

$$\mu_1-\mu_2 \le 0.2813$$

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.

10-7. $\bar{x}_1 = 89.6 \quad \bar{x}_2 = 92.5$

$\sigma_1^2 = 1.5 \quad \sigma_2^2 = 1.2$

$n_1 = 15 \quad n_2 = 20$

a) 95% confidence interval:

$$\left(\bar{x}_1-\bar{x}_2\right)-z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}} \le \mu_1-\mu_2 \le \left(\bar{x}_1-\bar{x}_2\right)+z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$$

$$(89.6 - 92.5) - 1.96\sqrt{\frac{1.5}{15} + \frac{1.2}{20}} \le \mu_1 - \mu_2 \le (89.6 - 92.5) + 1.96\sqrt{\frac{1.5}{15} + \frac{1.2}{20}}$$
$$-3.684 \le \mu_1 - \mu_2 \le -2.116$$

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

$$z_0 = \frac{(\bar{x}_1 - \bar{x}_2) - \Delta_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

6) Reject H_0 if $z_0 < -z_\alpha = -1.645$

7) $\bar{x}_1 = 89.6$ $\quad \bar{x}_2 = 92.5$

$\sigma_1^2 = 1.5$ $\quad \sigma_2^2 = 1.2$

$n_1 = 15$ $\quad n_2 = 20$

$$z_0 = \frac{(89.6 - 92.5) - 0}{\sqrt{\frac{(1.5)^2}{15} + \frac{(1.2)^2}{20}}} = -7.254$$

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 = $P(z \le -7.25) = 1 - P(z \le 7.25) = 1 - 1 \cong 0$

10-9. 95% level of confidence, $E = 1$, and $z_{0.025} = 1.96$

$$n \cong \left(\frac{z_{0.025}}{E}\right)^2 (\sigma_1^2 + \sigma_2^2) = \left(\frac{1.96}{1}\right)^2 (1.5 + 1.2) = 10.37,$$ $n = 11$, use $n_1 = n_2 = 11$

10-11.

Catalyst 1	Catalyst 2
$\bar{x}_1 = 65.22$	$\bar{x}_2 = 68.42$
$\sigma_1 = 3$	$\sigma_2 = 3$
$n_1 = 10$	$n_2 = 10$

a) 95% confidence interval on $\mu_1 - \mu_2$, the difference in mean active concentration

$$(\bar{x}_1 - \bar{x}_2) - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \le \mu_1 - \mu_2 \le (\bar{x}_1 - \bar{x}_2) + z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$$(65.22 - 68.42) - 1.96\sqrt{\frac{(3)^2}{10} + \frac{(3)^2}{10}} \le \mu_1 - \mu_2 \le (65.22 - 68.42) + 1.96\sqrt{\frac{(3)^2}{10} + \frac{(3)^2}{10}}$$
$$-5.83 \le \mu_1 - \mu_2 \le -0.57$$

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 that the mean active concentration depends on the choice of catalyst.

10-13. 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

$$z_0 = \frac{(\bar{x}_1 - \bar{x}_2) - \Delta_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

6) Reject H_0 if $z_0 < -z_{\alpha/2} = -1.96$ or $z_0 > z_{\alpha/2} = 1.96$

7) $\bar{x}_1 = 750.2$ $\bar{x}_2 = 756.88$ $\delta = 0$

$\sigma_1 = 20$ $\sigma_2 = 20$

$n_1 = 15$ $n_2 = 8$

$$z_0 = \frac{(750.2 - 756.88) - 0}{\sqrt{\frac{(20)^2}{15} + \frac{(20)^2}{8}}} = -2.385$$

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 of the parameter under test that is specified in the null hypothesis falls outside the confidence interval, this is equivalent to rejecting the null hypothesis.

Section 10-3

10-17 a) 1) The parameter of interest is the difference in mean rod diameter, $\mu_1 - \mu_2$, with $\Delta_0 = 0$

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 $-t_{0.025,30} = -2.042$ or $t_0 > t_{\alpha/2, n_1+n_2-2}$ where $t_{0.025,30} = 2.042$

7)) $\bar{x}_1 = 8.73$ $\bar{x}_2 = 8.68$

$s_1^2 = 0.35$ $s_2^2 = 0.40$

$n_1 = 15$ $n_2 = 17$

$$s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}} = \sqrt{\frac{14(0.35) + 16(0.40)}{30}} = 0.614$$

$$t_0 = \frac{(8.73 - 8.68)}{0.614\sqrt{\frac{1}{15} + \frac{1}{17}}} = 0.230$$

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: $t_{0.025,30} = 2.042$

$$(\bar{x}_1 - \bar{x}_2) - t_{\alpha/2,n_1+n_2-2}(s_p)\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} \le \mu_1 - \mu_2 \le (\bar{x}_1 - \bar{x}_2) + t_{\alpha/2,n_1+n_2-2}(s_p)\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$

$$(8.73 - 8.68) - 2.042(0.614)\sqrt{\frac{1}{15}+\frac{1}{17}} \le \mu_1 - \mu_2 \le (8.73 - 8.68) + 2.042(0.643)\sqrt{\frac{1}{15}+\frac{1}{17}}$$

$$-0.394 \le \mu_1 - \mu_2 \le 0.494$$

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.

10-21. a) 1) The parameter of interest is the difference in mean etch rate, $\mu_1 - \mu_2$, with $\Delta_0 = 0$

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) $\bar{x}_1 = 9.97$ $\quad \bar{x}_2 = 10.4$ $\quad s_1 = 0.422$ $\quad s_2 = 0.231$ $\quad n_1 = 10$ $\quad n_2 = 10$

$$s_p = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}} = \sqrt{\frac{9(0.422)^2 + 9(0.231)^2}{18}} = 0.340$$

$$t_0 = \frac{(9.97 - 10.4)}{0.340\sqrt{\frac{1}{10}+\frac{1}{10}}} = -2.83$$

8) Since $-2.83 < -2.101$ reject the null hypothesis and conclude the two machines mean etch rates do significantly differ at $\alpha = 0.05$.

b) P-value = $2P(t < -2.83)$ $\quad 2(0.005) <$ P-value $< 2(0.010) = 0.010 <$ P-value < 0.020

c) 95% confidence interval: $t_{0.025,18} = 2.101$

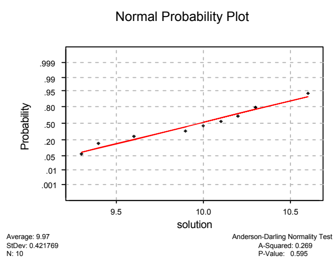
$$(\bar{x}_1 - \bar{x}_2) - t_{\alpha/2,n_1+n_2-2}(s_p)\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} \le \mu_1 - \mu_2 \le (\bar{x}_1 - \bar{x}_2) + t_{\alpha/2,n_1+n_2-2}(s_p)\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$

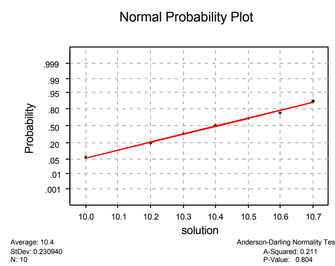
$$(9.97 - 10.4) - 2.101(.340)\sqrt{\frac{1}{10}+\frac{1}{10}} \le \mu_1 - \mu_2 \le (9.97 - 10.4) + 2.101(.340)\sqrt{\frac{1}{10}+\frac{1}{10}}$$

$$-0.749 \le \mu_1 - \mu_2 \le -0.111$$

We are 95% confident that the mean etch rate for solution 2 exceeds the mean etch rate 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 from both samples fall approximately along straight lines. The equality of variances does not appear to be severely violated either since the slopes are approximately the same for both samples.





10-27 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

$$v = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)}{n_2 - 1}} = 26.98$$

$$v \cong 26$$

(truncated)

7) $\bar{x}_1 = 20$ $\bar{x}_2 = 15$ $\Delta_0 = 0$

$s_1 = 2$ $s_2 = 8$

$n_1 = 25$ $n_2 = 25$

$$t_0 = \frac{(20 - 15) - 0}{\sqrt{\frac{(2)^2}{25} + \frac{(8)^2}{25}}} = 3.03$$

8) Since 3.03 > 2.056 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 = $2P(t > 3.03)$, $2(0.0025) < \text{P-value} < 2(0.005)$

$0.005 < \text{P-value} < 0.010$

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

$$t_0 = \frac{(\overline{x}_1 - \overline{x}_2) - \Delta_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

6) Reject the null hypothesis if $t_0 > t_{0.05,27}$ where $t_{0.05,26} = 1.706$ since

7) $\overline{x}_1 = 20$ $\overline{x}_2 = 15$

$s_1 = 2$ $s_2 = 8$ $\Delta_0 = 0$

$n_1 = 25$ $n_2 = 25$

$$t_0 = \frac{(20-15)-0}{\sqrt{\frac{(2)^2}{25} + \frac{(8)^2}{25}}} = 3.03$$

8) Since 3.03 > 1.706 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.

10-29. If $\alpha = 0.01$, construct a 99% lower one-sided confidence interval on the difference to answer question 10-28.

$t_{0.005,19} = 2.878$

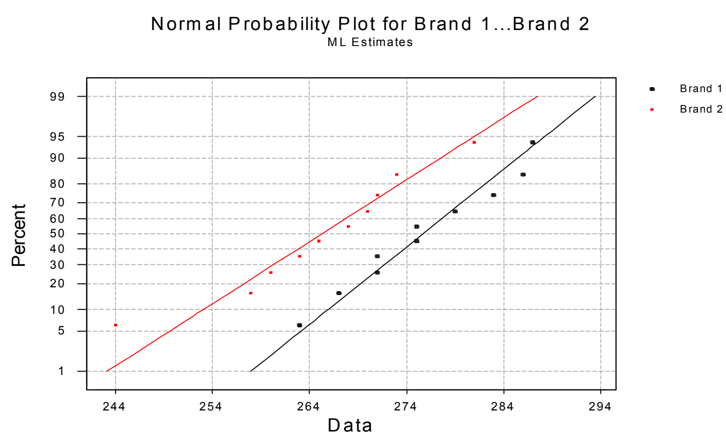
$$(\bar{x}_1 - \bar{x}_2) - t_{\alpha/2,v}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \le \mu_1 - \mu_2 \le (\bar{x}_1 - \bar{x}_2) + t_{\alpha/2,v}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$$(103.5 - 99.7) - 2.878\sqrt{\frac{(10.2)^2}{12} + \frac{(20.1)^2}{13}} \le \mu_1 - \mu_2 \le (103.5 - 99.7) - 2.878\sqrt{\frac{(10.2)^2}{12} + \frac{(20.1)^2}{13}}$$

$$-14.34 \le \mu_1 - \mu_2 \le 21.94.$$

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.

10-31 a.)



b . 1) The parameter of interest is the difference in mean overall distance, $\mu_1 - \mu_2$, with $\Delta_0 = 0$

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) $\bar{x}_1 = 275.7$ $\quad \bar{x}_2 = 265.3$

$$s_p = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}}$$

$s_1 = 8.03$ $\quad s_2 = 10.04$

$$= \sqrt{\frac{9(8.03)^2 + 9(10.04)^2}{20}} = 9.09$$

$n_1 = 10$ $\quad n_2 = 10$

$$t_0 = \frac{(275.7 - 265.3)}{9.09\sqrt{\frac{1}{10} + \frac{1}{10}}} = 2.558$$

8) Since 2.558>2.101 reject the null hypothesis and conclude that the data do not support the claim that both brands have the same mean overall distance at $\alpha = 0.05$. It appears that brand 1 has the higher mean differnce.

c.)P-value = $2P(t < 2.558)$ P-value $\approx 2(0.01)=0.02$

d.) $d = \frac{5}{2(9.09)} 0.275$ $\quad \beta=0.95$ $\quad$ Power =1-0.95=0.05

e.) $1-\beta=0..75$ $\quad \beta=0..27$ $\quad d = \frac{3}{2(9.09)} = 0.165$ $\quad n^*=100$ $\quad n = \frac{100+1}{2} = 50.5$

Therefore, $n=51$

f.) $$(\bar{x}_1 - \bar{x}_2) - t_{\alpha,v} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \le \mu_1 - \mu_2 \le (\bar{x}_1 - \bar{x}_2) + t_{\alpha,v} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$(275.7 - 265.3) - 2.101(9.09)\sqrt{\frac{1}{10} + \frac{1}{10}} \le \mu_1 - \mu_2 \le (275.7 - 265.3) + 2.101(9.09)\sqrt{\frac{1}{10} + \frac{1}{10}}$$

$$1.86 \le \mu_1 - \mu_2 \le 18.94$$

Section 10-4

10-37 $\bar{d} = 868.375$ $\quad s_d = 1290$, $n = 8$ $\quad$ 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 this confidence interval contains zero, we are 99% confident there is no significant difference between the two brands of tire.

10-39. 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.

Section 10-5

10-47. 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) $H_1 : \sigma_1^2 \ne \sigma_2^2$

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) $n_1 = 10$ $\quad n_2 = 16$
$s_1 = 4.7$ $\quad s_2 = 5.8$

$$f_0 = \frac{(4.7)^2}{(5.8)^2} = 0.657$$

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.

10-51 a) 90% confidence interval for the ratio of variances:

$$\left(\frac{s_1^2}{s_2^2}\right) f_{1-\alpha/2, n_1-1, n_2-1} \le \frac{\sigma_1^2}{\sigma_2^2} \le \left(\frac{s_1^2}{s_2^2}\right) f_{\alpha/2, n_1-1, n_2-1}$$

$$\left(\frac{(0.6)^2}{(0.8)^2}\right) 0.156 \le \frac{\sigma_1^2}{\sigma_2^2} \le \left(\frac{(0.6)^2}{(0.8)^2}\right) 6.39$$

$$0.08775 \le \frac{\sigma_1^2}{\sigma_2^2} \le 3.594$$

b) 95% confidence interval:

$$\left(\frac{s_1^2}{s_2^2}\right) f_{1-\alpha/2,n_1-1,n_2-1} \le \frac{\sigma_1^2}{\sigma_2^2} \le \left(\frac{s_1^2}{s_2^2}\right) f_{\alpha/2,n_1-1,n_2-1}$$

$$\left(\frac{(0.6)^2}{(0.8)^2}\right) 0.104 \le \frac{\sigma_1^2}{\sigma_2^2} \le \left(\frac{(0.6)^2}{(0.8)^2}\right) 9.60$$

$$0.0585 \le \frac{\sigma_1^2}{\sigma_2^2} \le 5.4$$

The 95% confidence interval is wider than the 90% confidence interval.

c) 90% lower-sided confidence interval:

$$\left(\frac{s_1^2}{s_2^2}\right) f_{1-\alpha,n_1-1,n_2-1} \le \frac{\sigma_1^2}{\sigma_2^2}$$

$$\left(\frac{(0.6)^2}{(0.8)^2}\right) 0.243 \le \frac{\sigma_1^2}{\sigma_2^2}$$

$$0.137 \le \frac{\sigma_1^2}{\sigma_2^2}$$

10-55 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) $H_1: \sigma_1^2 \ne \sigma_2^2$

4) $\alpha = 0.01$

5) The test statistic is

$$f_0 = \frac{s_1^2}{s_2^2}$$

6) Reject the null hypothesis if $f_0 < f_{0.995,10,12}$ where $f_{0.995,10,12}$ =0.1766 or $f_0 > f_{0.005,10,12}$ where $f_{0.005,10,12} = 2.91$

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 0.1766 >0.2575 > 5.0855 do not reject the null hypothesis and conclude the thickness variances are not equal at the 0.01 level of significance.

10-59 1) The parameters of interest are the overall distance standard deviations, σ_1, σ_2

2) $H_0: \sigma_1^2 = \sigma_2^2$

3) $H_1: \sigma_1^2 \ne \sigma_2^2$

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) $n_1 = 10$ $n_2 = 10$ $s_1 = 8.03$ $s_2 = 10.04$

$$f_0 = \frac{(8.03)^2}{(10.04)^2} = 0.640$$

8) Since 0.248 < 0.640 < 4.04 do not reject the null hypothesis and conclude there is no evidence to support the claim that there is a difference in the standard deviation of the overall distance of the two brands at the 0.05 level of significance.

95% confidence interval:

$$\left(\frac{s_1^2}{s_2^2}\right) f_{1-\alpha/2,n_1-1,n_2-1} \le \frac{\sigma_1^2}{\sigma_2^2} \le \left(\frac{s_1^2}{s_2^2}\right) f_{\alpha/2,n_1-1,n_2-1}$$

$$(0.640)0.248 \le \frac{\sigma_1^2}{\sigma_2^2} \le (0.640)4.03$$

$$0.159 \le \frac{\sigma_1^2}{\sigma_2^2} \le 2.579$$

Since the value 1 is contained within this interval, we are 95% confident there is no significant difference in the standard deviation of the overall distance of the two brands at the 0.05 level of significance.

Section 10-6

10-61. 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) $n_1 = 300$ $n_2 = 300$

$x_1 = 15$ $x_2 = 8$

$\hat{p}_1 = 0.05$ $\hat{p}_2 = 0.0267$ $\hat{p} = \frac{15+8}{300+300} = 0.0383$

$$z_0 = \frac{0.05 - 0.0267}{\sqrt{0.0383(1-0.0383)\left(\frac{1}{300} + \frac{1}{300}\right)}} = 1.49$$

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

10-63. 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

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

10-67 95% confidence interval on the difference:

$$(\hat{p}_1-\hat{p}_2)-z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}+\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}\le p_1-p_2\le(\hat{p}_1-\hat{p}_2)+z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}+\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

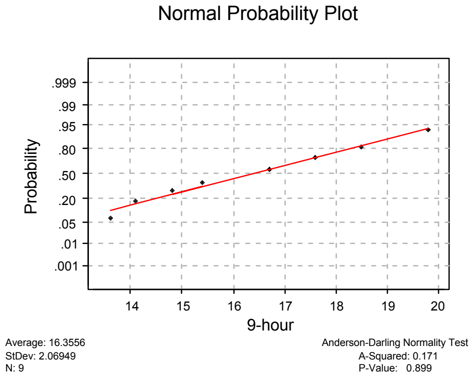
$$(0.77-0.6675)-1.96\sqrt{\frac{0.77(1-0.77)}{500}+\frac{0.6675(1-0.6675)}{400}}\le p_1-p_2\le(0.77-0.6675)+1.96\sqrt{\frac{0.77(1-0.77)}{500}+\frac{0.6675(1-0.6675)}{400}}$$

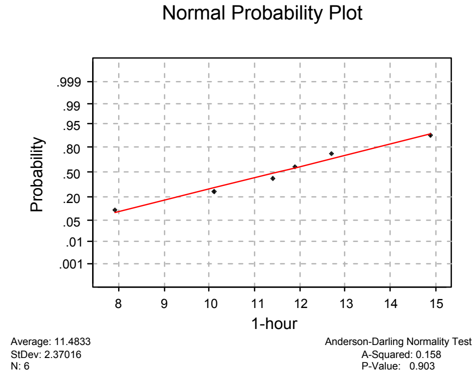
$$0.0434\le p_1-p_2\le 0.1616$$

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 0.0434 and 0.1616.

Supplemental Exercises

10-69 a) Assumptions that must be met are normality, equality of variance, independence of the observations and of the populations. 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. Independence of the observations for each sample is assumed. It is also reasonable to assume that the two populations are independent.





b) $\bar{x}_1 = 16.36$ $\quad\bar{x}_2 = 11.486$

$s_1 = 2.07$ $\quad s_2 = 2.37$

$n_1 = 9$ $\quad n_2 = 6$

99% confidence interval: $t_{\alpha/2,n_1+n_2-2} = t_{0.005,13}$ where $t_{0.005,13} = 3.012$

$$s_p = \sqrt{\frac{8(2.07)^2 + 5(2.37)^2}{13}} = 2.19$$

$$(\bar{x}_1 - \bar{x}_2) - t_{\alpha/2,n_1+n_2-2}(s_p)\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} \le \mu_1 - \mu_2 \le (\bar{x}_1 - \bar{x}_2) + t_{\alpha/2,n_1+n_2-2}(s_p)\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$

$$(16.36 - 11.486) - 3.012(2.19)\sqrt{\frac{1}{9}+\frac{1}{6}} \le \mu_1 - \mu_2 \le (16.36 - 11.486) + 3.012(2.19)\sqrt{\frac{1}{9}+\frac{1}{6}}$$

$1.40 \le \mu_1 - \mu_2 \le 8.36$

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).

10-73 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) $\quad \hat{p} = \frac{x_1 + x_2}{n_1 + n_2} = 0.000356$

$\hat{p}_2 = \frac{x_2}{n_2} = \frac{33}{200745} = 0.00016$ (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 $\alpha = 0.01$.

c) The conclusions are the same since z_0 is so large it exceeds $z_{\alpha/2}$ in both cases.

10-79.

$$n = \frac{\left(2.575\sqrt{\frac{(0.9+0.6)(0.1+0.4)}{2}} + 1.28\sqrt{0.9(0.1)+0.6(0.4)}\right)^2}{(0.9-0.6)^2}$$

$$= \frac{5.346}{0.09} = 59.4$$

$$n = 60$$

10-81. $H_0 : \mu_1 = \mu_2$

$H_1 : \mu_1 \neq \mu_2$

$n_1 = n_2 = n$

$\beta = 0.10$

$\alpha = 0.05$

Assume normal distribution and $\sigma_1^2 = \sigma_2^2 = \sigma^2$

$$\mu_1 = \mu_2 + \sigma$$

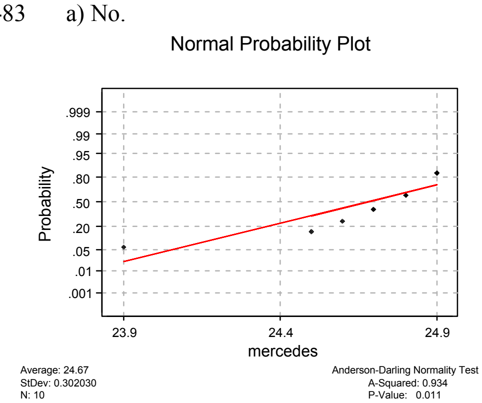
$$d = \frac{|\mu_1 - \mu_2|}{2\sigma} = \frac{\sigma}{2\sigma} = \frac{1}{2}$$

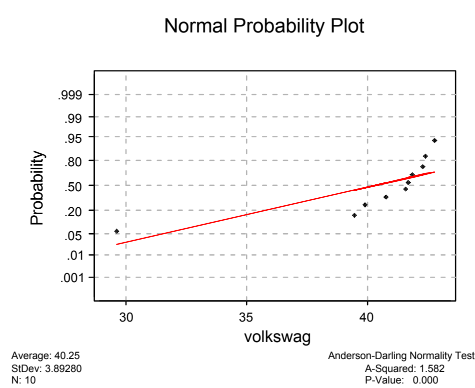
From Chart VI (e), $n^* = 50$

$$n = \frac{n^* + 1}{2} = \frac{50+1}{2} = 25.5$$

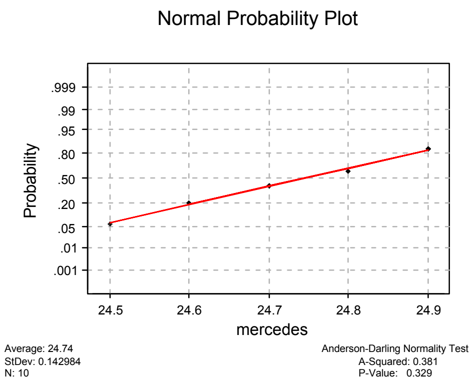
$n_1 = n_2 = 26$

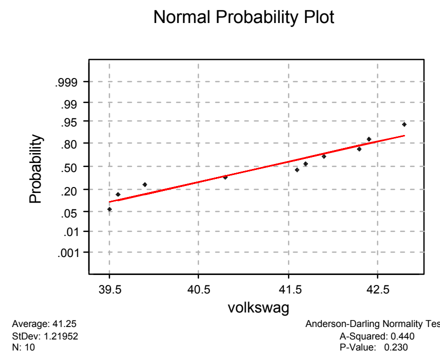
10-83 a) No.





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. Note that one unusual 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

$$s_V^2 = 1.49 \qquad f_{9,9,0.025} = 4.03$$

$$s_M^2 = 0.0204 \qquad f_{9,9,0.975} = \frac{1}{f_{9,9,0.025}} = \frac{1}{4.03} = 0.248$$

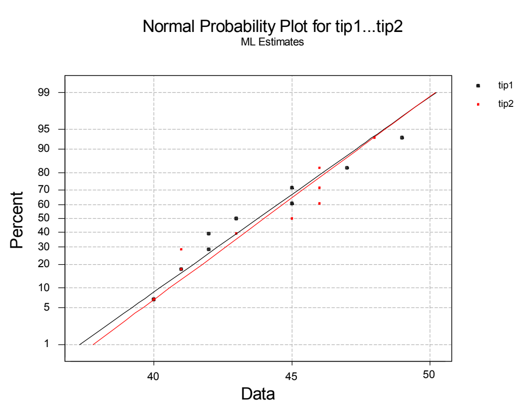
$$\left(\frac{s_V^2}{s_M^2}\right) f_{9,9,0.975} < \frac{\sigma_V^2}{\sigma_M^2} < \left(\frac{s_V^2}{s_M^2}\right) f_{9,9,0.025}$$

$$\left(\frac{1.49}{0.0204}\right) 0.248 < \frac{\sigma_V^2}{\sigma_M^2} < \left(\frac{1.49}{0.0204}\right) 4.03$$

$$18.124 < \frac{\sigma_V^2}{\sigma_M^2} < 294.35$$

Since the does not include the value of unity, we are 95% confident that there is evidence to reject the claim that the variability in mileage performance is different for the two types of vehicles. There is evidence that the variability is greater for a Volkswagen than for a Mercedes.

10-85 a) Underlying distributions appear to be normal since the data fall along a straight line on the normal probability plots. The slopes appear to be similar, so it is reasonable to assume that $\sigma_1^2 = \sigma_2^2$.



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

$$t_0 = \frac{(\bar{x}_1 - \bar{x}_2) - \delta}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

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) $\bar{x}_1 = 752.7$ $\quad \bar{x}_2 = 755.6$ $\quad s_p = \sqrt{\frac{9(1.252)^2 + 9(0.843)^2}{18}} = 1.07$

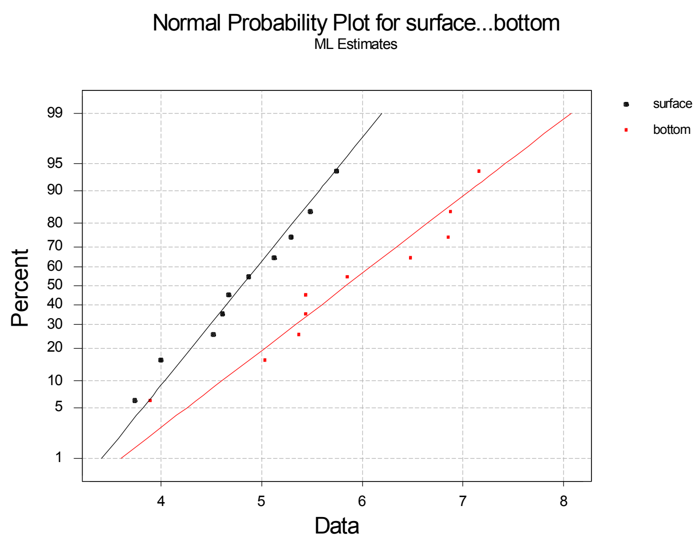
$s_1 = 1.252$ $\quad s_2 = 0.843$

$n_1 = 10$ $\quad n_2 = 10$

$$t_0 = \frac{(752.7 - 755.6) - 0}{1.07\sqrt{\frac{1}{10} + \frac{1}{10}}} = -6.06$$

8) Since $-6.06 < -2.101$, reject H_0 and conclude there is a significant difference between the two wineries with respect to the mean fill volumes.

10-89 a.) The data from both depths appear to be normally distributed, but the slopes are not equal. Therefore, it may not be assumed that $\sigma_1^2 = \sigma_2^2$.



b.) 1) The parameter of interest is the difference in mean HCB concentration, $\mu_1 - \mu_2$, with $\Delta_0 = 0$

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 $t_0 < -t_{0.025,15}$ where $-t_{0.025,15} = -2.131$ or $t_0 > t_{0.025,15}$ where $t_{0.025,15} =$ 2.131 since

$$v = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 - 1}} = 15.06$$

$$v \cong 15$$

(truncated)

7) $\bar{x}_1 = 4.804$ $\bar{x}_2 = 5.839$ $s_1 = 0.631$ $s_2 = 1.014$

$n_1 = 10$ $n_2 = 10$

$$t_0 = \frac{(4.804 - 5.839)}{\sqrt{\frac{(0.631)^2}{10} + \frac{(1.014)^2}{10}}} = -2.74$$

8) Since –2.74 < -2.131 reject the null hypothesis and conclude that the data support the claim that the mean HCB concentration is different at the two depths sampled at the 0.05 level of significance.

b) P-value = 2P(t < -2.74), 2(0.005) < P-value < 2(0.01)

0.001 < P-value < 0.02

c) $\Delta = 2$ $\alpha = 0.05$ $n_1 = n_2 = 10$ $d = \frac{2}{2(1)} = 1$

From Chart VI (e) we find $\beta = 0.20$, and then calculate Power = 1- β = 0.80

d.) $\Delta = 2$ $\alpha = 0.05$ $d = \frac{2}{2(1)} = 0.5$, $\beta = 0.0$

From Chart VI (e) we find n*=50 and $n = \frac{50 + 1}{2} = 25.5$, so n=26

Chapter 11 Selected Problem Solutions

Section 11-2

11-1. 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) $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$

$\hat{y} = 48.012962 - 2.3298017(4.3) = 37.99$

c) $\hat{y} = 48.012962 - 2.3298017(3.7) = 39.39$

d) $e = y - \hat{y} = 46.1 - 39.39 = 6.71$

11-5. 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
```

$\hat{\sigma}^2 = 8.76775$

If the calculations were to be done by hand use Equations (11-7) and (11-8).

$\hat{y} = 13.3202 + 3.32437x$

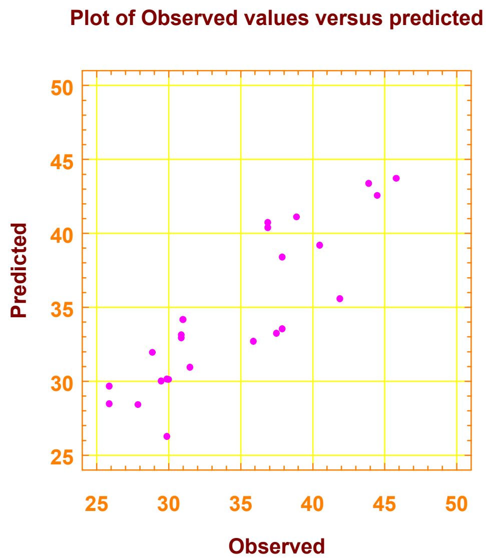
b) $\hat{y} = 13.3202 + 3.32437(7.5) = 38.253$

c) $\hat{y} = 13.3202 + 3.32437(5.8980) = 32.9273$

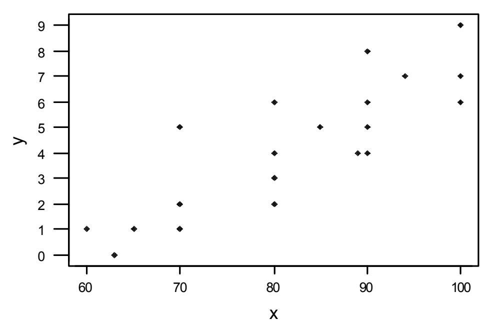
$\hat{y} = 32.9273$

$e = y - \hat{y} = 30.9 - 32.9273 = -2.0273$

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 indicates that the simple linear regression model provides a reasonable fit to the data.



11-9. a) Yes, a linear regression would seem appropriate, but one or two points appear to be outliers.



Predictor	Coef	SE Coef	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%

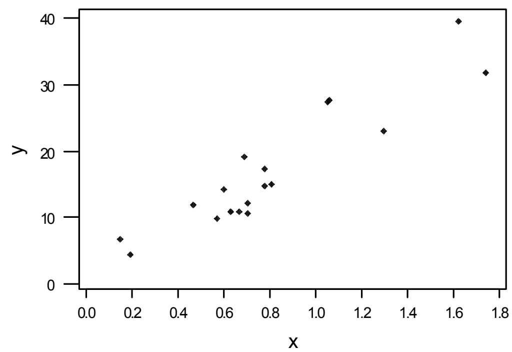
Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	88.520	88.520	44.66	0.000
Residual Error	18	35.680	1.982		
Total	19	124.200			

b) $\hat{\sigma}^2 = 1.9818$ and $\hat{y} = -9.8131 + 0.171484x$

c) $\hat{y} = 4.76301$ at x = 85

11-11. a) Yes, a linear regression would seem appropriate.



Predictor	Coef	SE Coef	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
Residual Error	16	220.9	13.8		
Total	17	1494.5			

b) $\hat{\sigma}^2 = 13.81$

$\hat{y} = 0.470467 + 20.5673x$

c) $\hat{y} = 0.470467 + 20.5673(1) = 21.038$

d) $\hat{y} = 10.1371 \quad e = 1.6629$

Section 11-4

11-21. Refer to ANOVA of Exercise 11-5

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 11-5

$$t_0 = \frac{3.32437}{0.390276} = 8.518$$

8) Since 8.518 > 2.074 reject H_0 and conclude the model is useful $\alpha = 0.05$.

b) 1) The parameter of interest is the slope, β_1

2) $H_0\!:\beta_1 = 0$

3) $H_1\!:\beta_1 \neq 0$

4) $\alpha = 0.05$

5) The test statistic is $f_0 = \dfrac{MS_R}{MS_E} = \dfrac{SS_R / 1}{SS_E / (n-2)}$

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 at a significance $\alpha = 0.05$.

The F-statistic is the square of the t-statistic. The F-test is a restricted to a two-sided test, whereas the t-test could be used for one-sided alternative hypotheses.

c) $se(\hat{\beta}_1) = \sqrt{\dfrac{\hat{\sigma}^2}{S_{xx}}} = \sqrt{\dfrac{8.7675}{57.5631}} = .39027$

$$se(\hat{\beta}_0) = \sqrt{\hat{\sigma}^2\left[\frac{1}{n} + \frac{\bar{x}}{S_{xx}}\right]} = \sqrt{8.7675\left[\frac{1}{24} + \frac{6.4049^2}{57.5631}\right]} = 2.5717$$

d) 1) The parameter of interest is the intercept, β_0.

2) $H_0 : \beta_0 = 0$

3) $H_1 : \beta_0 \neq 0$

4) $\alpha = 0.05$, using t-test

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 $-t_{0.025,22} = -2.074$ or $t_0 > t_{0.025,22} = 2.074$

7) Using the results from Exercise 11-5

$$t_0 = \frac{13.3201}{2.5717} = 5.2774$$

8) Since $5.2774 > 2.074$ reject H_0 and conclude the intercept is not zero at $\alpha = 0.05$.

11-25. Refer to ANOVA of Exercise 11-9

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, reject H_0. P-value = 0.000003.

b) $se(\hat{\beta}_1) = 0.0256613$

$se(\hat{\beta}_0) = 2.13526$

c) $H_0 : \beta_0 = 0$

$H_1 : \beta_0 \neq 0$

$\alpha = 0.05$

$$t_0 = -4.59573$$
$$t_{.025,18} = 2.101$$
$$|t_0| > t_{\alpha/2,18}$$
Therefore, reject H_0. P-value = 0.00022.

Sections 11-5 and 11-6

11-31. $t_{\alpha/2,n-2} = t_{0.025,12} = 2.179$

a) 95% confidence interval on β_1.

$$\hat{\beta}_1 \pm t_{\alpha/2,n-2} se(\hat{\beta}_1)$$
$$-2.3298 \pm t_{.025,12}(0.2697)$$
$$-2.3298 \pm 2.179(0.2697)$$
$$-2.9175. \le \beta_1 \le -1.7421.$$

b) 95% confidence interval on β_0.

$$\hat{\beta}_0 \pm t_{.025,12} se(\hat{\beta}_0)$$
$$48.0130 \pm 2.179(0.5959)$$
$$46.7145 \le \beta_0 \le 49.3114.$$

c) 95% confidence interval on μ when $x_0 = 2.5$.

$$\hat{\mu}_{Y|x_0} = 48.0130 - 2.3298(2.5) = 42.1885$$
$$\hat{\mu}_{Y|x_0} \pm t_{.025,12}\sqrt{\hat{\sigma}^2\left(\tfrac{1}{n} + \tfrac{(x_0-\bar{x})^2}{S_{xx}}\right)}$$
$$42.1885 \pm (2.179)\sqrt{1.844\left(\tfrac{1}{14} + \tfrac{(2.5-3.0714)^2}{25.3486}\right)}$$
$$42.1885 \pm 2.179(0.3943)$$
$$41.3293 \le \hat{\mu}_{Y|x_0} \le 43.0477$$

d) 95% on prediction interval when $x_0 = 2.5$.

$$\hat{y}_0 \pm t_{.025,12}\sqrt{\hat{\sigma}^2\left(1 + \tfrac{1}{n} + \tfrac{(x_0-\bar{x})^2}{S_{xx}}\right)}$$
$$42.1885 \pm 2.179\sqrt{1.844\left(1 + \tfrac{1}{14} + \tfrac{(2.5-3.0714)^2}{25.348571}\right)}$$
$$42.1885 \pm 2.179(1.1808)$$
$$38.2489 \le y_0 \le 46.1281$$

It is wider because it depends on both the error associated with the fitted model as well as that with the future observation.

11-35.

```
          99 percent confidence intervals for coefficient estimates
--------------------------------------------------------------------------------
                          Estimate  Standard error     Lower Limit     Upper Limit
CONSTANT                  -6.33550         1.66765        -11.6219        -1.05011
Temperature                9.20836         0.03377         9.10130         9.93154
--------------------------------------------------------------------------------
```

a) $9.10130 \le \beta_1 \le 9.31543$

b) $-11.6219 \le \beta_0 \le -1.04911$

c) $500.124 \pm (2.228)\sqrt{3.774609\left(\tfrac{1}{12} + \tfrac{(55-46.5)^2}{3308.9994}\right)}$

$$500.124 \pm 1.4037586$$
$$498.72024 \le \hat{\mu}_{Y|x_0} \le 501.52776$$

d) $500.124 \pm (2.228)\sqrt{3.774609\left(1 + \tfrac{1}{12} + \tfrac{(55-46.5)^2}{3308.9994}\right)}$

$$500.124 \pm 4.5505644$$

$$495.57344 \le y_0 \le 504.67456$$

It is wider because the prediction interval includes error for both the fitted model and from that associated with the future observation.

11-41 a) $-43.1964 \le \beta_1 \le -30.7272$

b) $2530.09 \le \beta_0 \le 2720.68$

c) $1886.154 \pm (2.101)\sqrt{9811.21(\frac{1}{20} + \frac{(20-13.3375)^2}{1114.6618})}$

1886.154 ± 62.370688

$1823.7833 \le \mu_{y|x_0} \le 1948.5247$

d) $1886.154 \pm (2.101)\sqrt{9811.21(1 + \frac{1}{20} + \frac{(20-13.3375)^2}{1114.6618})}$

1886.154 ± 217.25275

$1668.9013 \le y_0 \le 2103.4067$

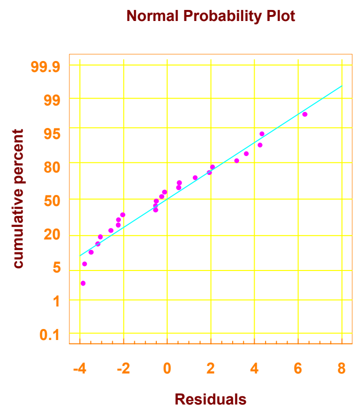
Section 11-7

11-43. Use the Results of exercise 11-5 to answer the following questions.

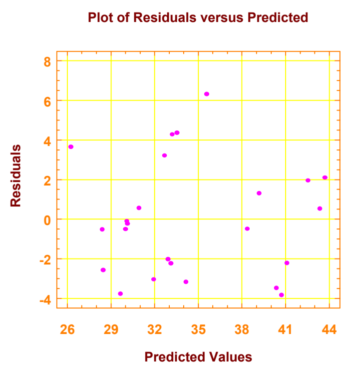
a)

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.



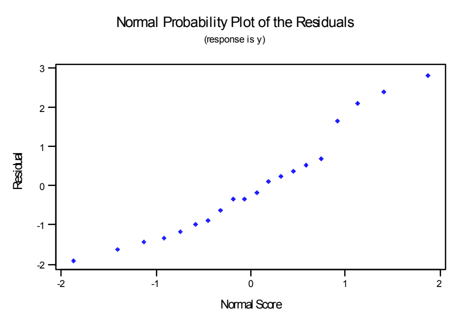
c) No serious departure from assumption of constant variance. This is evident by the random pattern of the residuals.



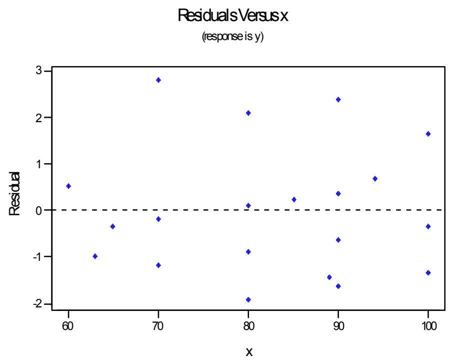
d) $R^2 \equiv 76.73\%$;

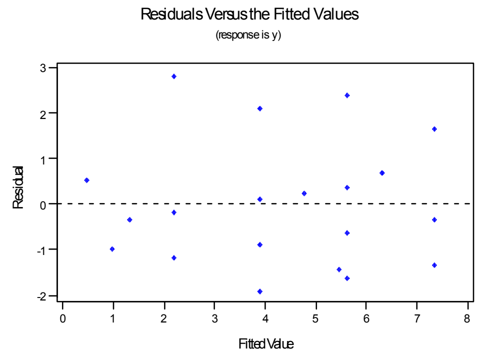
11-47. a) $R^2 = 71.27\%$

b) No major departure from normality assumptions.



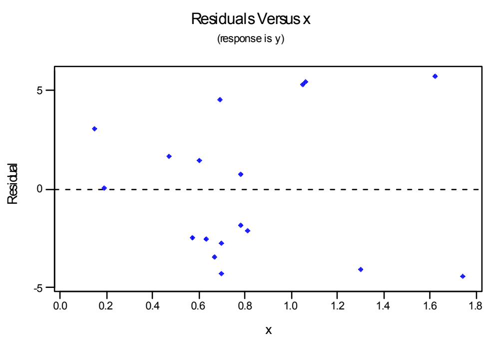
c) Assumption of constant variance appears reasonable.

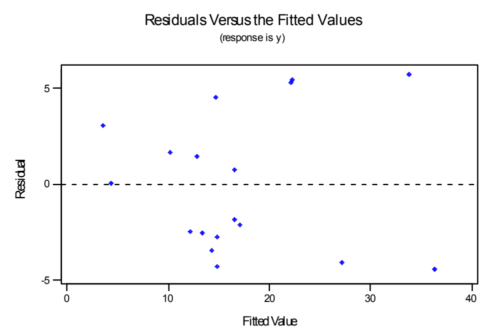




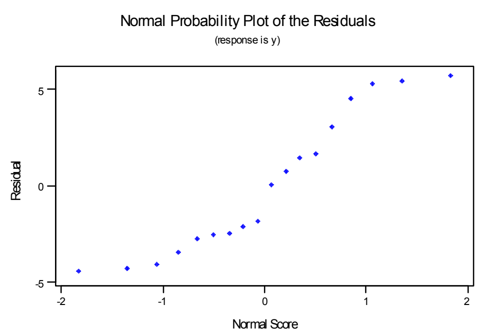
11-49. a) $R^2 = 85.22\%$

b) Assumptions appear reasonable, but there is a suggestion that variability increases with $\hat{y}$.





c) Normality assumption may be questionable. There is some " bending" away from a straight line in the tails of the normal probability plot.



Section 11-10

11-55. a) $\hat{y} = -0.0280411 + 0.990987x$

b) $H_0 : \beta_1 = 0$

$H_1 : \beta_1 \neq 0$ $\alpha = 0.05$

$f_0 = 79.838$

$f_{.05,1,18} = 4.41$

$f_0 >> f_{\alpha,1,18}$

Reject H_0.

c) $r = \sqrt{0.816} = 0.903$

d) $H_0 : \rho = 0$

$H_1 : \rho \neq 0$ $\alpha = 0.05$

$t_0 = \frac{R\sqrt{n-2}}{\sqrt{1-R^2}} = \frac{0.90334\ \sqrt{18}}{\sqrt{1-0.816}} = 8.9345$

$t_{.025,18} = 2.101$

$t_0 > t_{\alpha/2,18}$

Reject H_0.

e) $H_0 : \rho = 0.5$

$H_1 : \rho \neq 0.5$ $\alpha = 0.05$

$z_0 = 3.879$

$z_{.025} = 1.96$

$z_0 > z_{\alpha/2}$

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 $z_{.025} = 1.96$.

$0.7677 \leq \rho \leq 0.9615$.

11-59 n = 50 r = 0.62

a) $H_0 : \rho = 0$

$H_1 : \rho \neq 0$ $\alpha = 0.01$

$t_0 = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \frac{0.62\sqrt{48}}{\sqrt{1-(0.62)^2}} = 5.475$

$t_{.005,48} = 2.683$

$t_0 > t_{0.005,48}$

Reject H_0. P-value $\cong 0$

b) $\tanh(\text{arctanh } 0.62 - \frac{z_{.005}}{\sqrt{47}}) \leq \rho \leq \tanh(\text{arctanh } 0.62 + \frac{z_{.005}}{\sqrt{47}})$ where $z_{.005} = 2.575$.

$0.3358 \leq \rho \leq 0.8007$.

c) Yes.

11-61. a) $r = 0.933203$

a) $H_0 : \rho = 0$

$H_1 : \rho \neq 0$ $\alpha = 0.05$

$t_0 = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \frac{0.933203\ \sqrt{15}}{\sqrt{1-(0.8709)}} = 10.06$

$t_{.025,15} = 2.131$

$t_0 > t_{\alpha/2,15}$

Reject H_0.

c) $\hat{y} = 0.72538 + 0.498081x$

$H_0 : \beta_1 = 0$

$H_1 : \beta_1 \neq 0$ $\alpha = 0.05$

$f_0 = 101.16$

$f_{.05,1,15} = 4.545$

$f_0 >> f_{\alpha,1,15}$

Reject H_0. Conclude that the model is significant at $\alpha = 0.05$. This test and the one in part b are identical.

d) $H_0: \beta_0 = 0$

$H_1: \beta_0 \neq 0 \qquad \alpha = 0.05$

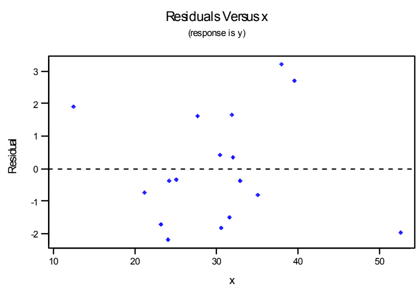
$t_0 = 0.468345$

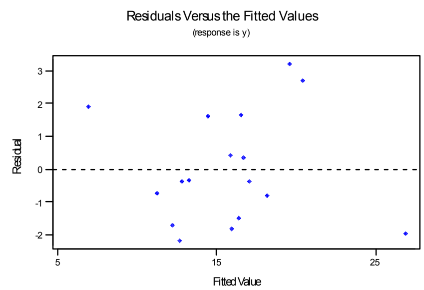
$t_{.025,15} = 2.131$

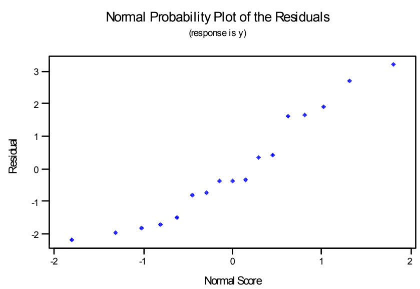
$t_0 \not> t_{\alpha/2,15}$

Do not reject H_0. We cannot conclude β_0 is different from zero.

e) No serious problems with model assumptions are noted.







Supplemental

11-65. a) $\hat{y} = 93.34 + 15.64x$

b) $H_0: \beta_1 = 0$

$H_1: \beta_1 \neq 0 \qquad \alpha = 0.05$

$f_0 = 12.872$

$f_{0.05,1,14} = 4.60$

$f_0 > f_{0.05,1,14}$

Reject H_0. Conclude that $\beta_1 \neq 0$ at $\alpha = 0.05$.

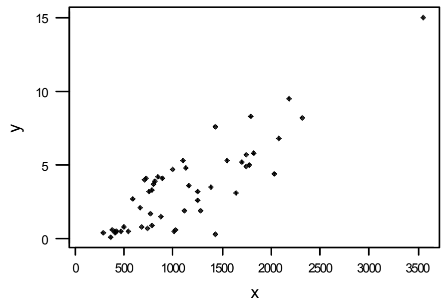
c) $(7.961 \leq \beta_1 \leq 23.322)$

d) $(74.758 \leq \beta_0 \leq 111.923)$

e) $\hat{y} = 93.34 + 15.64(2.5) = 132.44$

$$132.44 \pm 2.145\sqrt{136.27\left[\tfrac{1}{16}+\tfrac{(2.5-2.325)^2}{7.017}\right]}$$
$$132.44 \pm 6.26$$
$$126.18 \le \hat{\mu}_{Y|x_0=2.5} \le 138.70$$

11-67 a)



b) $\hat{y} = -0.8819 + 0.00385x$

c) $H_0 : \beta_1 = 0$

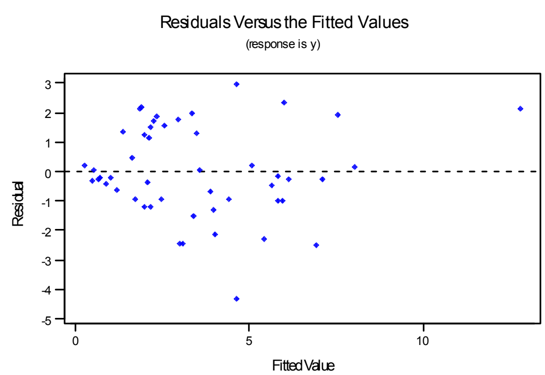
$H_1 : \beta_1 \ne 0$ $\quad\quad \alpha = 0.05$

$f_0 = 122.03$

$f_0 > f_{\alpha,1,48}$

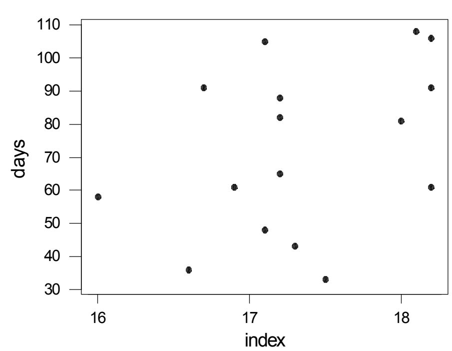
Reject H_0. Conclude that regression model is significant at $\alpha = 0.05$

d) No, it seems the variance is not constant, there is a funnel shape.



e) $\hat{y}^* = 0.5967 + 0.00097x$. Yes, the transformation stabilizes the variance.

11-71 a)



b) The regression equation is
$\hat{y} = -193 + 15.296x$

Predictor	Coef	SE Coef	T	P
Constant	-193.0	163.5	-1.18	0.258
x	15.296	9.421	1.62	0.127

S = 23.79   R-Sq = 15.8%   R-Sq(adj) = 9.8%

Analysis of Variance

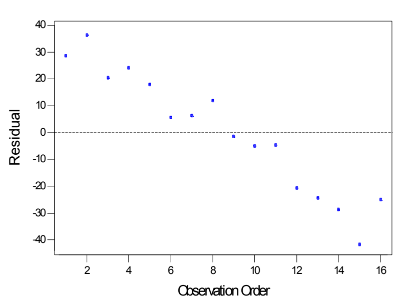
Source	DF	SS	MS	F	P
Regression	1	1492.6	1492.6	2.64	0.127
Error	14	7926.8	566.2		
Total	15	9419.4			

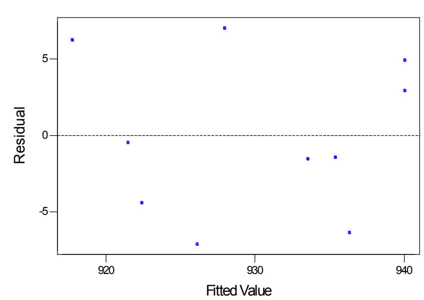
Cannot reject H_o; therefore we conclude that the model is not significant. Therefore the seasonal meteorological index (x) is not a reliable predictor of the number of days that the ozone level exceeds 0.20 *ppm* (y).

c) 95% CI on β_1

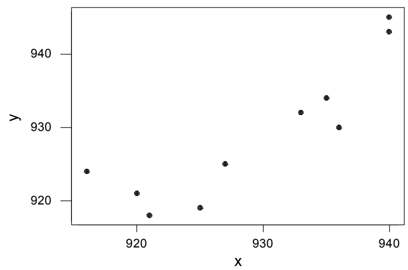
$$\hat{\beta}_1 \pm t_{\alpha/2,n-2} se(\hat{\beta}_1)$$
$$15.296 \pm t_{.025,12}(9.421)$$
$$15.296 \pm 2.145(9.421)$$
$$-4.912 \le \beta_1 \le 35.504$$

d) The normality plot of the residuals is satisfactory. However, the plot of residuals versus run order exhibits a strong downward trend. This could indicate that there is another variable should be included in the model, one that changes with time.





11-75 a)



b) $\hat{y} = 33.3 + 0.9636x$

c)

Predictor	Coef	SE Coef	T	P
Constant	33.3	171.7	0.19	0.851
x	0.9639	0.1848	5.22	0.001

S = 4.805 R-Sq = 77.3% R-Sq(adj) = 74.4%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	628.18	628.18	27.21	0.001
Residual Error	8	184.72	23.09		
Total	9	812.90			

Reject the hull hypothesis and conclude that the model is significant. 77.3% of the variability is explained by the model.

d) $H_0 : \beta_1 = 1$

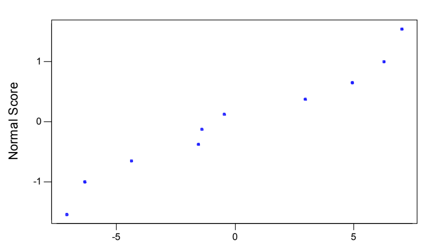
$H_1 : \beta_1 \neq 1$ α=.05

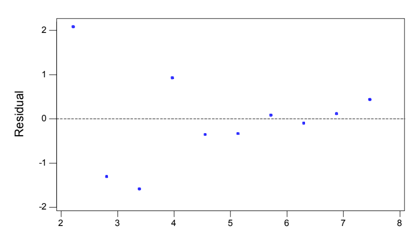
$$t_0 = \frac{\hat{\beta}_1 - 1}{se(\hat{\beta}_1)} = \frac{0.9639 - 1}{0.1848} = -0.1953$$

$$t_{a/2,n-2} = t_{.025,8} = 2.306$$

Since $t_0 > -t_{a/2,n-2}$, we cannot reject H_o and we conclude that there is not enough evidence to reject the claim that the devices produce different temperature measurements. Therefore, we assume the devices produce equivalent measurements.

e) The residual plots to not reveal any major problems.





Chapter 12 Selected Problem Solutions

Section 12-1

12-1. 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) $\hat{\beta} = \begin{bmatrix} 171.054 \\ 3.713 \\ -1.126 \end{bmatrix}$, so $\hat{y} = 171.054 + 3.714x_1 - 1.126x_2$

c) $\hat{y} = 171.054 + 3.714(18) - 1.126(43) = 189.481$

12-5.

```
Predictor       Coef       StDev          T        P
Constant      33.449       1.576      21.22    0.000
x1         -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) $\hat{\sigma}^2 = 8.03$

c) $\hat{y} = 33.4491 - 0.05435(300) + 1.07822(2) = 19.30$ mpg.

12-7.

```
Predictor          Coef      SE Coef          T        P
Constant         383.80        36.22      10.60    0.002
X1              -3.6381       0.5665      -6.42    0.008
X2             -0.11168      0.04338      -2.57    0.082
S = 12.35        R-Sq = 98.5%     R-Sq(adj) = 97.5%

Analysis of Variance
Source              DF          SS          MS          F         P
Regression           2       29787       14894      97.59     0.002
Residual Error       3         458         153
Total                5       30245
```

a) $\hat{y} = 383.80 - 3.6381x_1 - 0.1119x_2$

b) $\hat{\sigma}^2 = 153.0$, $se(\hat{\beta}_0) = 36.22$, $se(\hat{\beta}_1) = 0.5665$, and $se(\hat{\beta}_2) = .04338$

c) $\hat{y} = 383.80 - 3.6381(25) - 0.1119(1000) = 180.95$

d)

```
Predictor          Coef      SE Coef          T        P
Constant          484.0        101.3       4.78    0.041
X1               -7.656        3.846      -1.99    0.185
X2              -0.2221       0.1129      -1.97    0.188
X1*X2          0.004087     0.003871       1.06    0.402
S = 12.12        R-Sq = 99.0%     R-Sq(adj) = 97.6%
Analysis of Variance
Source              DF          SS          MS          F         P
Regression           3     29951.4      9983.8      67.92     0.015
Residual Error       2       294.0       147.0
Total                5     30245.3
```

$\hat{y} = 484.0 - 7.656x_1 - 0.222x_2 - 0.0041x_{12}$

e) $\hat{\sigma}^2 = 147.0$, $se(\hat{\beta}_0) = 101.3$, $se(\hat{\beta}_1) = 3.846$, $se(\hat{\beta}_2) = 0.113$ and $se(\hat{\beta}_{12}) = 0.0039$

f) $\hat{y} = 484.0 - 7.656(25) - 0.222(1000) - 0.0041(25)(1000) = -31.3$

The predicted value is smaller

12-9.

```
Predictor        Coef      SE Coef         T        P
Constant        47.17        49.58      0.95    0.356
x1             -9.735        3.692     -2.64    0.018
x2             0.4283       0.2239      1.91    0.074
x3             18.237        1.312     13.90    0.000

S = 3.480       R-Sq = 99.4%     R-Sq(adj) = 99.3%

Analysis of Variance
Source            DF          SS          MS         F         P
Regression         3       30532       10177    840.55     0.000
Residual Error    16         194          12
Total             19       30725
```

a) $y = 47.174 - 9.7352x_1 + 0.4283x_2 + 18.2375x_3$

b) $\hat{\sigma}^2 = 12$

c) $se(\hat{\beta}_0) = 49.5815$, $se(\hat{\beta}_1) = 3.6916$, $se(\hat{\beta}_2) = 0.2239$, and $se(\hat{\beta}_3) = 1.312$

d) $y = 47.174 - 9.7352(14.5) + 0.4283(220) + 18.2375(5) = 91.43$

Section 12-2

12-13. 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

$$SS_T = 371595.6 - \frac{(1916)^2}{10} = 4490$$

$$\mathbf{X'y} = \begin{bmatrix} \sum y_i \\ \sum x_{i1}y_i \\ \sum x_{i2}y_i \end{bmatrix} = \begin{bmatrix} 1030 \\ 21310 \\ 44174 \end{bmatrix}$$

$$\hat{\boldsymbol{\beta}}'\mathbf{X'y} = \begin{bmatrix} 171.054 & 3.713 & -1.126 \end{bmatrix} \begin{bmatrix} 1916 \\ 43550.8 \\ 104736.8 \end{bmatrix} = 371535.9$$

$$SS_R = 371535.9 - \frac{1916^2}{10} = 4430.38$$

$$SS_E = SS_T - SS_R = 4490 - 4430.38 = 59.62$$

$$f_0 = \frac{\frac{SS_R}{k}}{\frac{SS_E}{n-p}} = \frac{4430.38/2}{59.62/7} = 260.09$$

$$f_{0.05,2,7} = 4.74$$

$$f_0 > f_{0.05,2,7}$$

Reject H_0 and conclude that the regression model is significant at $\alpha = 0.05$.

b) $H_0: \beta_1 = 0$ $\quad\quad\quad\quad$ $\beta_2 = 0$

$H_1 : \beta_1 \neq 0$ $\qquad\qquad\qquad$ $\beta_2 \neq 0$

$$t_0 = \frac{\hat{\beta}_1}{se(\hat{\beta}_1)} \qquad\qquad t_0 = \frac{\hat{\beta}_2}{se(\hat{\beta}_2)}$$

$$= \frac{3.713}{0.1934} = 19.20 \qquad\qquad = \frac{-1.126}{0.0861} = -13.08$$

$t_{\alpha/2,7} = t_{.025,7} = 2.365$

Reject H_0 $\qquad\qquad\qquad$ Reject H_0

Both regression coefficients are significant

12-17. 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 β_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 β_6 is significant at $\alpha = 0.05$.

No, only x_1 contributes significantly to the regression.

12-21. a) $H_0 : \beta_1 = \beta_2 = \beta_{12} = 0$

H_1 at least one $\beta_j \neq 0$

$\alpha = 0.05$

$f_0 = 67.92$

$f_{\alpha,3,2} = f_{.05,3,2} = 19.16$

$f_0 \not> f_{\alpha,3,2}$

Reject H_0

b) $H_0 : \beta_{12} = 0$

$H_1 : \beta_{12} \neq 0$

$\alpha = 0.05$

$$SSR(\beta_{12} \mid \beta_1, \beta_2) = 29951.4 - 29787 = 164.4$$
$$f_0 = \frac{SSR}{MS_E} = \frac{164.4}{153} = 1.07$$
$$f_{.05,1,2} = 18.51$$
$$f_0 \not> f_{\alpha,1,2}$$
Do not reject H_0

c) $\hat{\sigma}^2 = 147.0$

$\hat{\sigma}^2$ (no interaction term) = 153.0

$MS_E(\hat{\sigma}^2)$ was reduced in the interaction term model due to the addition of this term.

12-23. a) $H_0: \beta_1 = \beta_2 = \beta_3 = 0$ 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) $\alpha = 0.05$ $\quad t_{\alpha/2,n-p} = t_{.025,16} = 2.12$

$H_0: \beta_1 = 0$	$\beta_2 = 0$	$\beta_3 = 0$
$H_1: \beta_1 \neq 0$	$\beta_2 \neq 0$	$\beta_3 \neq 0$
$t_0 = -2.637$	$t_0 = 1.91$	$t_0 = 13.9$
$\lvert t_0 \rvert > t_{\alpha/2,16}$	$\lvert t_0 \rvert \not> t_{\alpha/2,16}$	$\lvert t_0 \rvert > t_{\alpha/2,16}$
Reject H_0	Do not reject H_0	Reject H_0

Sections 12-3 and 12-4

12-27. a) $-0.00657 \leq \beta_8 \leq -0.00122$

b) $\sqrt{\hat{\sigma}^2 \mathbf{x}_0'(\mathbf{X'X})^{-1}\mathbf{x}_0} = 0.497648 = se(\hat{\mu}_{Y|x_0})$

c) $\hat{\mu}_{Y|x_0} = -7.63449 + 0.00398(2000) + 0.24777(60) - 0.00389(1800) = 8.19$

$$\hat{\mu}_{Y|x_0} \pm t_{.025,24}\ se(\hat{\mu}_{Y|x_0})$$
$$8.19 \pm (2.064)(0.497648)$$
$$8.19 \pm 1.03$$
$$7.16 \leq \mu_{Y|x_0} \leq 9.22$$

12-29. a) 95 % CI on coeficients

$$\beta_1 \pm t_{a/2,n-p}(\hat{\beta}_1)$$
$$0.0972 \leq \beta_1 \leq 1.4174$$
$$-1.9646 \leq \beta_2 \leq 17.0026$$
$$-1.7953 \leq \beta_3 \leq 6.7613$$
$$-1.7941 \leq \beta_4 \leq 0.8319$$

b) $\hat{\mu}_{Y|x_0} = 290.44$ $se(\hat{\mu}_{Y|x_0}) = 7.61$ $t_{.025,7} = 2.365$

$$\hat{\mu}_{Y|x_0} \pm t_{\alpha/2,n-p} \ se(\hat{\mu}_{Y|x_0})$$
$$290.44 \pm (2.365)(7.61)$$
$$272.44 \le \mu_{Y|x_0} \le 308.44$$

c) $\hat{y}_0 \pm t_{\alpha/2,n-p}\sqrt{\hat{\sigma}^2(1+\mathbf{x}_0'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_0)}$

$$290.44 \pm 2.365(14.038)$$
$$257.25 \le y_0 \le 323.64$$

12-31 a)95% Confidence Interval on coefficients

$$-0.595 \le \beta_2 \le 0.535$$
$$0.229 \le \beta_3 \le 0.812$$
$$-0.216 \le \beta_4 \le 0.013$$
$$-7.982 \le \beta_5 \le 2.977$$

b) $\hat{\mu}_{Y|x_0} = 8.99568$ $se(\hat{\mu}_{Y|x_0}) = 0.472445$ $t_{.025,14} = 2.145$

$$\hat{\mu}_{Y|x_0} \pm t_{\alpha/2,n-p} \ se(\hat{\mu}_{Y|x_0})$$
$$8.99568 \pm (2.145)(0.472445)$$
$$7.982 \le \mu_{Y|x_0} \le 10.009$$

c) $y_0 = 8\ 99568$ $se(\hat{y}_0) = 1.00121$

$$8.99568 \pm 2.145(1.00121)$$
$$6.8481 \le y_0 \le 11.143$$

12-35. a) $0.3882 \le \beta_{Pts} \le 0.5998$

b) $\hat{y} = -5.767703 + 0.496501 x_{Pts}$

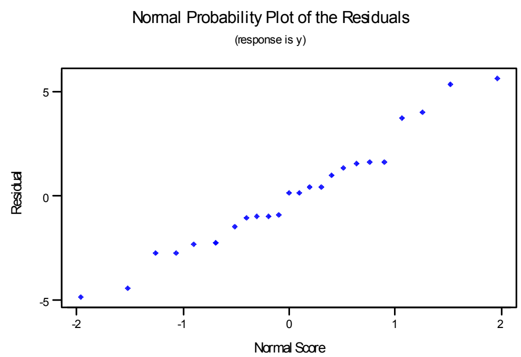
c) $0.4648 \le \beta_{Pts} \le 0.5282$

d) The simple linear regression model has the shorter interval. Yes, the simple linear regression model in this case is preferable.

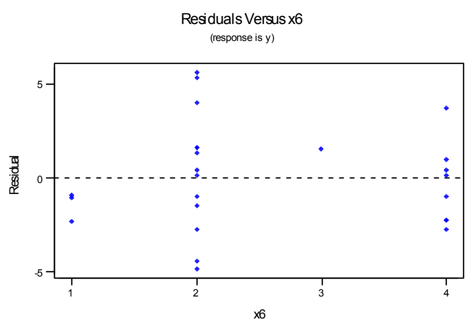
Section 12-5

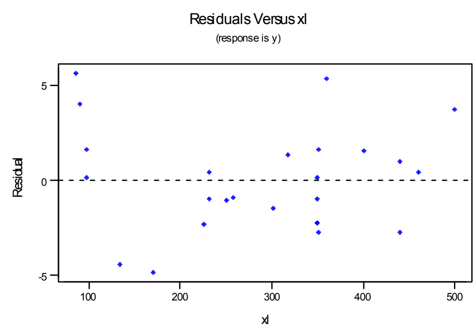
12-37. a) $r^2 = 0.82897$

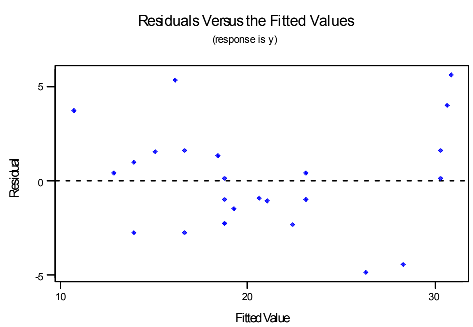
b) Normality assumption appears valid.



c) Assumption of constant variance appears reasonable.







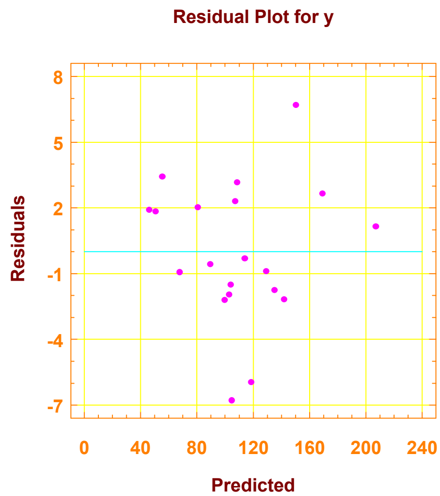
d) Yes, observations 7, 10, and 18

12-39. a) $r^2 = 0.985$

b) $r^2 = 0.990$

r^2 increases with addition of interaction term. No, adding additional regressor will always increase r^2

12-41 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)

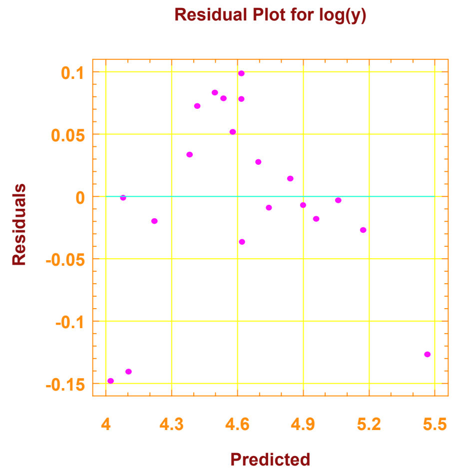
Model fitting results for: log(y)

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.

$$\hat{y}^* = 6.22489 - 0.16647x_1 - 0.000228x_2 + 0.157312x_3$$

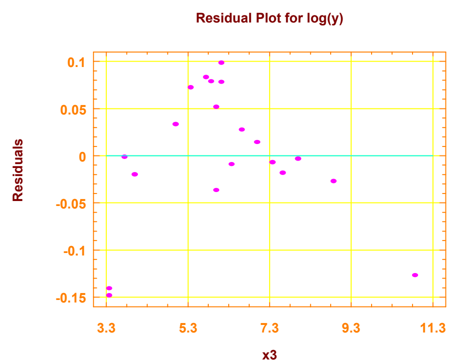
d)



Plot exhibits curvature

There is curvature in the plot. The plot does not give much more information as to which model is preferable.

e)

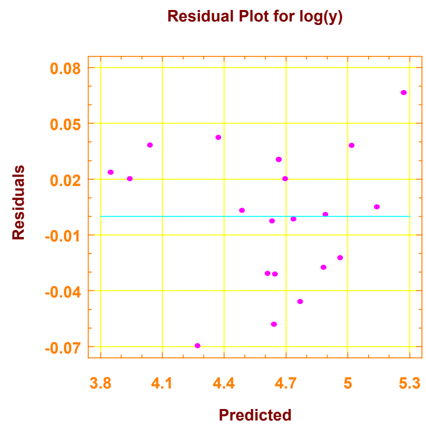


Plot exhibits curvature

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
```



Using 1/x3

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 response data.

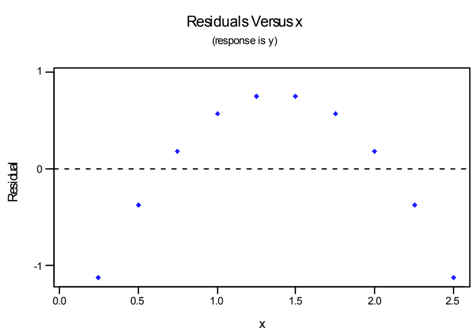
Section 12-6

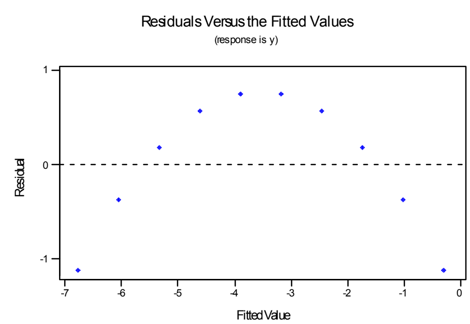
12-47. a) $\hat{y} = -1.633 + 1.232x - 1.495x^2$

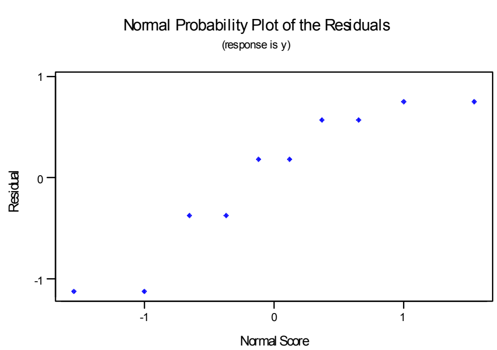
b) $f_0 = 1858613$, reject H_0

c) $t_0 = -601.64$, reject H_0

d) Model is acceptable, observation number 10 has large leverage.







12-49. $\hat{y} = 759.395 - 90.783x' - 47.166(x')^2$, 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) $\hat{y} = 759.395 - 90.783\left(\frac{x-297.125}{11.9336}\right) - 47.166\left(\frac{x-297.125}{11.9336}\right)^2$

$$\hat{y} = 759.395 - 7.607(x - 297.125) - 0.331(x - 297.125)^2$$

$$\hat{y} = -26204.14 + 189.09x - 0.331x^2$$

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 total variability explained is the same for both standardized and un-standardized models. Therefore, R^2 is the <u>same</u> for both models.

$y' = \beta_0^* + \beta_1^* x' + \beta_{11}^*(x')^2$ where $y' = \dfrac{y - \bar{y}}{S_y}$ and $x' = \dfrac{x - \bar{x}}{S_x}$ $y' = \beta_0^* + \beta_1^* x' + \beta_{11}^*(x')^2$

12-51 a)

```
Predictor        Coef      SE Coef          T        P
Constant       -1.769        1.287      -1.37    0.188
x1             0.4208       0.2942       1.43    0.172
x2             0.2225       0.1307       1.70    0.108
x3           -0.12800      0.07025      -1.82    0.087
x1x2         -0.01988      0.01204      -1.65    0.118
x1x3         0.009151     0.007621       1.20    0.247
x2x3         0.002576     0.007039       0.37    0.719
x1^2         -0.01932      0.01680      -1.15    0.267
x2^2         -0.00745      0.01205      -0.62    0.545
x3^3         0.000824     0.001441       0.57    0.575

S = 0.06092     R-Sq = 91.7%     R-Sq(adj) = 87.0%

Analysis of Variance

Source            DF          SS          MS        F        P
Regression         9    0.655671    0.072852    19.63    0.000
Residual Error    16    0.059386    0.003712
Total             25    0.715057
```

$$\hat{y} = -1.769 + 0.421x_1 + 0.222x_2 - 0.128x_3 - 0.02x_1x_2 + 0.009x_1x_3 + 0.003x_2x_3 - 0.019x_1^2 - 0.007x_2^2 + 0.001x_3^2$$

b) H_0 : all $\beta_1 = \beta_2 = \beta_3 = \text{K} = \beta_{33} = 0$

H_1: at least 1 $\beta_j \neq 0$

$f_0 = 19.628$

$f_{.05,9,16} = 2.54$

$f_0 > f_{\alpha,9,16}$

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 $\beta_{ij} \neq 0$

$$f_0 = \frac{SS_R(\beta_{11},\beta_{22},\beta_{33},\beta_{12},\beta_{13},\beta_{23} \mid \beta_1,\beta_2,\beta_3,\beta_0)/r}{MS_E} = \frac{\frac{0.0359}{6}}{0.003712} = 1.612$$

$$f_{.05,6,16} = 2.74$$

$$f_0 \not> f_{.05,6,16}$$

Do not reject H_0

$$SS_R(\beta_{12},\beta_{13},\beta_{23},\beta_{11},\beta_{22},\beta_{33} \mid \beta_1,\beta_2,\beta_3,\beta_0) = SS_R(\beta_1,\beta_2,\beta_3,\beta_{12},\beta_{13},\beta_{23},\beta_{11},\beta_{22},\beta_{33} \mid \beta_0) -$$

$$SS_R(\beta_1\beta_2\beta_3 \mid \beta_0)$$

$$= 0.65567068 - 0.619763$$

$$= 0.0359$$

Reduced Model: $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$

12-55. a) The min. MS_E equation is $x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8$

$MS_E = 6.58$ $c_p = 5.88$

The 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) $\hat{y} = 0.341 + 2.862x_5 + 0.246x_8 - 0.010x_{10}$

$MS_E = 7.97$ $C_p = 5.02$

e) Minimum C_p and backward elimination result in the same model. Stepwise and forward selection result in the same model. Because it is much smaller, the minimum C_p model seems preferable.

12-61. a) Min. C_p

$\hat{y} = -3.517 + 0.486x_1 - 0.156x_9$

$C_p = -1.67$

b) Min MS_E model is x_1, x_7, x_9, $MS_E = 1.67$, $C_p = -0.77$

$y = -5\ 964 + 0\ 495x_1 + 0\ 025x_7 - 0\ 163x_9$

c) Max. adjusted R^2 model is x_1, x_7, x_9, Adj. $R^2 = 0.98448$ Yes, same as Min. MS_E model.

Supplemental Exercises

12-65. 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) $\alpha = 0.01$ $t_{.005,36} = 2.72$

$H_0{:}\beta_3^* = 0$	$H_0{:}\beta_4 = 0$	$H_0{:}\beta_5 = 0$
$H_1{:}\beta_3^* \neq 0$	$H_1{:}\beta_4 \neq 0$	$H_1{:}\beta_5 \neq 0$
$t_0 = -1.32$	$t_0 = 19.97$	$t_0 = 2.48$
$\lvert t_0 \rvert \ngtr t_{\alpha/2,36}$	$\lvert t_0 \rvert > t_{\alpha/2,36}$	$\lvert t_0 \rvert \ngtr t_{\alpha/2,36}$
Do not reject H_0	Reject H_0	Do not reject H_0

Only regressor x_4 is significant

c) Curvature is evident in the residuals vs. regressor plots from this model.

12-67. 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 pleasing than those in Exercise 12-66.

12-69. 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 parts b.&c. = 714.20 $C_p = 5.57$

12-71. 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 12-65: $\hat{y} = 19.69 - 1.27x_3^* + 0.005x_4 + 0.0004x_5$

12-73. a) $R^2 = \dfrac{SS_R}{SS_T}$

$$SS_R = R^2(SS_T) = 0.94(0.50) = 0.47$$
$$SS_E = SS_T - SS_R = 0.5 - 0.47 = 0.03$$
$$H_0 : \beta_1 = \beta_2 = ... = \beta_6 = 0$$
$H_1 : \beta_j \neq 0$ 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) $k = 5$ $n = 14$ $p = 6$ $R^2 = 0.92$

$$SS_R' = R^2(SS_T) = 0.92(0.50) = 0.46$$
$$SS_E' = SS_T - SS_R' = 0.5 - 0.46 = 0.04$$
$$SS_R(\beta_j, \beta_{i,i=1,2,\Lambda .6,i\neq j} \mid \beta_0) = SS_R(full) - SS_R(reduced)$$
$$= 0.47 - 0.46$$
$$= 0.01$$
$$f_0 = \frac{SS_R(\beta_j \mid \beta_{i,i=1,2,\Lambda ,6,i\neq j} \mid \beta_0)/r}{SS_E'/(n-p)} = \frac{0.01/1}{0.04/8} = 2$$
$$f_{.05,1,8} = 5.32$$
$$f_0 \not> f_{\alpha,1,8}$$

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) $MS_E(reduced) = \dfrac{SS_E}{n-p} = \dfrac{0.04}{8} = 0.005$

$$MS_E(full) = \frac{0.03}{7} = 0.00429$$

No, the MS_E is larger for the reduced model, although not by much. Generally, if adding a variable to a model reduces the MS_E it is an indication that the variable may be useful in explaining the response variable. Here the decrease in MS_E is not very great because the added variable had no real explanatory power.

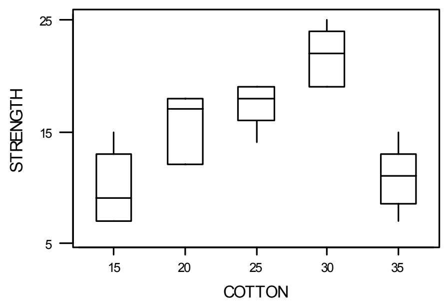
Chapter 13 Selected Problem Solutions

Section 13-2

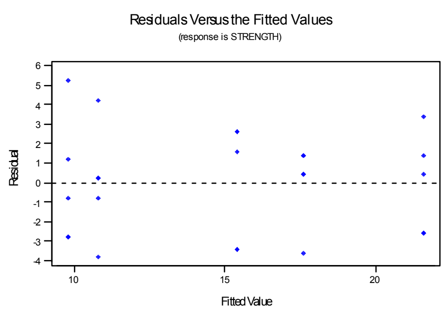
13-1. a) Analysis of Variance for STRENGTH

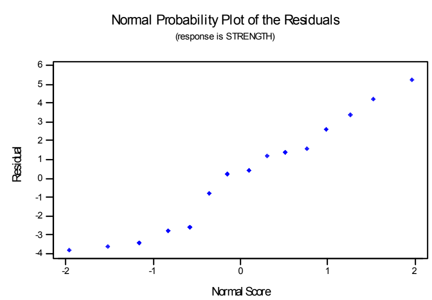
Source	DF	SS	MS	F	P
COTTON	4	475.76	118.94	14.76	0.000
Error	20	161.20	8.06		
Total	24	636.96			

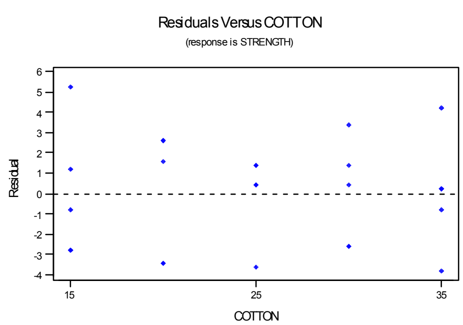
Reject H_0 and conclude that cotton percentage affects mean breaking strength.
b) Tensile strength seems to increase to 30% cotton and declines at 35% cotton.



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







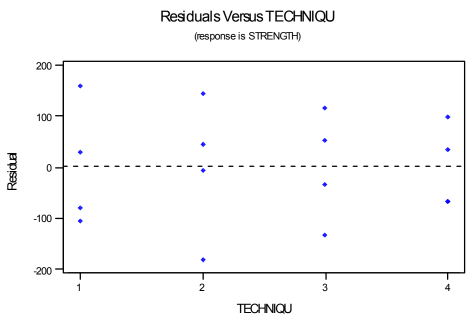
13-3. a) Analysis of Variance for STRENGTH

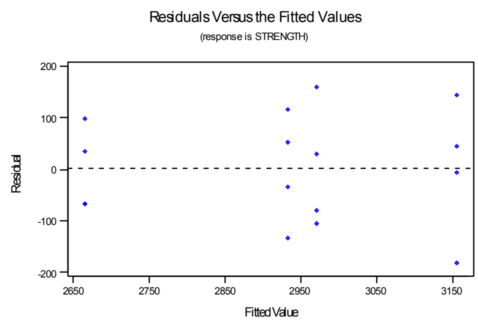
Source	DF	SS	MS	F	P
TECHNIQU	3	489740	163247	12.73	0.000
Error	12	153908	12826		
Total	15	643648			

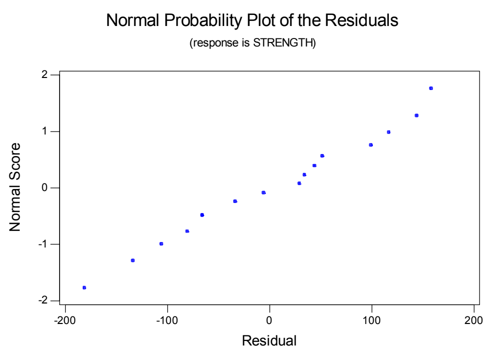
Reject H_0. Techniques affect the mean strength of the concrete.

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

c) Residuals are acceptable





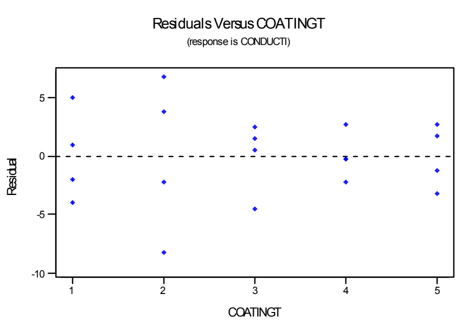


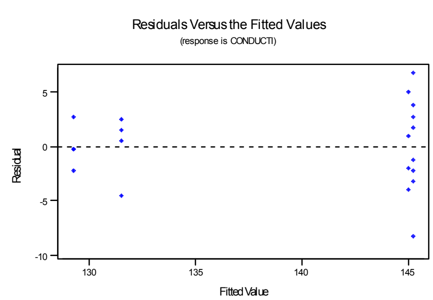
13-5. a) Analysis of Variance for CONDUCTIVITY

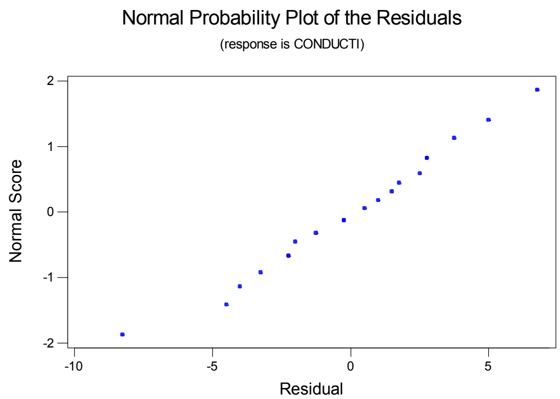
Source	DF	SS	MS	F	P
COATINGTYPE	4	1060.5	265.1	16.35	0.000
Error	15	243.3	16.2		
Total	19	1303.8			

Reject H_0 ; P-value $\cong 0$.

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







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

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

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

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

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

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

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

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

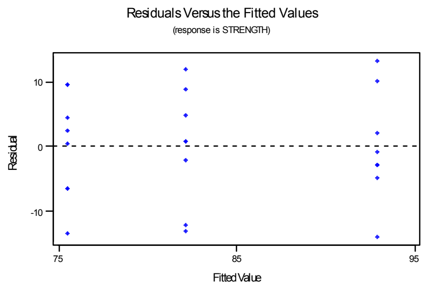
13-9. a) Analysis of Variance for STRENGTH

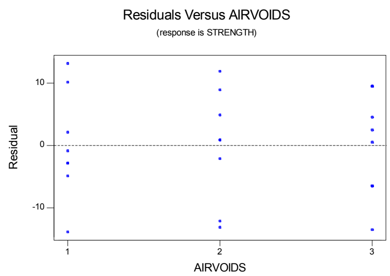
Source	DF	SS	MS	F	P
AIRVOIDS	2	1230.3	615.1	8.30	0.002
Error	21	1555.8	74.1		
Total	23	2786.0			

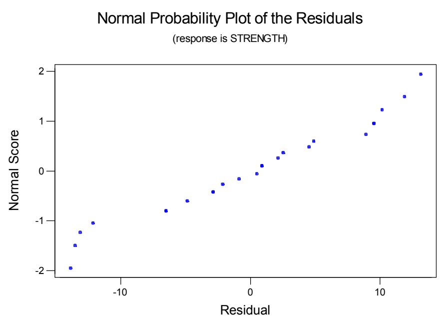
Reject H_0

b) *P-value* = 0.002

c) The residual plots show that the assumptions of equality of variance is reasonable. The normal probability plot has some curvature in the tails.







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

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

$$8.229 - 2.080\sqrt{\frac{74.1}{8}} \le \mu_3 \le 8.229 + 2.080\sqrt{\frac{74.1}{8}}$$

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

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

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

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

$$8.42 \le \mu_1 - \mu_4 \le 26.38$$

Section 13-3

13-21 a) Analysis of Variance for OUTPUT

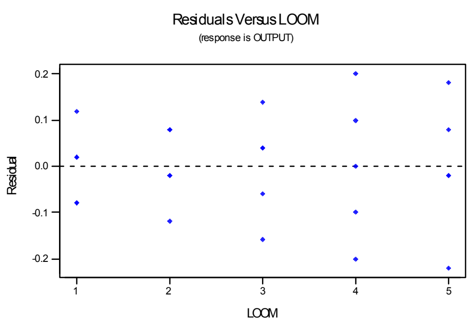
Source	DF	SS	MS	F	P
LOOM	4	0.3416	0.0854	5.77	0.003
Error	20	0.2960	0.0148		
Total	24	0.6376			

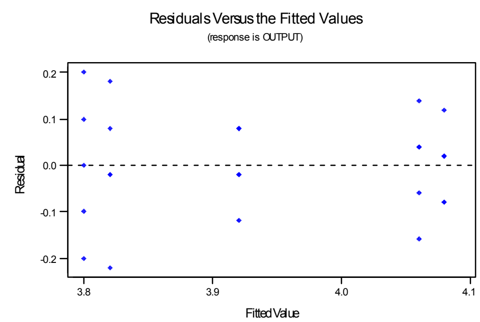
Reject H_0, and conclude that there are significant differences among the looms.

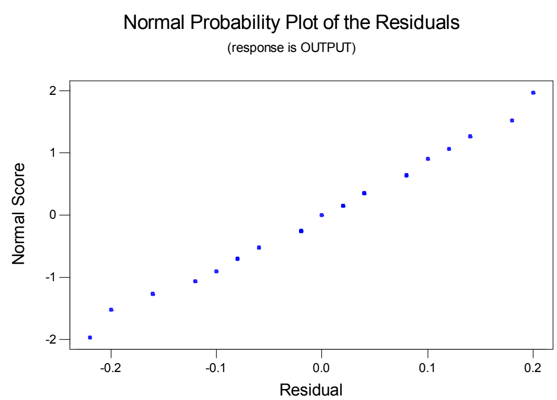
b) $\hat{\sigma}_{\tau}^{2} = \dfrac{MS_{Treatments} - MS_E}{n} = \dfrac{0.0854 - 0.0148}{5} = 0.01412$

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

d) Residuals plots are acceptable





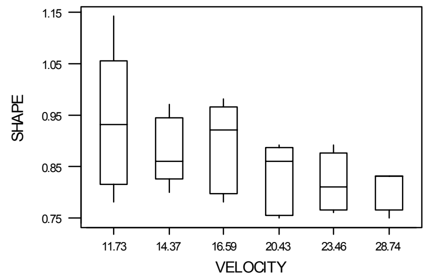


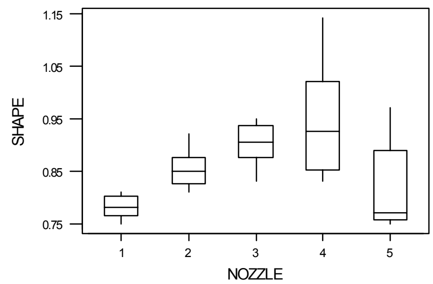
Section 13-4

13-25. a) Analysis of Variance for SHAPE

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

Reject H_0, and conclude that nozzle type affects the mean shape measurement.



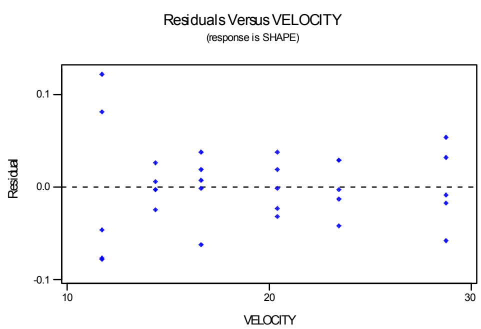


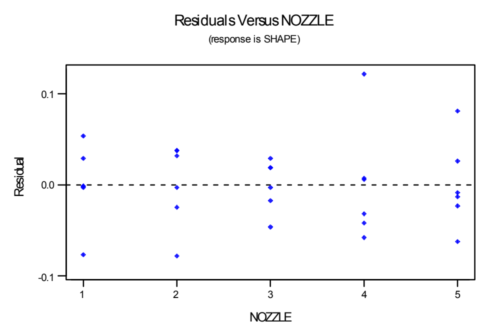
b) Fisher's pairwise comparisons

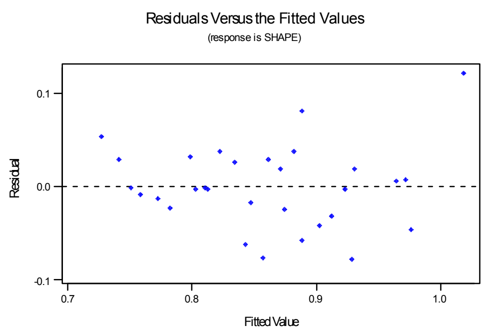
```
   Family error rate = 0.268
   Individual error rate = 0.0500
   Critical value = 2.060
   Intervals for (column level mean) - (row level mean)
                1            2            3            4
       2   -0.15412
             0.01079
       3   -0.20246    -0.13079
            -0.03754     0.03412
       4   -0.24412    -0.17246    -0.12412
            -0.07921    -0.00754     0.04079
       5   -0.11412    -0.04246     0.00588     0.04754
             0.05079     0.12246     0.17079     0.21246
```

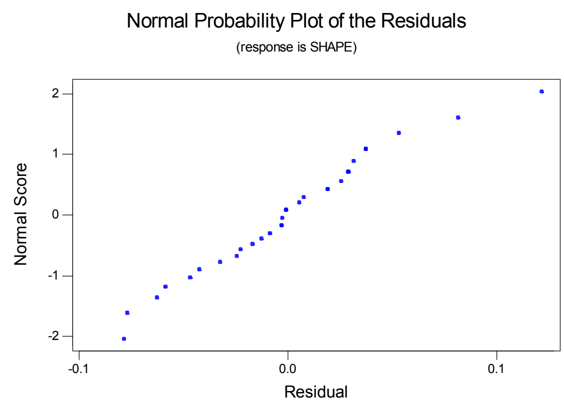
There are significant differences between nozzle types 1 and 3, 1 and 4, 2 and 4, 3 and 5, and 4 and 5.

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









Supplemental Exercises

13-31. a)
```
Analysis of Variance for RESISTANCE
Source      DF        SS        MS        F        P
ALLOY        2   10941.8    5470.9    76.09    0.000
Error       27    1941.4      71.9
Total       29   12883.2
```

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

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

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

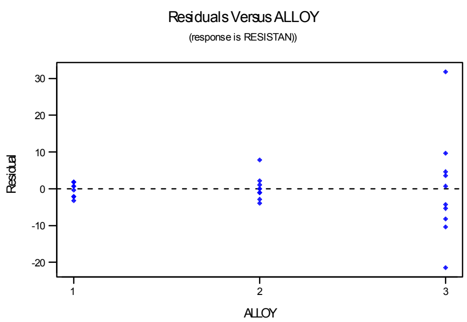
c) 99% confidence interval on the mean contact resistance for alloy 3

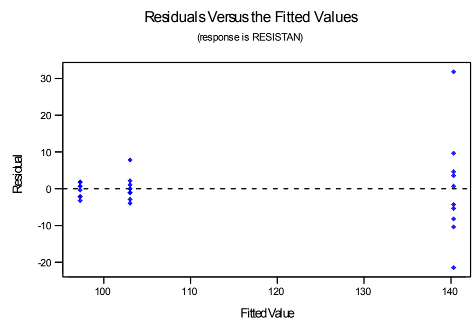
$$\bar{y}_3 - t_{0.005,271}\sqrt{\frac{MS_E}{n}} \le \mu_i \le \bar{y}_3 + t_{0.005,27}\sqrt{\frac{MS_E}{n}}$$

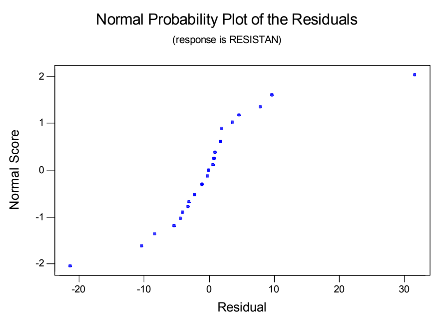
$$140.4 - 2.771\sqrt{\frac{71.9}{10}} \le \mu_3 \le 140.4 - 2.771\sqrt{\frac{71.9}{10}}$$

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

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







13-35. a)Analysis of Variance for VOLUME

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

Reject H_0.

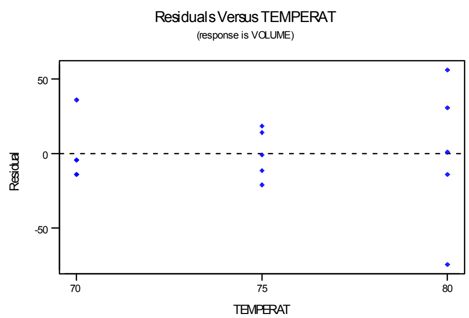
b) *P-value* = 0.007

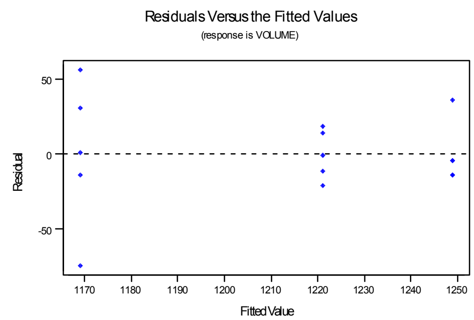
c) Fisher's pairwise comparisons

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

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

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



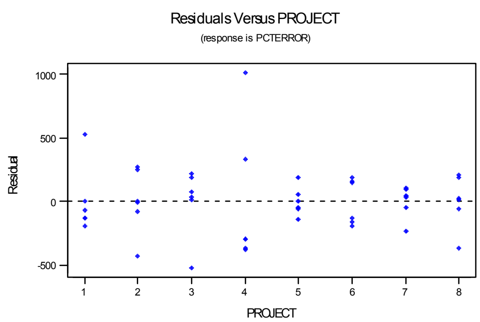


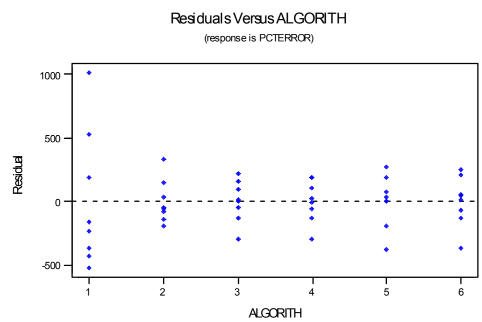
13-37. a) Analysis of Variance for PCTERROR

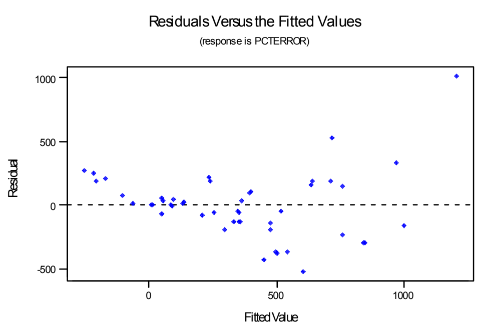
Source	DF	SS	MS	F	P
ALGORITH	5	2825746	565149	6.23	0.000
PROJECT	7	2710323	387189	4.27	0.002
Error	35	3175290	90723		
Total	47	8711358			

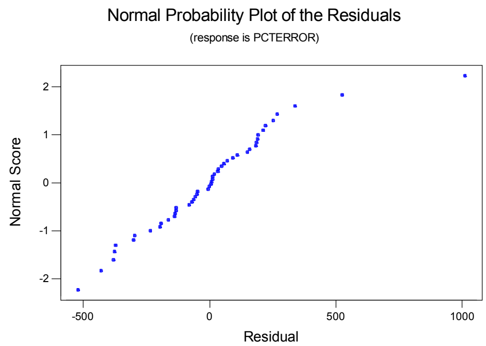
Reject H_0 , the algorithm is significant.

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









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

13-39 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 = $1 - \beta$
5	3.317	20	0.10	0.90

The sample size should be approximately n = 5

Chapter 14 Selected Problem Solutions

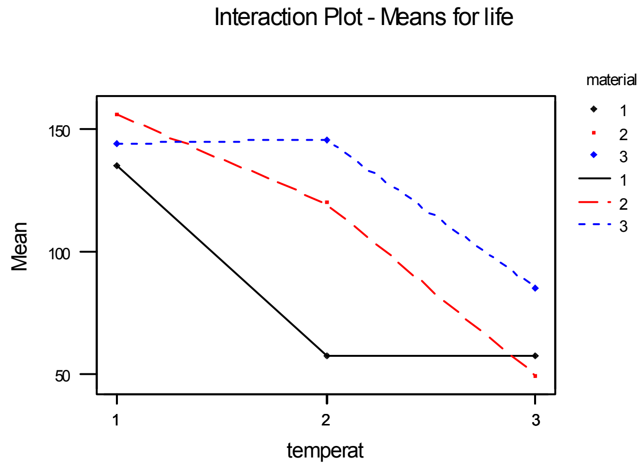
Section 14-3

14-1. a) Analysis of Variance for life

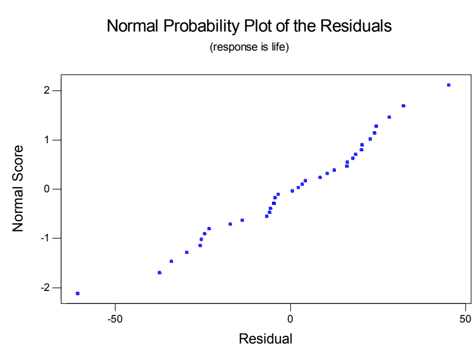
Source	DF	SS	MS	F	P
material	2	10683.7	5341.9	7.91	0.002
temperat	2	39118.7	19559.4	28.97	0.000
material*temperat	4	9613.8	2403.4	3.56	0.019
Error	27	18230.7	675.2		
Total	35	77647.0			

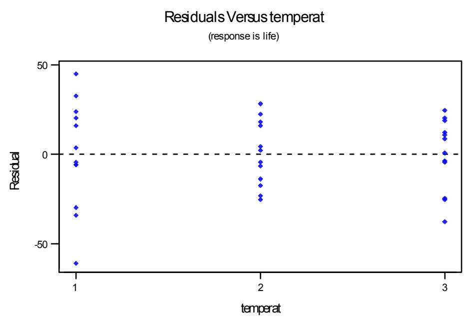
Main factors and interaction are all significant.

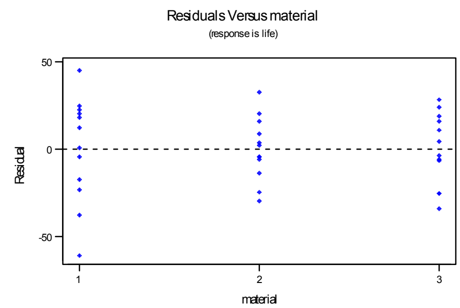
b)The mean life for material 2 is the highest at temperature level 1, in the middle at temperature level 2 and the lowest at temperature level 3. This shows that there is an interaction.



c) There appears to be slightly more variability at temperature 1 and material 1. The normal probability plot shows that the assumption of normality is reasonable.







14-3 a) $H_0 : \tau_1 = \tau_2 = 0$

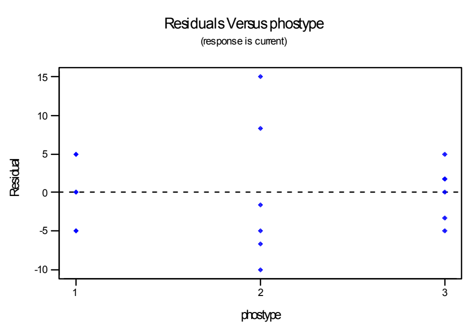
H_1 : at least one $\tau_i \neq 0$

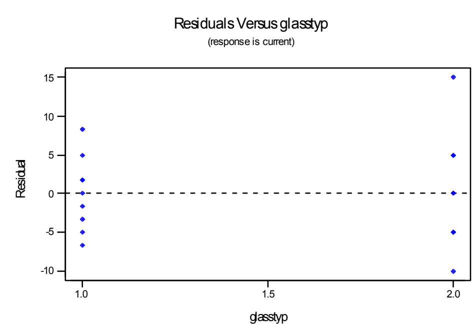
b) Analysis of Variance for current

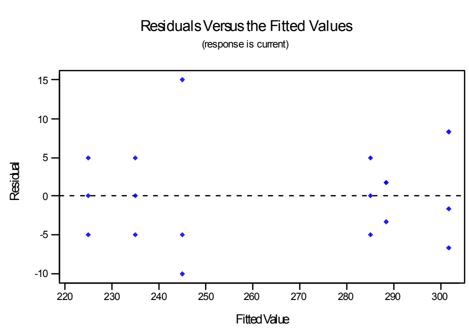
Source	DF	SS	MS	F	P
glasstyp	1	14450.0	14450.0	273.79	0.000
phostype	2	933.3	466.7	8.84	0.004
glasstyp*phostype	2	133.3	66.7	1.26	0.318
Error	12	633.3	52.8		
Total	17	16150.0			

Main effects are significant, the interaction is not significant. Glass type 1 and phosphor type 2 lead to the high mean current (brightness).

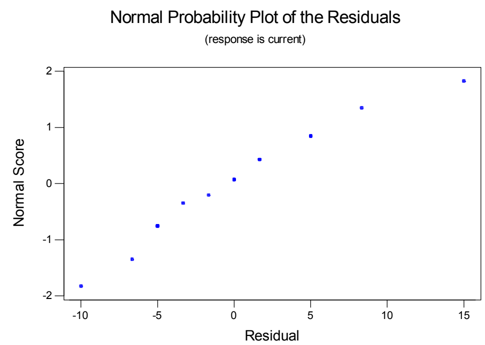
c) There appears to be slightly more variability at phosphor type 2 and glass type 2. The normal plot of the residuals shows that the assumption of normality is reasonable.







14-7 The ratio



$$T = \frac{\bar{y}_{\cdot i\cdot} - \bar{y}_{\cdot j\cdot} - (\mu_i - \mu_j)}{\sqrt{2MS_E / n}}$$ has a t distribution with $ab(n\text{-}1)$ degrees of freedom

Therefore, the (1-α)% confidence interval on the difference in two treatment means is

$$\bar{y}_{\cdot i\cdot} - \bar{y}_{\cdot j\cdot} - t_{a/2,ab(n-1)}\sqrt{\frac{2MS_E}{n}} \le \mu_i - \mu_j \le \bar{y}_{\cdot i\cdot} - \bar{y}_{\cdot j\cdot} + t_{a/2,ab(n-1)}\sqrt{\frac{2MS_E}{n}}$$

For exercise 14-6, the mean warping at 80% copper concentration is 21.0 and the mean warping at 60% copper concentration is 18.88 a=4, b=4, n=2 and MS_E=6.78. The degrees of freedom are (4)(4)(1)=16

$$(21.0 - 18.88) - 2.120\sqrt{\frac{2 * 6.78}{2}} \le \mu_3 - \mu_2 \le (21.0 - 18.88) + 2.120\sqrt{\frac{2 * 6.78}{2}}$$
$$-3.40 \le \mu_3 - \mu_2 \le 7.64$$

Therefore, there is no significant differences between the mean warping values at 80% and 60% copper concentration.

<u>Section 14-4</u>

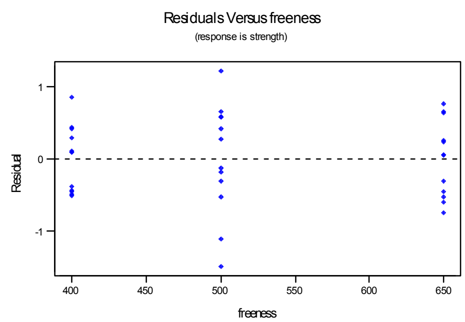
14-11 Parts a. and b.

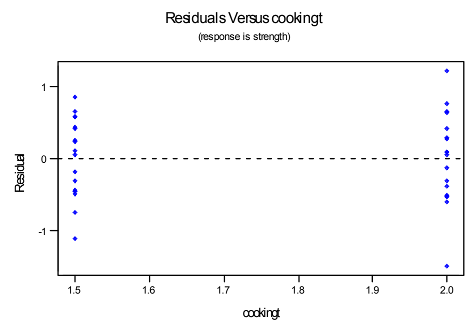
Analysis of Variance for strength

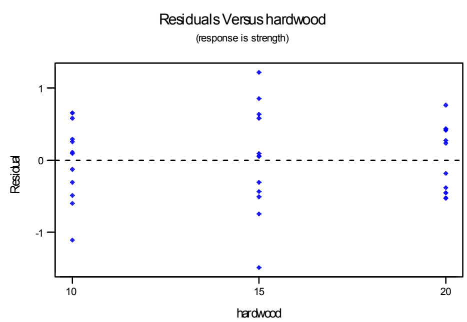
Source	DF	SS	MS	F	P
hardwood	2	8.3750	4.1875	7.64	0.003
cookingtime	1	17.3611	17.3611	31.66	0.000
freeness	2	21.8517	10.9258	19.92	0.000
hardwood*cookingtime	2	3.2039	1.6019	2.92	0.075
hardwood*freeness	4	6.5133	1.6283	2.97	0.042
cookingtime*freeness	2	1.0506	0.5253	0.96	0.399
Error	22	12.0644	0.5484		
Total	35	70.4200			

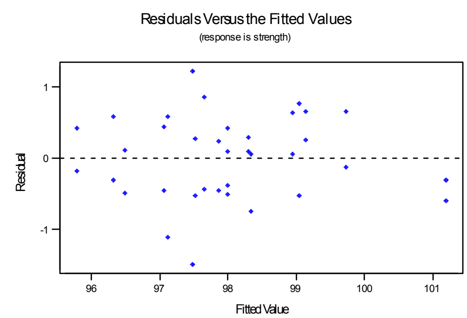
All main factors are significant. The interaction of hardwood * freeness is also significant.

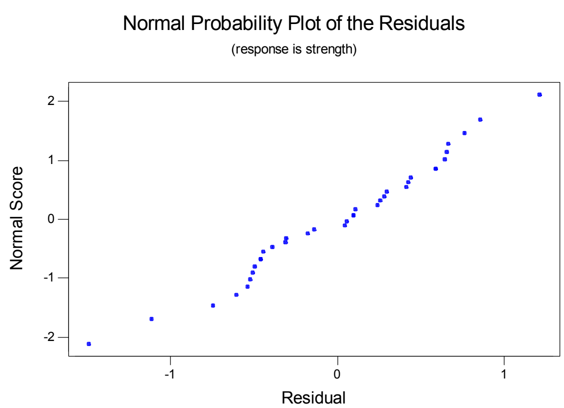
c) The residual plots show no serious problems with normality or equality of variance











Section 14-5

14-13 a) Analysis of Variance for life (coded units)

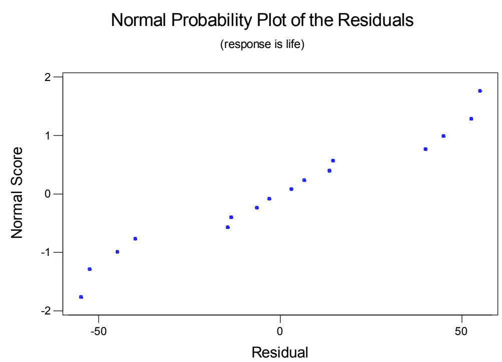
Source	DF	SS	MS	F	P
speed	1	1332	1332	0.49	0.502
hardness	1	28392	28392	10.42	0.010
angle	1	20592	20592	7.56	0.023
speed*hardness	1	506	506	0.19	0.677
speed*angle	1	56882	56882	20.87	0.000
hardness*angle	1	2352	2352	0.86	0.377
Error	9	24530	2726		
Total	15	134588			

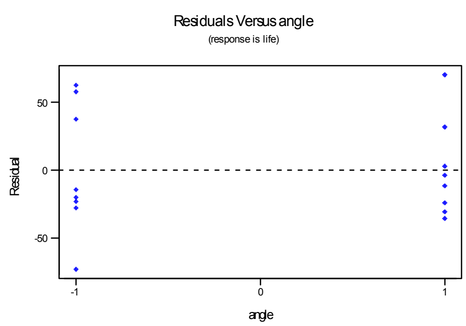
b) Estimated Effects and Coefficients for life (coded units)

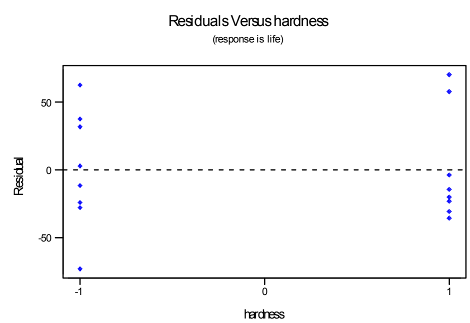
Term	Effect	Coef	SE Coef	T	P
Constant		413.13	12.41	33.30	0.000
speed	18.25	9.12	12.41	0.74	0.483
hardness	84.25	42.12	12.41	3.40	0.009
angle	71.75	35.87	12.41	2.89	0.020
speed*hardness	-11.25	-5.63	12.41	-0.45	0.662
speed*angle	-119.25	-59.62	12.41	-4.81	0.001
hardness*angle	-24.25	-12.12	12.41	-0.98	0.357
speed*hardness*angle	-34.75	-17.37	12.41	-1.40	0.199

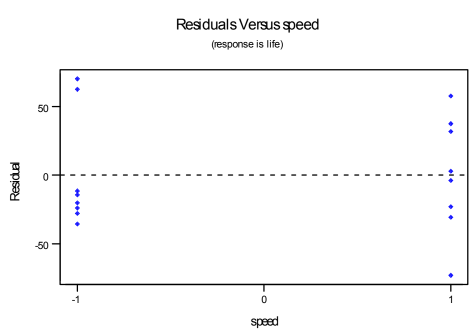
$$\hat{y} = 413.125 + 9.125x_1 + 45.12x_2 + 35.87x_3 - 59.62x_{13}$$

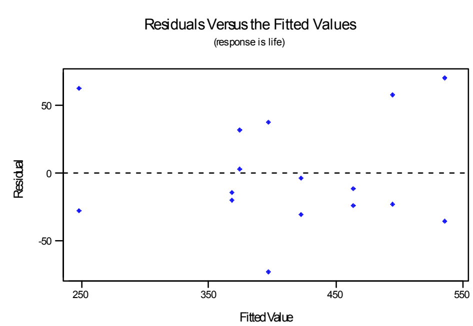
c) Analysis of the residuals shows that all assumptions are reasonable.



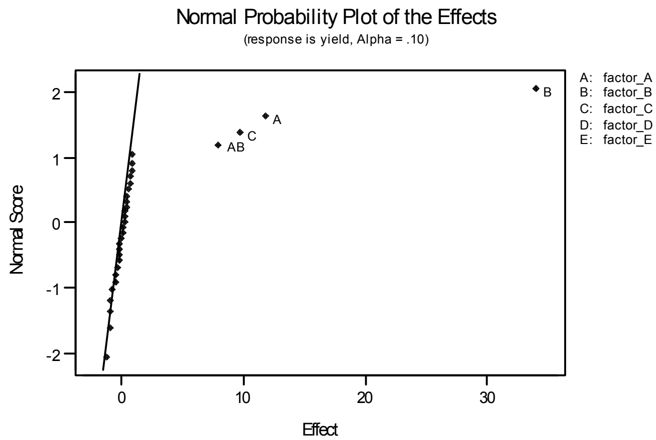








14-19. a) Factors A, B, C, and the interaction AB appear to be significant from the normal probability plot of the effects.



b)

Analysis of Variance for yield (coded units)

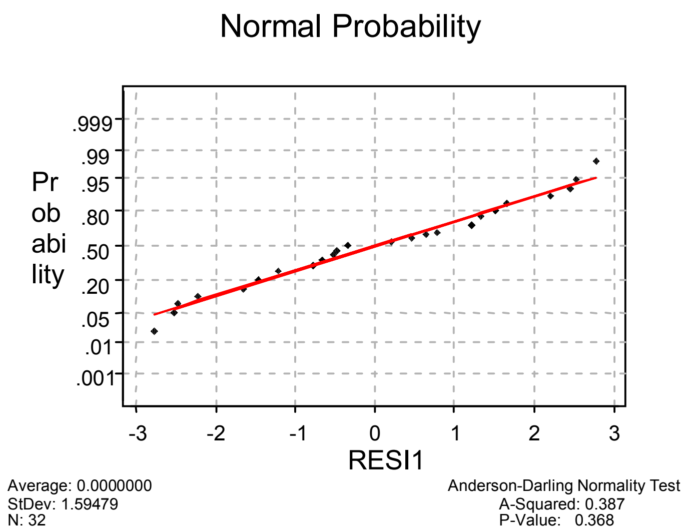
Term	Effect	Coef	StDev Coef	T	P
Constant		30.5312	0.2786	109.57	0.000
factor_A	11.8125	5.9063	0.2786	21.20	0.000
factor_B	9.6875	4.8437	0.2786	17.38	0.000
factor_D	-0.8125	-0.4063	0.2786	-1.46	0.164
factor_E	0.4375	0.2187	0.2786	0.79	0.444
factor_A*factor_B	7.9375	3.9687	0.2786	14.24	0.000
factor_A*factor_C	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

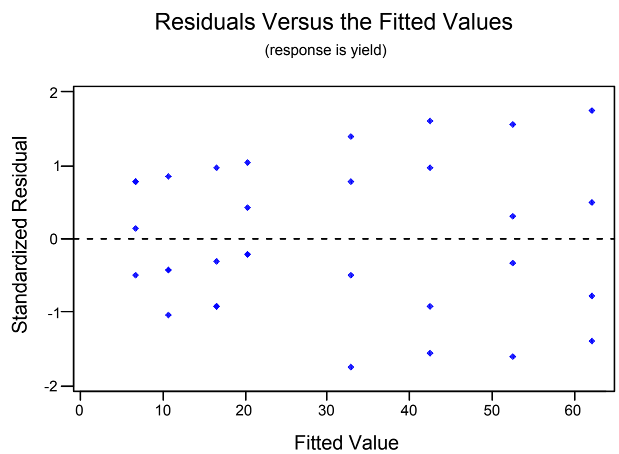
Analysis of Variance for yield

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)

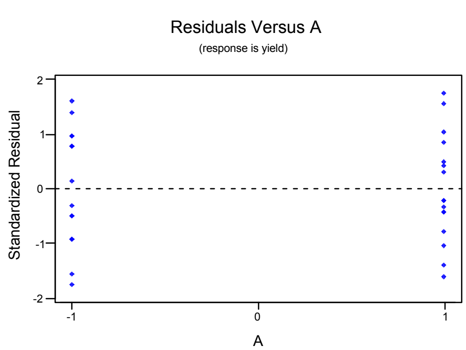
c) The normal probability plot of the residuals is satisfactory. However their variance appears to increase as the fitted value increases.

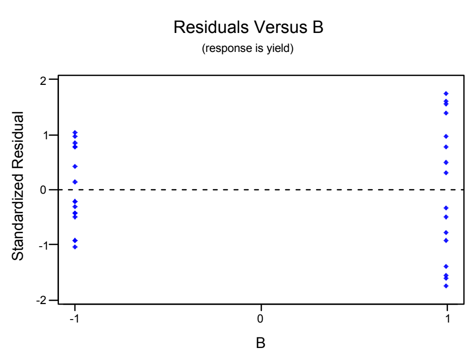


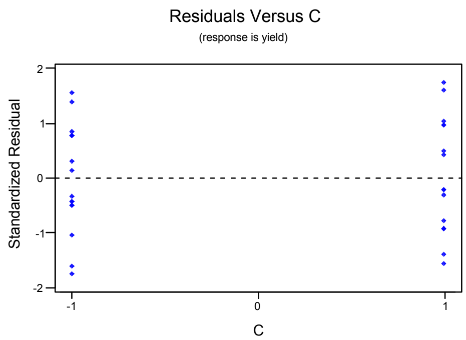


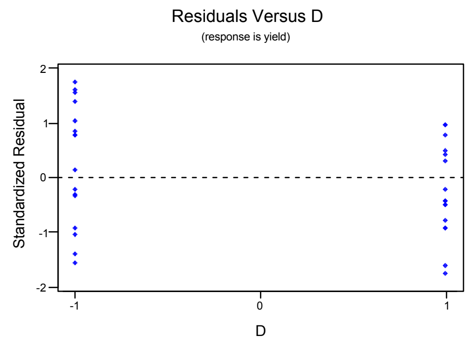
.

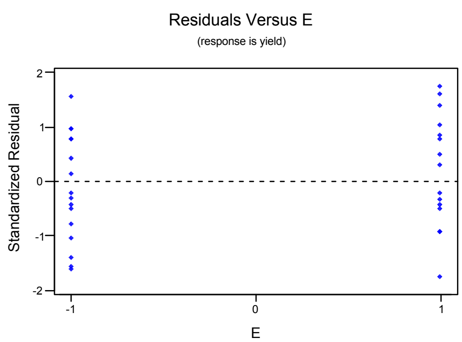
d) All plots support the constant variance assumption , although there is a very slight indication that variability is greater at the high level of factor B.



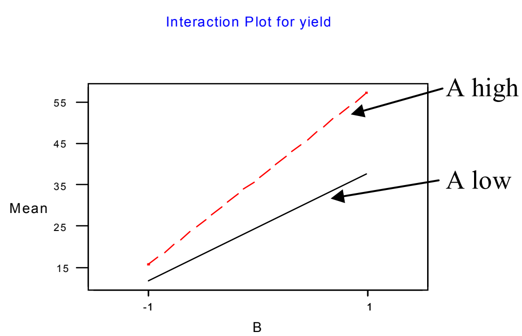






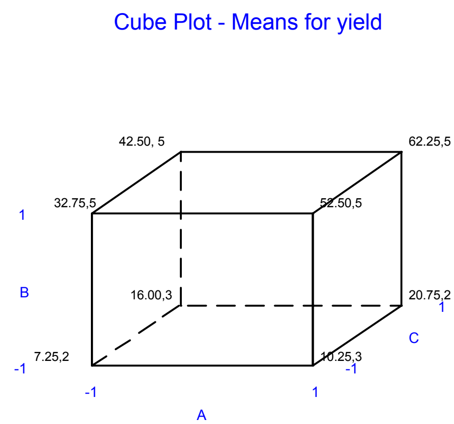


e) The AB interaction appears to be significant. The interaction plot from MINITAB indicates that a high level of A and of B increases the mean yield, while low levels of both factors would lead to a reduction in the mean yield.

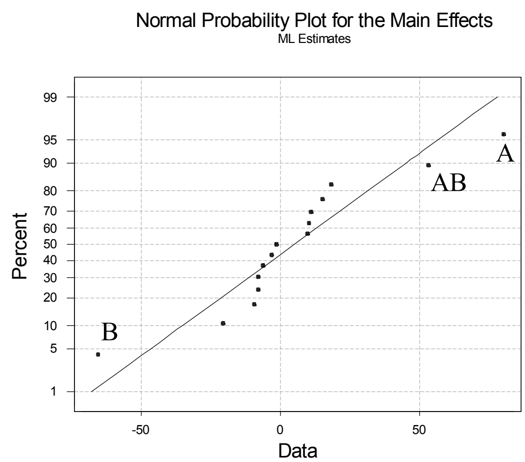


f.) To increase yield and therefor optimize the process, we would want to set A, B, and C at their high levels.

g) It is evident from the cube plot that we should run the process with all factors set at their high level.



14-21



b) Based on the normal probability plot of the effects, factors A, B and AB are significant. The model would include these three factors.

c) The estimated model is: $\hat{y} = 400 + 40.124x_1 - 32.75x_2 + 26.625x_{12}$

Section 14-6

14-25 Model with four blocks

BLOCK	A	B	C	D	var_1
1	-1	-1	-1	-1	190
1	1	-1	1	-1	181
1	-1	1	-1	1	187
1	1	1	1	1	180
2	1	-1	-1	-1	174
2	-1	-1	1	-1	177
2	1	1	-1	1	185
2	-1	1	1	1	187
3	-1	1	-1	-1	181
3	1	1	1	-1	173
3	-1	-1	-1	1	198
3	1	-1	1	1	179
4	1	1	-1	-1	183
4	-1	1	1	-1	188
4	1	-1	-1	1	172
4	-1	-1	1	1	199

```
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
```

Factor A is the only significant factor.

14-29 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
```

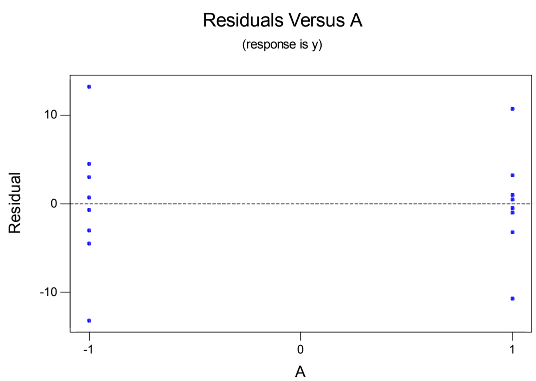
3		-10.88	4.560	-2.38	0.054
A	-45.25	-22.62	2.633	-8.59	0.000
B	-1.50	-0.75	2.633	-0.28	0.785
C	14.50	7.25	2.633	2.75	0.033
A*B	19.00	9.50	2.633	3.61	0.011
A*C	-14.50	-7.25	2.633	-2.75	0.033
B*C	-9.25	-4.63	2.633	-1.76	0.130

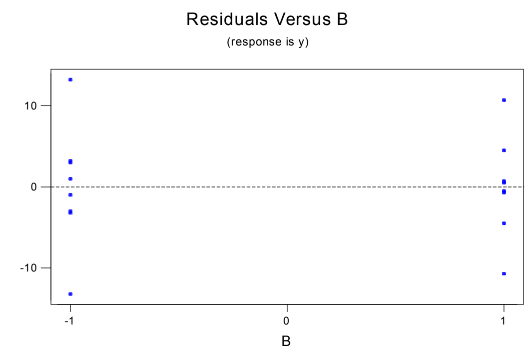
Analysis of Variance for y

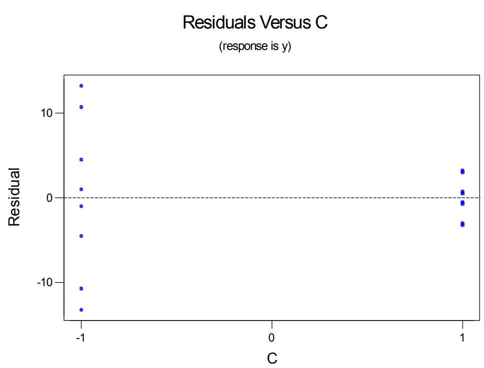
Source	DF	Seq SS	Adj SS	Adj MS	F	P
Blocks	3	1502.8	1502.8	500.9	4.52	0.055
Main Effects	3	9040.2	9040.2	3013.4	27.17	0.001
2-Way Interactions	3	2627.2	2627.2	875.7	7.90	0.017
Residual Error	6	665.5	665.5	110.9		
Total	15	13835.7				

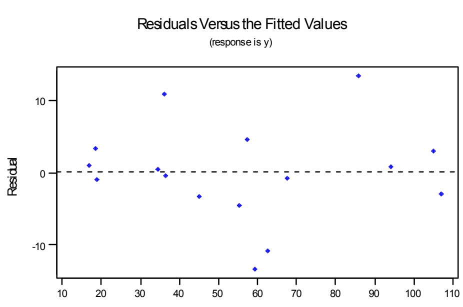
Factors A, C, AB, and AC are significant.

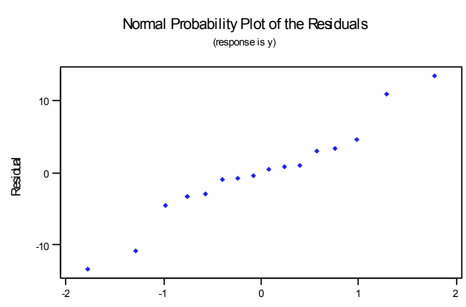
b) Analysis of the residuals shows that the model is adequate. There is more variability on the response associated with the low setting of factor C, but that is the only problem.









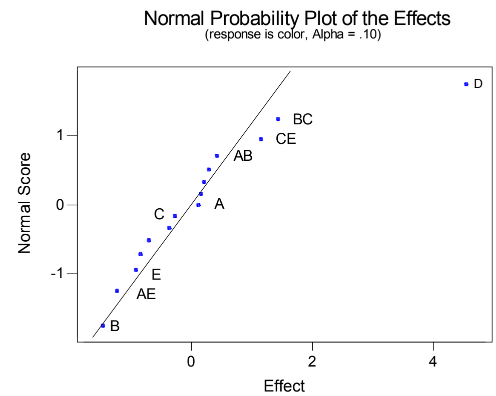


c.) Some of the information from the experiment is lost because the design is run in 4 blocks. This causes us to lose information on the ABC interaction even though we have replicated the experiment twice. If it is possible to run the experiment in only 2 blocks, there would be information on all interactions.

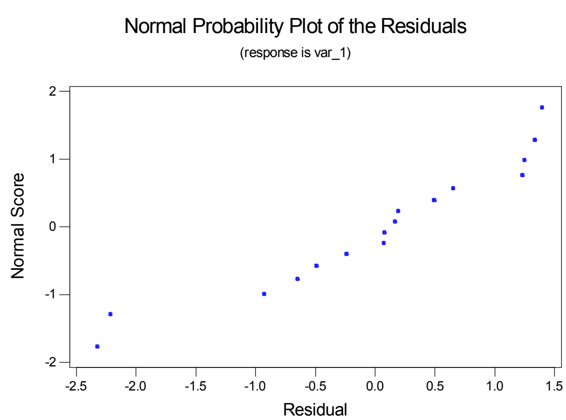
d) To have data on all interactions, we could run the experiment so that each replicate is a block (therefore only 2 blocks).

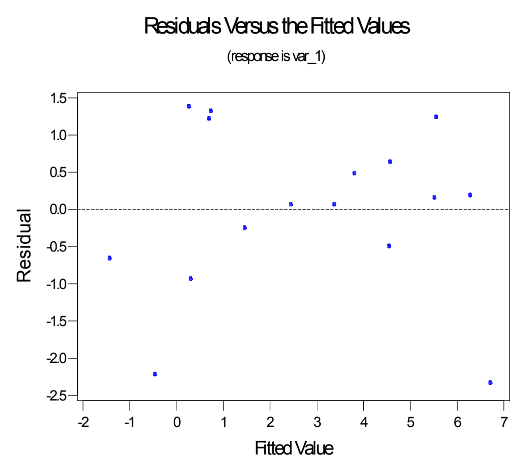
Section 14-7

14-31 a) Factors A, B and D are active factors.



b) There are no serious problems with the residual plots. The normal probability plot has a little bit of curvature at the low end and there is a little more variability at the lower and higher ends of the fitted values.



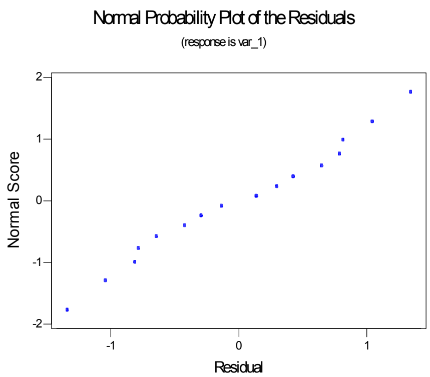


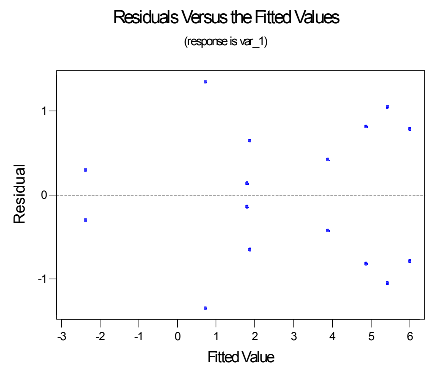
c) Part a. indicates that only A,B, and D are important. In these factors only, the design is a 2^3 with two replicates.

```
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
```

Factors A, B, D, AB and AD are significant.



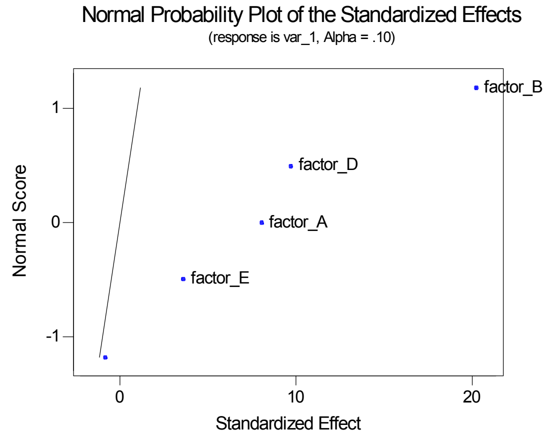


The normal probability plot and the plot of the residuals versus fitted values are satisfactory.

14-35 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. This is different than the design that results when C and E are dropped from the 2^{6-2} in Table 14-28 which results in a full factorial because, the factors ABDF do not form a word in the complete defining relation

14-37 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




```
Estimated Effects and Coefficients for var_1 (coded units)

Term          Effect      Coef     SE Coef       T       P
Constant                3025.00      90.14   33.56   0.001
factor_A     1450.00     725.00      90.14    8.04   0.015
factor_B     3650.00    1825.00      90.14   20.25   0.002
factor_C     -150.00     -75.00      90.14   -0.83   0.493
factor_D     1750.00     875.00      90.14    9.71   0.010
factor_E      650.00     325.00      90.14    3.61   0.069

Analysis of Variance for var_1 (coded units)

Source                DF      Seq SS     Adj SS      Adj MS       F       P
Main Effects           5    37865000   37865000     7573000  116.51   0.009
Residual Error         2      130000     130000       65000
Total                  7    37995000
```

Factors A, B and D are significant.

Supplemental Exercises

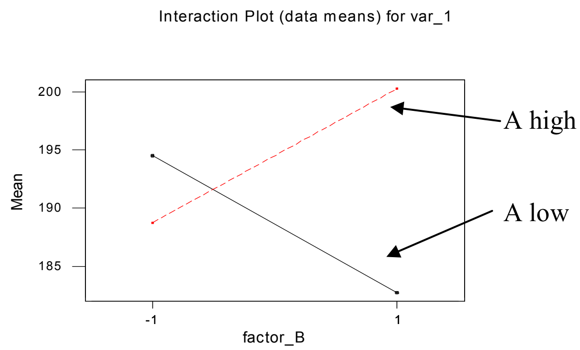
14-41 a

```
Estimated Effects and Coefficients for var_1 (coded units)

Term                  Effect      Coef     SE Coef       T       P
Constant                       191.563       1.158  165.49   0.000
factor_A (PH)          5.875     2.937       1.158    2.54   0.026
factor_B (CC)         -0.125    -0.062       1.158   -0.05   0.958
factor_A*factor_B     11.625     5.812       1.158    5.02   0.000

Analysis of Variance for var_1 (coded units)

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    540.562    540.562     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
```

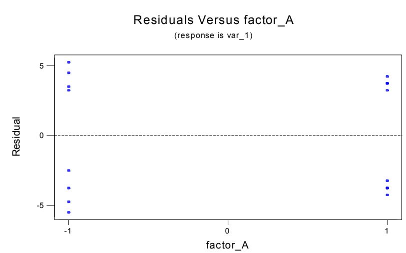
The main effect of pH and the interaction of pH and Catalyst Concentration (CC) are significant at the 0.05 level of significance.

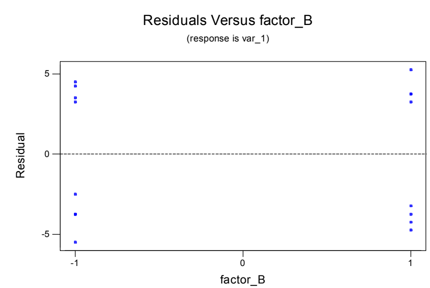
The model used is viscosity = $191.563 + 2.937x_1 - 0.062x_2 + 5.812x_{12}$

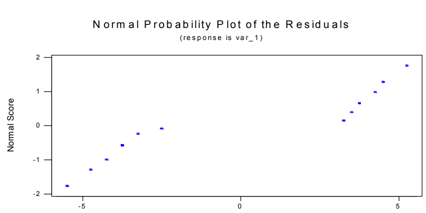
b.) The interaction plot shows that there is a strong interaction. When Factor A is at its low level, the mean response is large at the low level of B and is small at the high level of B. However, when A is at its high level, the results are opposite.

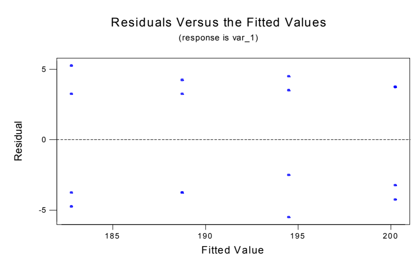


c.) The plots of the residuals show that the equality of variance assumption is reasonable. However, there is a large gap in the middle of the normal probability plot. Sometimes, this can indicate that there is another variable that has an effect on the response but which is not included in the experiment. For example, in this experiment, note that the replicates in each cell have two pairs of values that are very similar, but there is a rather large difference in the mean values of the two pairs. (Cell 1 has 189 and 192 as one pair and 198 and 199 as the other.)







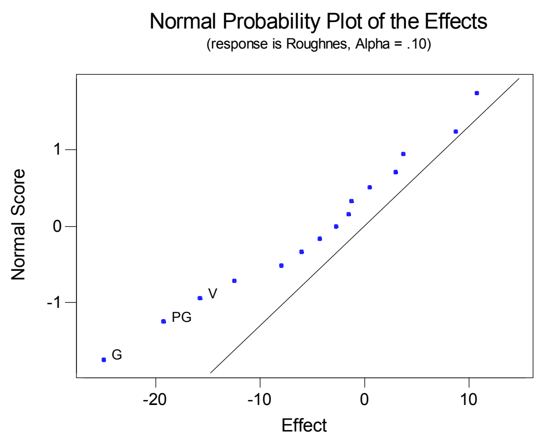


14-47 a)

Term	Effect
V	-15.75

F	8.75
P	10.75
G	-25.00
V*F	3.00
V*P	-8.00
V*G	-2.75
F*P	-6.00
F*G	3.75
P*G	-19.25
V*F*P	-1.25
V*F*G	0.50
V*P*G	-1.50
F*P*G	-12.50
V*F*P*G	-4.25

b)



According to the probability plot, factors V, P, and G and, PG are possibly significant.

Estimated Effects and Coefficients for roughnes (coded units)

Term	Effect	Coef	SE Coef	T	P
Constant		102.75	2.986	34.41	0.000
V	-15.75	-7.87	2.986	-2.64	0.046
F	8.75	4.37	2.986	1.46	0.203
P	10.75	5.37	2.986	1.80	0.132
G	-25.00	-12.50	2.986	-4.19	0.009
V*F	3.00	1.50	2.986	0.50	0.637
V*P	-8.00	-4.00	2.986	-1.34	0.238
V*G	-2.75	-1.38	2.986	-0.46	0.665
F*P	-6.00	-3.00	2.986	-1.00	0.361
F*G	3.75	1.88	2.986	0.63	0.558
P*G	-19.25	-9.63	2.986	-3.22	0.023

Analysis of Variance for roughnes (coded units)

Analysis of Variance for Roughnes (coded units)

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Main Effects	4	4260.7	4260.7	1065.2	7.46	0.024

2-Way Interactions	6	2004.7	2004.7	334.1	2.34	0.184
Residual Error	5	713.5	713.5	142.7		
Total	15	6979.0				

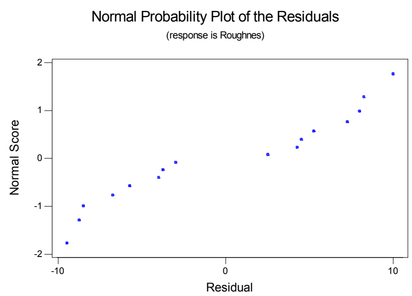
$$\hat{y} = 102.75 - 7.87x_1 + 5.37x_3 - 12.50x_4 - 9.63x_{34}$$

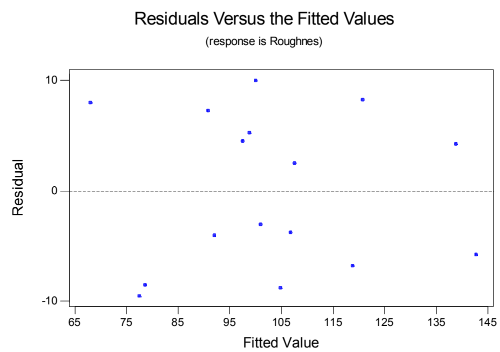
c) From the analysis, we see that water jet pressure (P), abrasive grain size (G), and jet traverse speed (V) are significant along with the interaction of water jet pressure and abrasive grain size

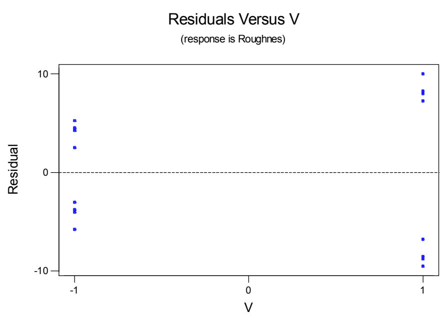
d) The residual plots appear to indicate the assumption of constant variance may not be met. The assumption of normality appears reasonable.

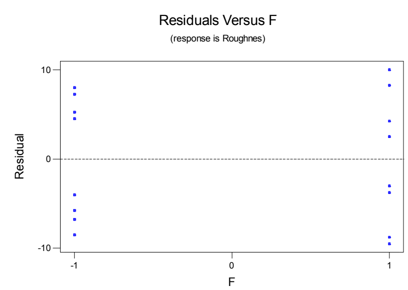
14-49 The design uses G=VPF as the generator.

Alias Structure

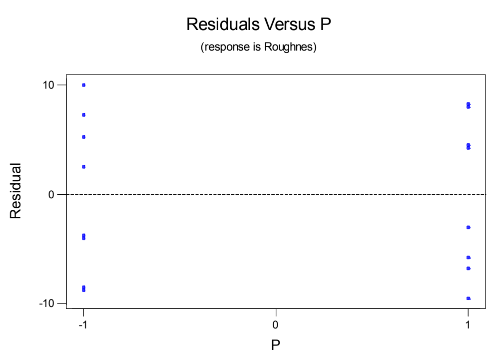


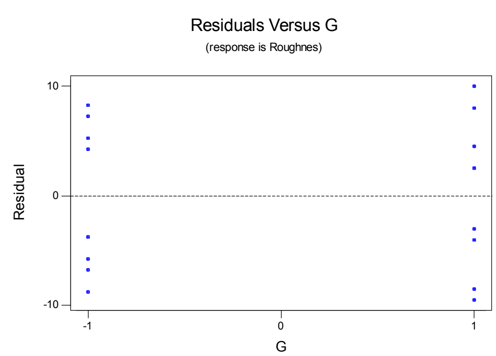






I + V*P*F*G





V + P*F*G
P + V*F*G
F + V*P*G
G + V*P*F

```
V*P + F*G
V*F + P*G
V*G + P*F


Estimated Effects and Coefficients for C9 (coded units)

Term         Effect      Coef     SE Coef        T      P
Constant                102.63      6.365    16.12  0.004
V            -14.75      -7.37      6.365    -1.16  0.366
P            -28.25     -14.12      6.365    -2.22  0.157
F             -1.25      -0.62      6.365    -0.10  0.931
G            -14.75      -7.38      6.365    -1.16  0.366
P*G           17.75       8.88      6.365     1.39  0.298

Analysis of Variance for C9 (coded units)

Source                 DF      Seq SS     Adj SS     Adj MS       F      P
Main Effects            4      2469.5     2469.5      617.4    1.90  0.373
2-Way Interactions      1       630.1      630.1      630.1    1.94  0.298
Residual Error          2       648.3      648.3      324.1
Total                   7      3747.9
```

The results do not show any significant factors. A lot of the information is lost due to the half-fraction of the design.

14-51 Design Generators: D = AB E = AC

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

Design

StdOrder	A	B	C	D	E
1	-1	-1	-1	1	1
2	1	-1	-1	-1	-1
3	-1	1	-1	-1	1
4	1	1	-1	1	-1
5	-1	-1	1	1	-1
6	1	-1	1	-1	1
7	-1	1	1	-1	-1
8	1	1	1	1	1

Chapter 15 Selected Problem Solutions

Section 15-2

15-1. 1. The parameter of interest is median of pH.

2. $H_0 : \tilde{\mu} = 7.0$

3 $H_1 : \tilde{\mu} \neq 7.0$

4. α=0.05

5. The test statistic is the observed number of plus differences or $r^+ = 8$.

6. We reject H_0 if the *P-value* corresponding to $r^+ = 8$ is less than or equal to α=0.05.

7. Using the binomial distribution with n=10 and p=0.5, *P-value* = $2P(R^* \geq 8 | p=0.5) = 0.109$

8. Conclusion: we cannot reject H_0. There is not enough evidence to reject the manufacturer's claim that the median of the pH is 7.0

15-5 **a**. 1. The parameter of interest is the median compressive strength

2. $H_0 : \tilde{\mu} = 2250$

3. $H_1 : \tilde{\mu} > 2250$

4. α=0.05

5. The test statistic is the observed number of plus differences or $r^+ = 7$.

6. We reject H_0 if the *P-value* corresponding to $r^+ = 7$ is less than or equal to α=0.05.

7. Using the binomial distribution with n=12 and p=0.5, *P-value* = $P(R^* \geq 7 | p=0.5) = .3872$

8.Conclusion: cannot reject H_0. The median compressive strength is not more than 2250.

b. 1. The parameter of interest is the median compressive strength

2. $H_0 : \tilde{\mu} = 2250$

3. $H_1 : \tilde{\mu} > 2250$

4. α=0.05

5. Test statistic is $z_0 = \dfrac{r^+ - 0.5n}{0.5\sqrt{n}}$

6. We reject H_0 if the $|Z_0| > Z_{0.025} = 1.96$

7. Computation: $z_0 = \dfrac{7 - 0.5(12)}{0.5\sqrt{12}} = 0.577$

8. Conclusion: cannot reject H_0. The median compressive strength is not more than 2250.

The *P-value* = $1 - \Phi(0.58) = 1 - .7190 = 0.281$

15-7. 1. The parameter of interest is the median titanium content

2. $H_0 : \tilde{\mu} = 8.5$

3. $H_1 : \tilde{\mu} \neq 8.5$

4. α=0.05

5. Test statistic is $z_0 = \dfrac{r^+ - 0.5n}{0.5\sqrt{n}}$

6. We reject H_0 if the $|Z_0| > Z_{0.025} = 1.96$

7. Computation: $z_0 = \dfrac{7 - 0.5(20)}{0.5\sqrt{20}} = -1.34$

8. Conclusion: cannot reject H_0. The median titanium content is 8.5.

The *P-value* = $2 \cdot P(|Z| > 1.34) = 0.1802$.

15-9. 1. The parameters of interest are the median hardness readings for the two tips

2. $H_0 : \tilde{\mu}_D = 0$

3. $H_1 : \tilde{\mu}_D \neq 0$

4. α=0.05

5. The test statistic is $r = min(r^+, r^-)$.

6. Since α=0.05 and n=8, Appendix,= Table VII gives the critical value of $r^*_{0.05}$ =2. We will reject H_0 in favor of H_1 if $r \le 1$.

7. $r^+ = 6$ and $r^- = 2$ and so $r = min(6,2) = 2$

8. Conclusion: cannot reject H_0. There is not significant difference in the tips.

15-11. 1. The parameters of interest are the median drying times for the two formulations.

2. $H_0 : \tilde{\mu}_D = 0$

3. $H_1 : \tilde{\mu}_D \neq 0$

4. α=0.05

5. Test statistic is $z_0 = \dfrac{r^+ - 0.5n}{0.5\sqrt{n}}$

6. We reject H_0 if the $|Z_0| > Z_{0.025} = 1.96$

7. Computation: $z_0 = \dfrac{15 - 0.5(20)}{0.5\sqrt{20}} = 2.24$

8. Conclusion: reject H_0. There is a difference in the median drying times between the two formulations.

The *P-value* = 2*P(|Z|>2.24) = 0.025.

15-17. a) $\alpha = P(Z>1.96) = 0.025$

b) $\beta = P\left(\dfrac{X}{\sigma/\sqrt{n}} = 1.96 \mid \mu = 1\right) = P(Z < -1.20) = 0.115$

c) The sign test that rejects if $R^- \le 1$ has $\alpha = 0.011$ based on the binomial distribution.

d) $\beta = P(R^- > 1 \mid \mu = 1) = 0.1587$. 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.

<u>Section 15-3</u>

15-21 1. The parameter of interest is the mean titanium content

2. $H_0 : \mu = 8.5$

3. $H_1 : \mu \neq 8.5$

4. α=0.05

5. The test statistic is $w = min(w^+, w^-)$.

6. We will reject H_0 if $w \le w^*_{0.05} = 52$, since α=0.05 and n=20, the value in Appendix A, Table VIII.

7. $w^+ = 71$ and $w^- = 102$ and so $w = min(71,102) = 71$

8. Conclusion: Since 71>52, we cannot reject H_0.

15-23 1. The parameter of interest is the mean titanium content

2. $H_0 : \mu = 2.5$

3. $H_1 : \mu < 2.5$

4. α=0.05

5. The test statistic $w = min(w^+, w^-)$.

6. We will reject H_0 if $w \le w^*_{0.05} = 65$, since α=0.05 and n=22 the value in Appendix A, Table VIII.

7. $w^+ = 225$ and $w^- = 8$ and so $w = min(225, 8) = 8$

8.Conclusion: Since 8 < 65, we reject H_0.

15-27. 1. The parameters of interest are the mean blood cholesterol levels.

2. $H_0 : \mu_D = 0$

3. $H_1 : \mu_D \neq 0$

4. α=0.05

5. The test statistic is $w = min(w^+, w^-)$.

6. We will reject H_0 is $w \le w^*_{0.05} = 25$, since α=0.05 and n=15, the value in Appendix A, Table VIII.

7. $w^+ = 118$ and $w^- = 1$ and so $w = min(118, 1) = 1$ Since $1 < 25$
8. Conclusion: Since $1 < 25$, we reject H_0.

Section 15-4

15-31. 1. The parameters of interest are the mean image brightness'.
2. $H_0 : \mu_1 = \mu_2$
3. $H_1 : \mu_1 > \mu_2$
4. α=0.025
5. The test statistic is $z_0 = \frac{W_1 - \mu_{w_1}}{\sigma_{w_1}}$
6. We will reject H_0 if $Z_0 > Z_{0.025} = 1.96$
7. $w_1 = 40$, μ_{w_1} =85.5 and $\sigma^2_{w_1}$ =128.25

$$z_0 = \frac{54 - 85.5}{11.32} = -2.78$$

Since $Z_0 < 1.96$, cannot reject H_0
8. Conclusion: do not reject H_0.

P-value = 0.9973

15-35. 1. The parameters of interest are the mean etch rates
2. $H_0 : \mu_1 = \mu_2$
3. $H_1 : \mu_1 \ne \mu_2$
4. α=0.025
5. The test statistic is $z_0 = \frac{W_1 - \mu w_1}{\sigma_{w_1}}$
6. We will reject H_0 if $|Z_0| > Z_{0.025} = 1.96$
7. $w_1 = 55$, μ_{w_1} =105 and $\sigma^2_{w_1}$ =175

$$z_0 = \frac{55 - 105}{13.23} = -3.77$$

Since $|Z_0| < 1.96$, reject H_0
8. Conclusion: reject H_0.

P-value = 0.0001

Section 15-5

15-37.
```
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

Supplemental

15-43. 1. The parameter of interest is median of surface finish.

2. $H_0 : \tilde{\mu} = 10.0$

3 $H_1 : \tilde{\mu} \neq 10.0$

4. α=0.05

5. The test statistic is the observed number of plus differences or $r^+ = 5$.

6. We reject H_0 if the *P-value* corresponding to $r^+ = 5$ is less than or equal to α=0.05.

7. Using the binomial distribution with n=10 and p=0.5, *P-value* = $2P(R^* \geq 5 | p=0.5)=1.0$

8. Conclusion: we cannot reject H_0. We cannot reject the claim that the median is 10 μin.

15-45. The parameter of interest is the median fluoride emissions

$H_0 : \tilde{\mu} = 6$

$H_1 : \tilde{\mu} < 6$

α=0.05

Using Minitab (Sign Rank Test)

```
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

15-47. 1. The parameters of interest are the median impurity levels.

2. $H_0 : \tilde{\mu}_D = 0$

3. $H_1 : \tilde{\mu}_D \neq 0$

4. α=0.01

5. The test statistic is $r = min(r^+, r^-)$.

6. Since α=0.01 and n=8, Appendix,= Table VII gives the critical value of $r^*_{0.01}$ =0. We will reject H_0 in favor of H_1 if $r \leq 10$.

7. $r^+ = 1$ and $r^- = 7$ and so $r=min(1,7) = 1$

8. Conclusion: cannot reject H_0. There is no significant difference in the impurity levels.

15-49. The parameter of interest is the median fluoride emissions

$H_0 : \mu = 6$

$H_1 : \mu < 6$

α=0.05

```
Using Minitab Wilcoxon signed-rank t test
Test of median = 6.000 versus median < 6.000
                N for   Wilcoxon            Estimated
             N   Test  Statistic         P     Median
y           15     13       19.0     0.035      5.000
```

Reject H_0

The Wilcoxon signed-rank test applies to symmetric continuous distributions. The test to applies to the mean of the distribution.

15-51. 1. The parameters of interest are the mean volumes

2. $H_0 : \mu_1 = \mu_2$

3. $H_1 : \mu_1 \neq \mu_2$

4. α=0.01

5. The test statistic is $w_2 = \frac{(n_1 + n_2)(n_1 + n_2 + 1)}{2} - w_1$

6. We will reject H_0 if $w \leq w^*_{0.01} = 71$, since α=0.01 and n_1=10 and n_2=10, the value in Appendix A, Table IX.

7. $w_1 = 42$ and $w_2 = 78$ and so since 42 is less than 78, we reject H_0

8. Conclusion: reject H_0

15-57.

```
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)
```

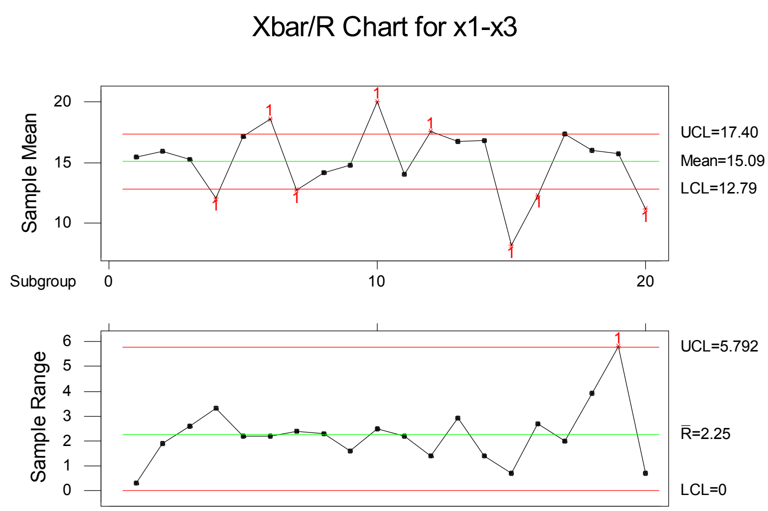
Reject H_0, P-value=0.0009

Chapter 16 Selected Problem Solutions

Section 16-5

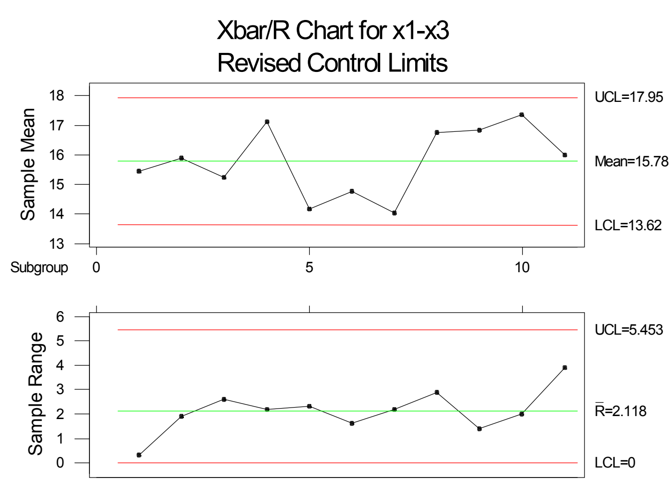
16-3 a)

```
                    X-bar and Range - Initial Study
                         Charting Problem 16-3
X-bar                                    | Range
-----                                    | -----
UCL:  +    3.0 sigma = 17.4              | UCL:  +    3.0 sigma =  5.792
Centerline           = 15.09             | Centerline           =  2.25
LCL:  -    3.0 sigma = 12.79             | LCL:  -    3.0 sigma =  0
                                         |

Test Results: X-bar One point more than 3.00 sigmas from center line.
Test Failed at points: 4 6 7 10 12 15 16 20

Test Results for R Chart:One point more than 3.00 sigmas from center line.
Test Failed at points: 19
```



.

b. Removed points 4, 6, 7, 10, 12, 15, 16, 19, and 20 and revised the control limits
The control limits are not as wide after being revised X-bar UCL=17.96, CL=15.78 LCL=13.62 and R UCL = 5.453, R-bar=2.118, LCL=0.
The X-bar control moved down.

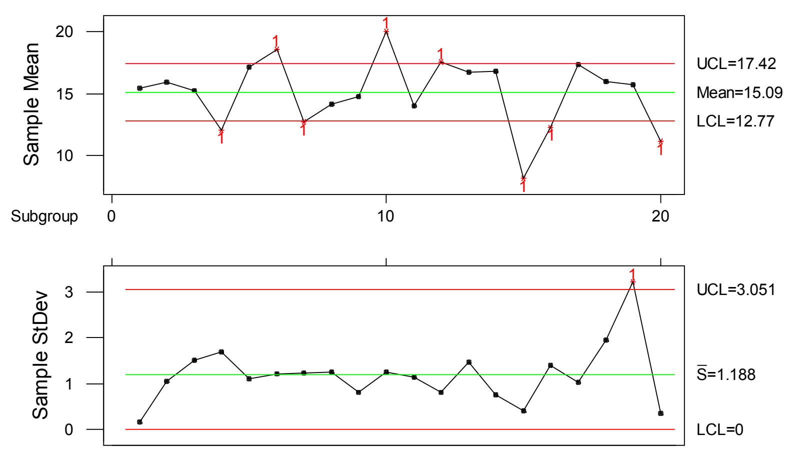


c)

```
                     X-bar and StDev - Initial Study
                       Charting Problem 16-3
X-bar                                    | StDev
-----                                    | -----
UCL:  +    3.0 sigma = 17.42             | UCL:  +    3.0 sigma =  3.051
Centerline           = 15.09             | Centerline           =  1.188
LCL:  -    3.0 sigma = 12.77             | LCL:  -    3.0 sigma =  0
                                         |

Test Results: X-bar One point more than 3.00 sigmas from center line.
Test Failed at points: 4 6 7 10 12 15 16 20

Test Results for S Chart:One point more than 3.00 sigmas from center line.
Test Failed at points: 19
```

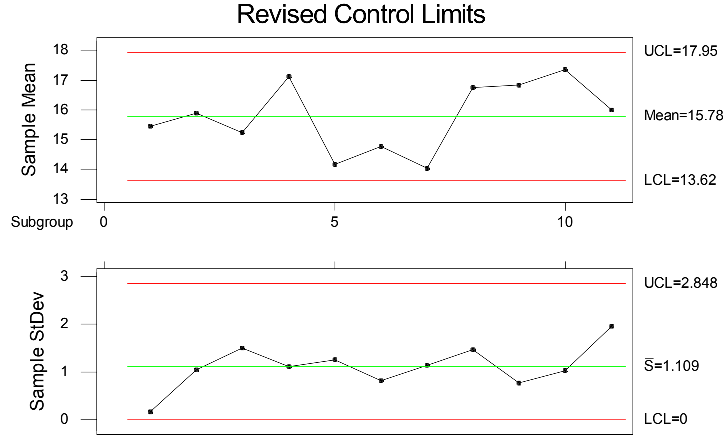
.

Xbar/S Chart for x1-x3



Removed points 4, 6, 7, 10, 12, 15, 16, 19, and 20 and revised the control limits
The control limits are not as wide after being revised X-bar UCL=17.95, CL=15.78
LCL=13.62 and S UCL = 2.848, S-bar=1.109, LCL=0.
The X-bar control moved down.

Xbar/S Chart for x1-x3
Revised Control Limits



16-5. a) $\bar{\bar{x}} = \dfrac{7805}{35} = 223 \qquad \bar{r} = \dfrac{1200}{35} = 34.286$

$\bar{x}$ 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

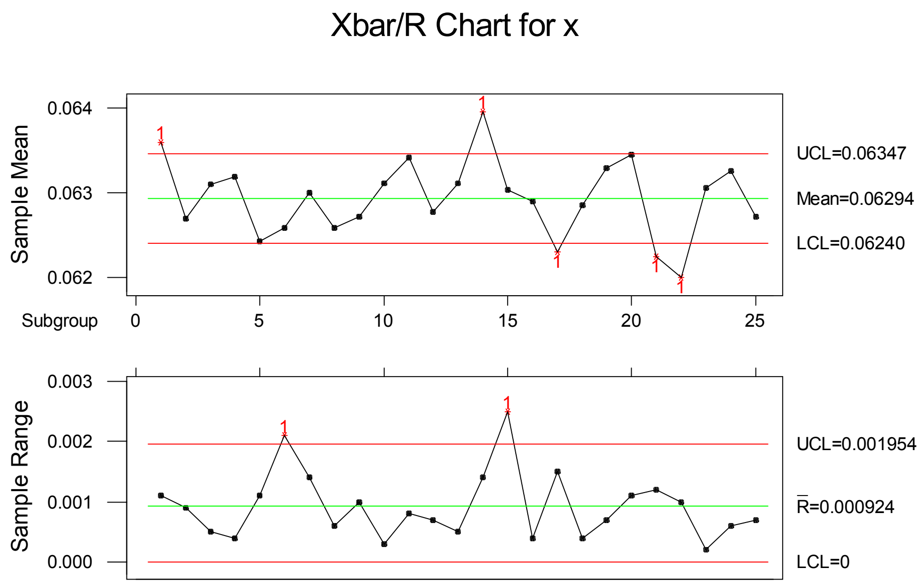
$$UCL = D_4\bar{r} = 2.115(34.286) = 72.51$$
$$CL = 34.286$$
$$LCL = D_3\bar{r} = 0(34.286) = 0$$

b)

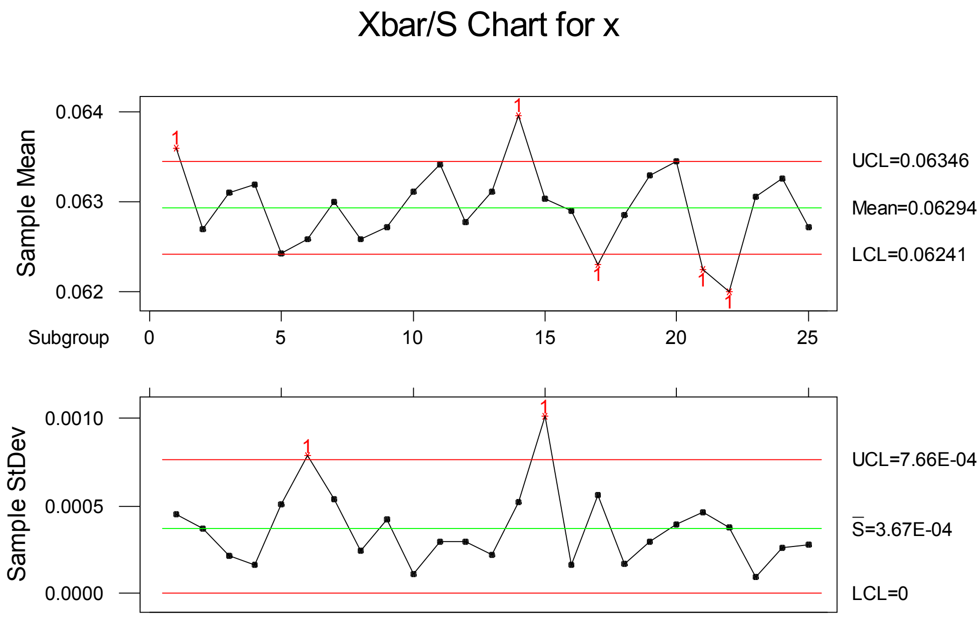
$$\hat{\mu} = \bar{\bar{x}} = 223$$
$$\hat{\sigma} = \frac{\bar{r}}{d_2} = \frac{34.286}{2.326} = 14.74$$

16-7 a)

```
                    X-bar and Range - Initial Study
                        Charting Problem 16-7
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
```



b)



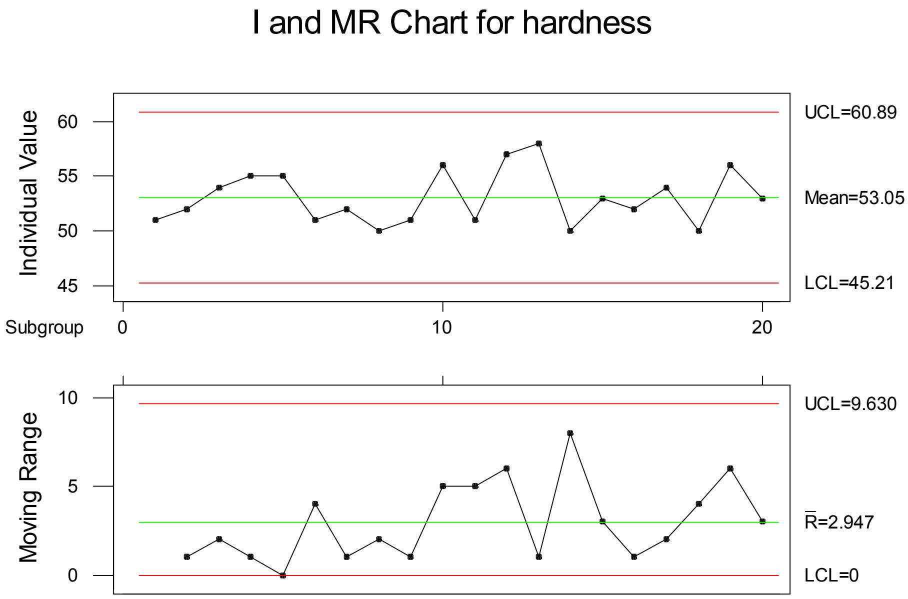
c) There are several points out of control. The control limits need to be revised.
The points are 1, 5, 14,17, 20,21, and 22; or outside the control limits of the R chart: 6 and 15

Section 16-6

16-9. 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       =  2.94737
```

There are no points beyond the control limits. The process appears to be in control.



b)

$$\hat{\mu} = \bar{x} = 53.05$$

$$\hat{\sigma} = \frac{\overline{mr}}{d_2} = \frac{2.94737}{1.128} = 2.613$$

Section 16-7

16-15. a) Assuming a normal distribution with $\hat{\mu}$ = 0.14.510 and $\hat{\sigma} = \frac{\bar{r}}{d_2} = \frac{0.344}{2.326} = 0.148$

$$\begin{aligned} P(X < LSL) &= P\left(Z < \frac{LSL - \hat{\mu}}{\hat{\sigma}}\right) \\ &= P\left(Z < \frac{14.00 - 14.51}{0.148}\right) \\ &= P(Z < -3.45) \\ &= 1 - P(Z < 3.45) \\ &= 1 - 0.99972 \\ &= 0.00028 \end{aligned}$$

$$
\begin{aligned}
P(X > USL) &= P\left(Z > \frac{USL - \hat{\mu}}{\hat{\sigma}}\right) \\
&= P\left(Z > \frac{15.00 - 14.51}{0.148}\right) \\
&= P(Z > 3.31) \\
&= 1 - P(Z < 3.31) \\
&= 1 - 0.99953 \\
&= 0.00047
\end{aligned}
$$

Therefore, the proportion nonconforming is given by
P(X<LSL) + P(X>USL) =0.00028 + 0.00047 = 0.00075

b)

$$PCR = \frac{USL - LSL}{6(\hat{\sigma})} = \frac{15.00 - 14.00}{6(0.148)} = 1.13$$

$$
\begin{aligned}
PCR_K &= \min\left[\frac{USL - \bar{\bar{x}}}{3\hat{\sigma}}, \frac{\bar{\bar{x}} - LSL}{3\hat{\sigma}}\right] \\
&= \min\left[\frac{15.00 - 14.51}{3(0.148)}, \frac{14.51 - 14.00}{3(0.148)}\right] \\
&= \min[1.104, 1.15] \\
&= 1.104
\end{aligned}
$$

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.

16-19 a) Assuming a normal distribution with $\hat{\mu}$ = 223 and $\hat{\sigma} = \frac{\bar{s}}{c_4} = \frac{13.58}{0.9213} = 14.74$

$$
\begin{aligned}
P(X < LSL) &= P\left(Z < \frac{LSL - \hat{\mu}}{\hat{\sigma}}\right) \\
&= P\left(Z < \frac{170 - 223}{14.74}\right) \\
&= P(Z < -3.60) \\
&= 0.00016
\end{aligned}
$$

$$
\begin{aligned}
P(X > USL) &= P\left(Z > \frac{USL - \hat{\mu}}{\hat{\sigma}}\right) \\
&= P\left(Z > \frac{270 - 223}{14.75}\right) \\
&= P(Z > 3.18) \\
&= 1 - P(Z < 3.18) \\
&= 1 - 0.99926 \\
&= 0.00074
\end{aligned}
$$

Probability of producing a part outside the specification limits is 0.00016+0.00074 = 0.0009

b

$$
PCR = \frac{USL - LSL}{6(\hat{\sigma})} = \frac{270 - 220}{6(14.75)} = 1.13
$$

$$
\begin{aligned}
PCR_K &= \min\left[\frac{USL - \bar{\bar{x}}}{3\hat{\sigma}}, \frac{\bar{\bar{x}} - LSL}{3\hat{\sigma}}\right] \\
&= \min\left[\frac{270 - 223}{3(14.75)}, \frac{223 - 170}{3(14.75)}\right] \\
&= \min[1.06, 1.19] \\
&= 1.06
\end{aligned}
$$

Since PCR exceeds unity, the natural tolerance limits lie inside the specification limits and very few defective units will be produced. The estimated proportion nonconforming is given by $P(X<LSL) + P(X>USL) = 0.00016 + 0.00074 = 0.0009$

16-23. Assuming a normal distribution with $\hat{\mu} = 500.6$ and $\hat{\sigma} = 17.17$

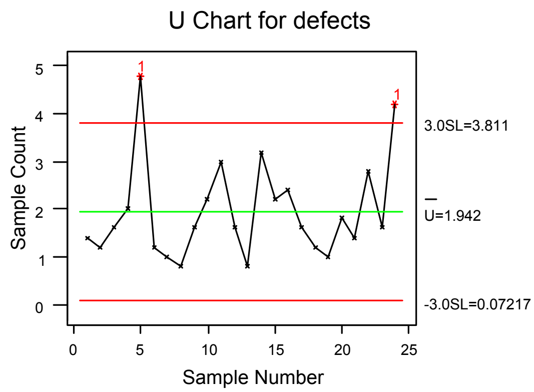
$$
PCR = \frac{USL - LSL}{6(\hat{\sigma})} = \frac{525 - 475}{6(17.17)} = 0.49
$$

$$
\begin{aligned}
PCR_K &= \min\left[\frac{USL - \bar{\bar{x}}}{3\hat{\sigma}}, \frac{\bar{\bar{x}} - LSL}{3\hat{\sigma}}\right] \\
&= \min\left[\frac{525 - 500.6}{3(17.17)}, \frac{500.6 - 475}{3(17.17)}\right] \\
&= \min[0.474, 0.50] \\
&= 0.474
\end{aligned}
$$

Since the process capability ratios are less than unity, the process capability appears to be poor.

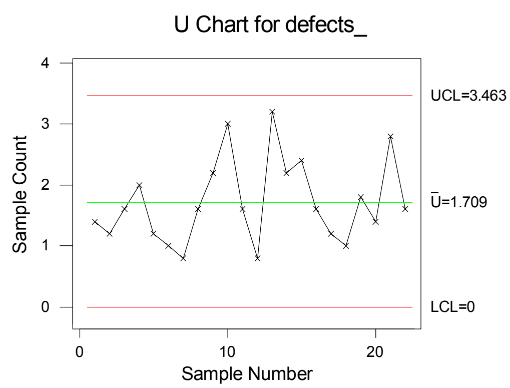
Section 16-8

16-25.



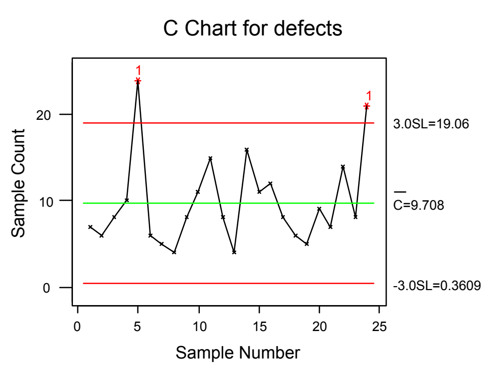
Samples 5 and 24 have out-of-control points. The limits need to be revised.

b)



The control limits are calculated without the out-of-control points. There are no points out of control for the revised limits.

16-27.



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.

Section 16-9

16-31. a) $\hat{\sigma}_{\bar{x}} = \dfrac{\hat{\sigma}}{\sqrt{n}} = \dfrac{2.4664}{\sqrt{5}} = 1.103$, $\mu = 36$

$$P(30.78 < \overline{X} < 37.404)$$
$$= P\left(\frac{30.78-36}{1.103} < \frac{\overline{X}-\mu}{\hat{\sigma}_{\bar{x}}} < \frac{37.404-36}{1.103}\right)$$
$$= P(-4.73 < Z < 1.27) = P(Z < 1.27) - P(Z < -4.73)$$
$$= 0.8980 - 0 = 0.8980$$

The probability that this shift will be detected on the next sample is p = 1–0.8980 = 0.1020.

b) $ARL = \dfrac{1}{p} = \dfrac{1}{0.102} = 9.8$

16-33. a) $\hat{\sigma} = \dfrac{\overline{R}}{d_2} = \dfrac{6.75}{2.059} = 3.28$ $\hat{\sigma}_{\bar{x}} = \dfrac{\hat{\sigma}}{\sqrt{n}} = \dfrac{3.28}{\sqrt{4}} = 1.64$, $\mu = 13$

$$P(5.795 < \overline{X} < 15.63)$$
$$= P\left(\frac{5.795-13}{1.64} < \frac{\overline{X}-\mu}{\sigma_{\bar{x}}} < \frac{15.63-13}{1.64}\right)$$
$$= P(-4.39 < Z < 1.60) = P(Z < 1.60) - P(Z < -4.39)$$
$$= 0.9452 - 0 = 0.9452$$

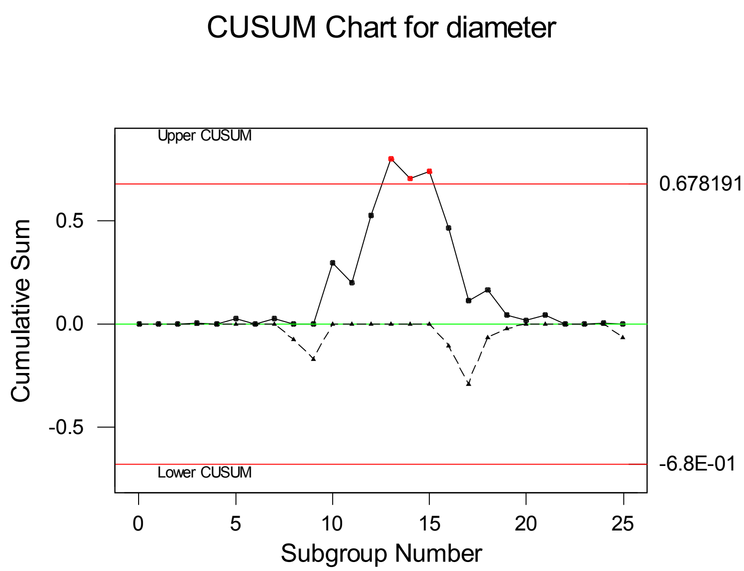
The probability that this shift will be detected on the next sample is p = 1–0.9452 = 0.0548.

b) $ARL = \dfrac{1}{p} = \dfrac{1}{0.0548} = 18.25$

Section 16-10

16-39. a) $\hat{\sigma} = 0.1695$

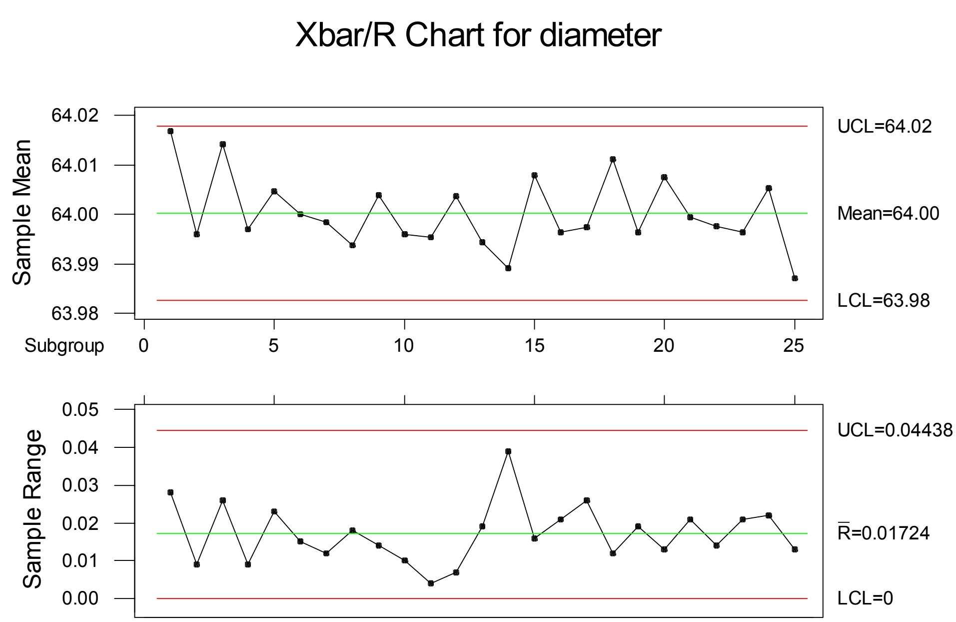
b) The process appears to be out of control at the specified target level.



Supplemental

16-43. 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) $\hat{\mu} = \overline{\overline{x}} = 64$ $\quad \hat{\sigma} = \dfrac{\overline{R}}{d_2} = \dfrac{0.01764}{1.693} = 0.0104$

c) $PCR = \dfrac{USL - LSL}{6\hat{\sigma}} = \dfrac{64.02 - 63.98}{6(0.0104)} = 0.641$

The process does not meet the minimum capability level of PCR ≥ 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 $PCR_k = \dfrac{\overline{\overline{x}} - LSL}{3\sigma} = 2.0$ for σ:

$$\frac{64 - 61}{3\sigma} = 2.0$$

$$6\sigma = 64. - 61$$

$$\sigma = \frac{64. - 61}{6} = 0.50$$

Therefore, the process variance would have to be decreased to $\sigma^2 = (0.50)^2 = 0.025$.

f) $\hat{\sigma}_{\overline{x}} = 0.0104$

$$P(63.98 < X < 64.02)$$

$$= P\left(\frac{63.98 - 64.01}{0.0104} < \frac{X - \mu}{\sigma_x} < \frac{64.02 - 64.01}{0.0104}\right)$$

$$= P(-2.88 < Z < 0.96) = P(Z < 0.96) - P(Z < -2.88)$$

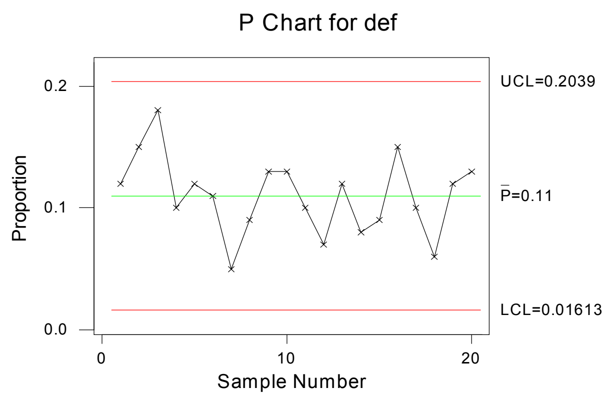
$$= 0.8315 - 0.0020 = 0.8295$$

The probability that this shift will be detected on the next sample is p = 1–0.8295 = 0.1705

$$ARL = \frac{1}{p} = \frac{1}{0.1705} = 5.87$$

16-45. a)

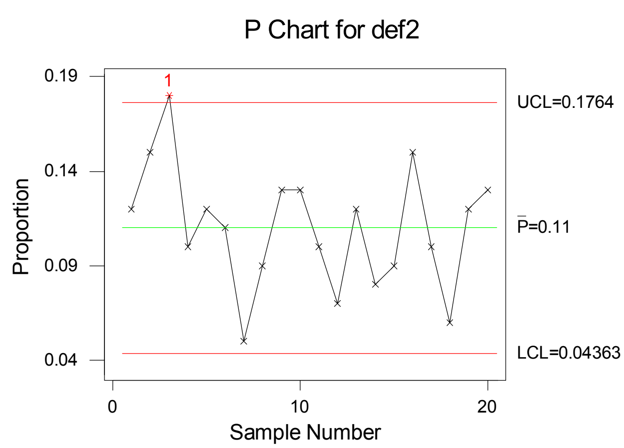
```
                              P Chart - Initial Study
P Chart
-----
UCL:  +    3.0 sigma = 0.203867
Centerline           = 0.11
LCL:  -    3.0 sigma = 0.0161331
out of limits = 0
Estimated
mean P  = 0.11
sigma   = 0.031289
```



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
```

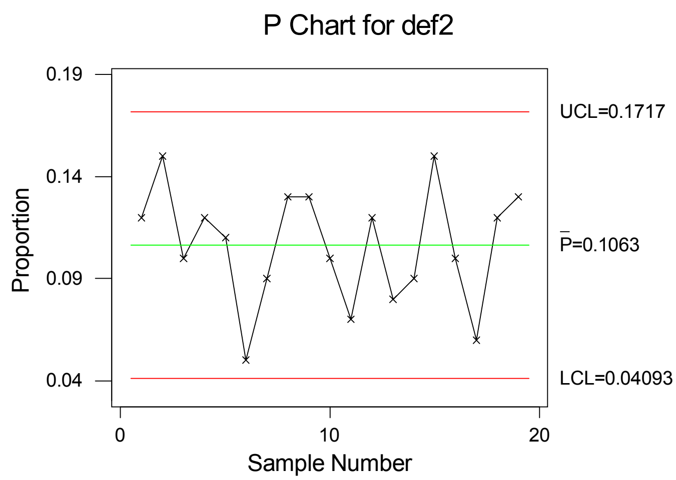


There is one point beyond the upper control limit. The process is out of control. The revised limits are:

```
                                   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
```

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.

16-49. ARL = 1/p where p is the probability a point falls outside the control limits.

a) $\mu = \mu_0 + \sigma$ and n = 1

$$p = P(\overline{X} > UCL) + P(\overline{X} < LCL)$$

$$= P\left(Z > \frac{\mu_0 + \frac{3\sigma}{\sqrt{n}} - \mu_0 - \sigma}{\sigma/\sqrt{n}} \right) + P\left(Z < \frac{\mu_0 - \frac{3\sigma}{\sqrt{n}} - \mu_0 - \sigma}{\sigma/\sqrt{n}} \right)$$

$$= P(Z > 3 - \sqrt{n}) + P(Z < -3 - \sqrt{n})$$

$$= P(Z > 2) + P(Z < -4) \qquad when \quad n = 1$$

$$= 1 - P(Z < 2) + [1 - P(Z < 4)] \quad = 1 - 0.97725 + [1 - 1] \quad = 0.02275$$

Therefore, ARL = 1/p = 1/0.02275 = 43.9.

b) $\mu = \mu_0 + 2\sigma$

$$P(\overline{X} > UCL) + P(\overline{X} < LCL)$$

$$= P\left(Z > \frac{\mu_0 + \frac{3\sigma}{\sqrt{n}} - \mu_0 - 2\sigma}{\sigma / \sqrt{n}} \right) + P\left(Z < \frac{\mu_0 - \frac{3\sigma}{\sqrt{n}} - \mu_0 - 2\sigma}{\sigma / \sqrt{n}} \right)$$

$$= P(Z > 3 - 2\sqrt{n}) + P(Z < -3 - 2\sqrt{n})$$

$$= P(Z > 1) + P(Z < -5) \qquad when \quad n = 1$$

$$= 1 - P(Z < 1) + [1 - P(Z < 5)]$$

$$= 1 - 0.84134 + [1 - 1]$$

$$= 0.15866$$

Therefore, ARL = 1/p = 1/0.15866 = 6.30.

c) $\mu = \mu_0 + 3\sigma$

$$P(\overline{X} > UCL) + P(\overline{X} < LCL)$$

$$= P\left(Z > \frac{\mu_0 + \frac{3\sigma}{\sqrt{n}} - \mu_0 - 3\sigma}{\sigma / \sqrt{n}} \right) + P\left(Z < \frac{\mu_0 - \frac{3\sigma}{\sqrt{n}} - \mu_0 - 3\sigma}{\sigma / \sqrt{n}} \right)$$

$$= P(Z > 3 - 3\sqrt{n}) + P(Z < -3 - 3\sqrt{n})$$

$$= P(Z > 0) + P(Z < -6) \qquad when \quad n = 1$$

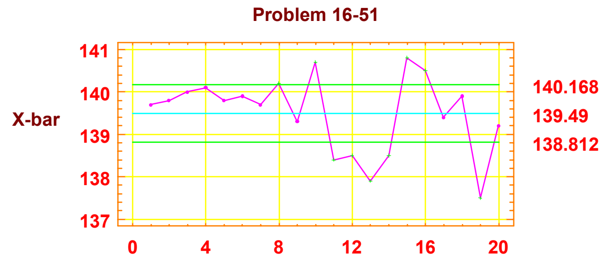
$$= 1 - P(Z < 0) + [1 - P(Z < 6)] = 1 - 0.50 + [1 - 1] = 0.50$$

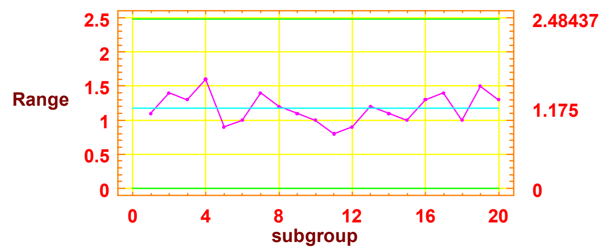
Therefore, ARL = 1/p = 1/0.50 = 2.00.

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.

16-51. 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
```

z

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) $\hat{\sigma}_{\overline{x}} = 0.529$

$$p = P(139.098 < X < 140.518 \mid \mu = 139.7)$$

$$= P\left(\frac{139.098-139.7}{0.529} < \frac{X-\mu}{\sigma_x} < \frac{140.518-139.7}{0.529}\right)$$

$$= P(-1.14 < Z < 1.55) = P(Z < 1.55) - P(Z < -1.14)$$

$$= P(Z < 1.55) - [1 - P(Z < 1.14)] = 0.93943 - [1 - 0.87285] = 0.8123$$

The probability that this shift will be detected on the next sample is 1–p = 1–0.8123 = 0.1877.

$$\text{ARL} = \frac{1}{1-\text{p}} = \frac{1}{0.1877} = 5.33$$