CHAPTER 2

Section 2-1

2-1. Let "a", "b" denote a part above, below the specification
$S = \{aaa, aab, aba, abb, baa, bab, bba, bbb\}$

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

$$S = \left\{\begin{matrix} eeee, eoee, oeee, ooee, \\ eeeo, eoeo, oeeo, ooeo, \\ eeoe, eooe, oeoe, oooe, \\ eeoo, eooo, oeoo, oooo \end{matrix}\right\}$$

2-3. Let "a" denote an acceptable power supply
Let "f" ,"m","c" denote a supply with a functional, minor, or cosmetic error, respectively.
$S = \{a, f, m, c\}$

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

2-5. If only the number of tracks with errors is of interest, then $S = \{0, 1, 2, \ldots, 24\}$

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

2-7. S is the sample space of 100 possible two digit integers.

2-8. Let an ordered pair of numbers, such as 43 denote the response on the first and second question. Then, S consists of the 25 ordered pairs $\{11, 12, \ldots, 55\}$

2-9. $S = \{0, 1, 2, \ldots,\}$ in ppb.

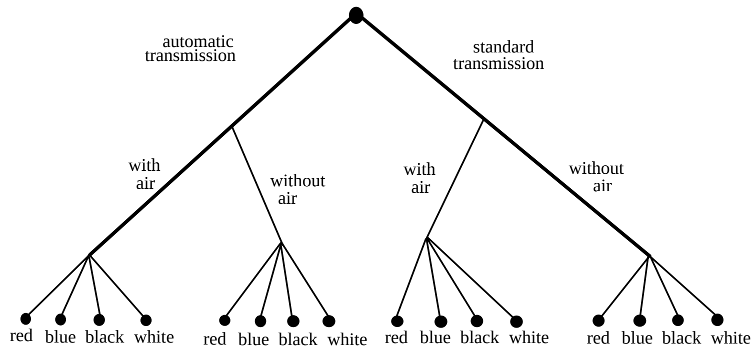
2-10. $S = \{0, 1, 2, \ldots,\}$ in milliseconds

2-11. $S = \{1.0, 1.1, 1.2, \ldots 14.0\}$

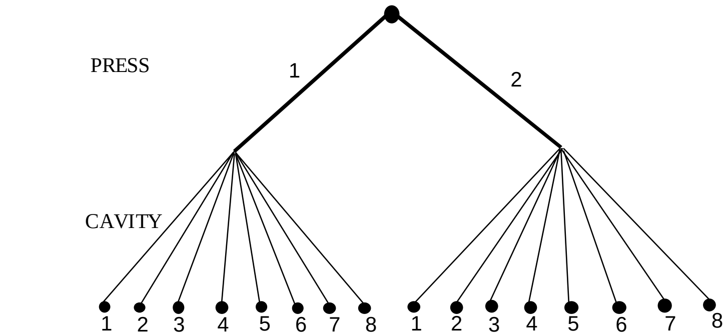
2-12. s = small, m = medium, l = large; S = {s, m, l, ss, sm, sl, ….}

2-13 $S = \{0, 1, 2, \ldots,\}$ in milliseconds.

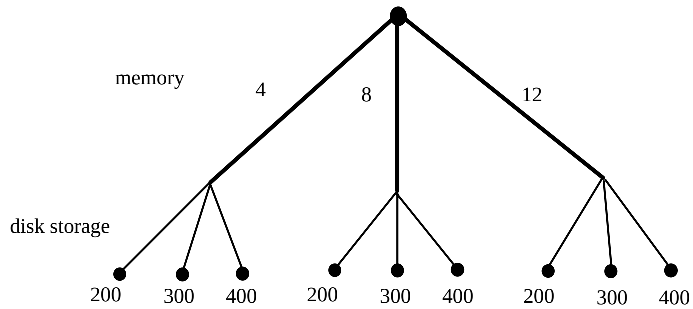
2-14.



2-15.



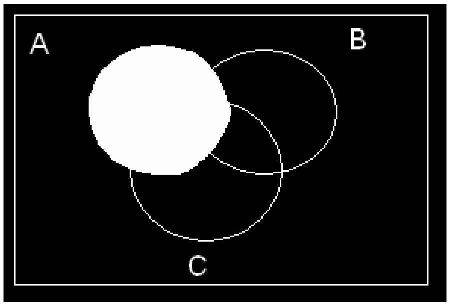
2-16.



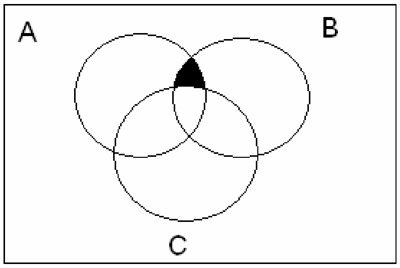
2-17. c = connect, b = busy, S = {c, bc, bbc, bbbc, bbbbc, …}

2-18. $S = \{s, fs, ffs, fffS, fffFS, fffFFS, fffFFFA\}$

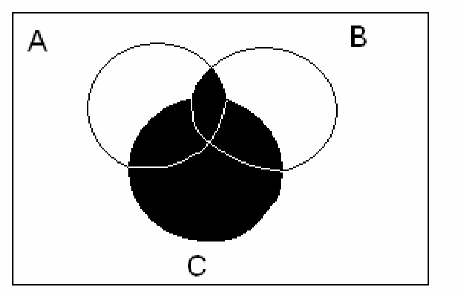
2-19 a.)



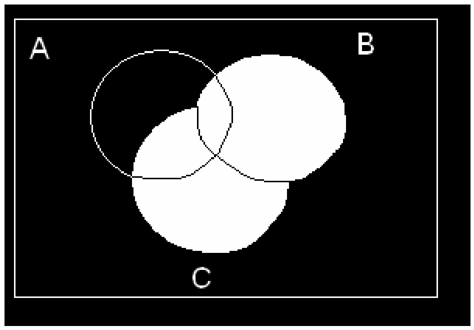
b.)



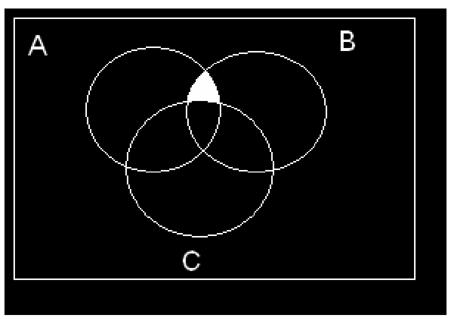
c.)



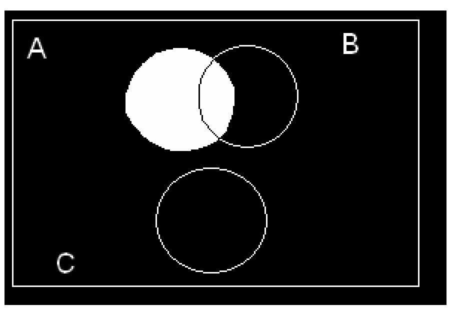
d.)



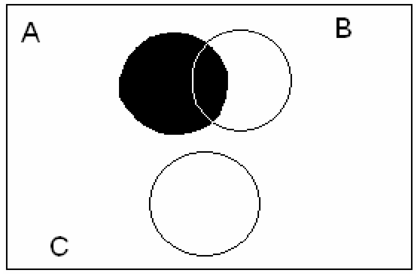
e.)



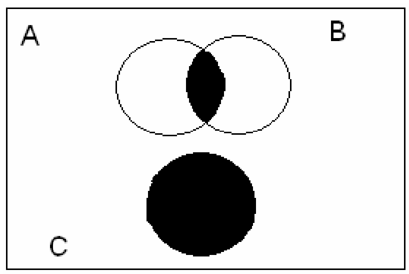
2.20 a.)



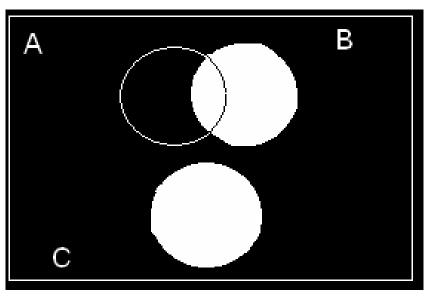
b.)



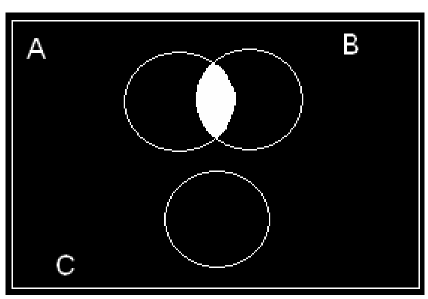
c.)



d.)



e.)



2-21. a) S = nonnegative integers from 0 to the largest integer that can be displayed by the scale.
Let X represent weight.
A is the event that $X > 11$
B is the event that $X \le 15$
C is the event that $8 \le X < 12$
S = {0, 1, 2, 3, …}

b) S

c) $11 < X \le 15$ or {12, 13, 14, 15}

d) $X \le 11$ or {0, 1, 2, …, 11}

e) S

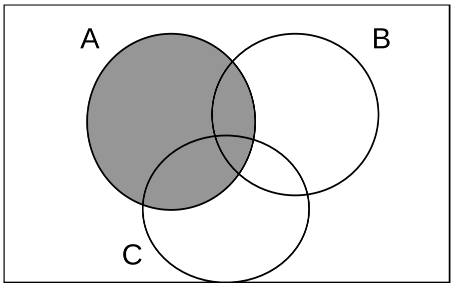
f) $A \cup C$ would contain the values of X such that: $X \ge 8$
Thus $(A \cup C)'$ would contain the values of X such that: $X < 8$ or {0, 1, 2, …, 7}

g) ∅

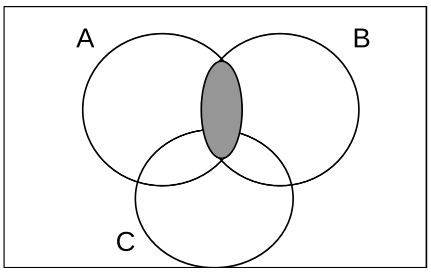
h) B′ would contain the values of X such that $X > 15$. Therefore, $B' \cap C$ would be the empty set. They have no outcomes in common or ∅

i) $B \cap C$ is the event $8 \le X < 12$. Therefore, $A \cup (B \cap C)$ is the event $X \ge 8$ or {8, 9, 10, …}

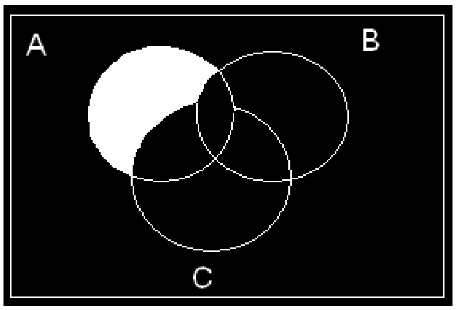
2-22. a)



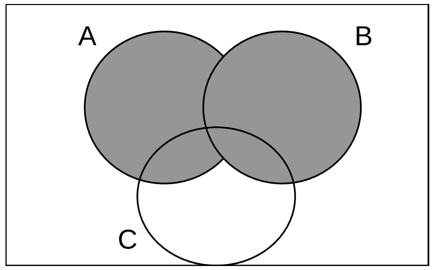
b)



c)



d.)



e.) If the events are mutually exclusive, then $A \cap B$ is equal to zero. Therefore, the process would not produce product parts with X=50 cm and Y=10 cm. The process would not be successful

2-23. Let "d" denoted a distorted bit and let "o" denote a bit that is not distorted.

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

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

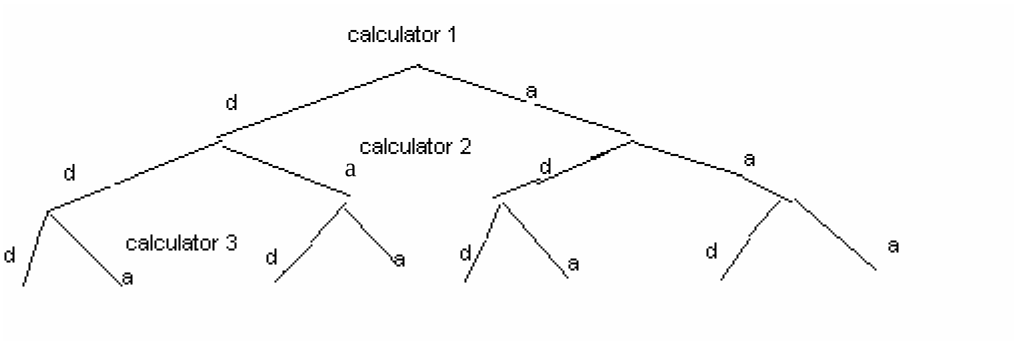
c) $A_1 = \left\{ \begin{array}{l} dddd, dodd, \\ dddo, dodo \\ ddod, dood \\ ddoo, dooo \end{array} \right\}$

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

e) $A_1 \cap A_2 \cap A_3 \cap A_4 = \{dddd\}$

f) $(A_1 \cap A_2) \cup (A_3 \cap A_4) = \{dddd, dodd, dddo, oddd, ddod, oodd, ddoo\}$

2-24. Let "d" denote a defective calculator and let "a" denote an acceptable calculator



a) $S = \{ddd, add, dda, ada, dad, aad, daa, aaa\}$

b) $A = \{ddd, dda, dad, daa\}$

c) $B = \{ddd, dda, add, ada\}$

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

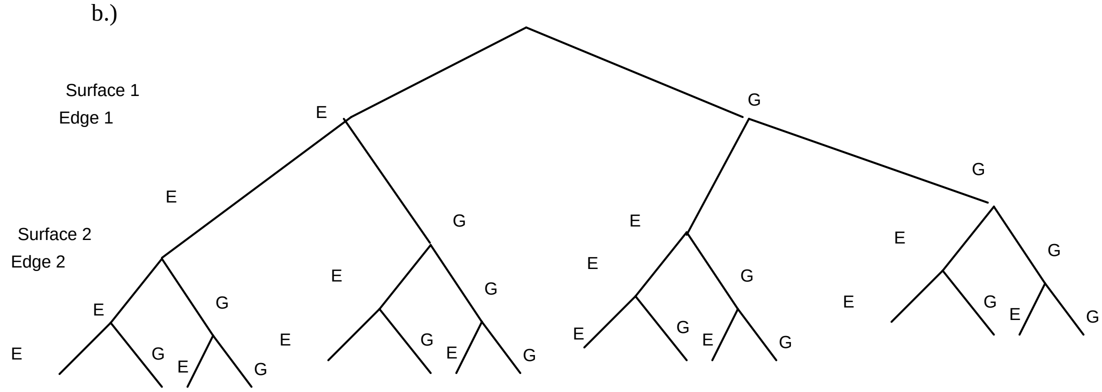
e) $B \cup C = \{ddd, dda, add, ada, dad, aad\}$

2-25. $2^{12} = 4096$

2-26. $A \cap B = 70$, $A' = 14$, $A \cup B = 95$

2-27. a.) $A' \cap B = 10$, $B' = 10$, $A \cup B = 92$

b.)



2-28. $A' \cap B = 55$, $B' = 23$, $A \cup B = 85$

2-29. a) $A' = \{x \mid x \geq 72.5\}$

b) $B' = \{x \mid x \leq 52.5\}$

c) $A \cap B = \{x \mid 52.5 < x < 72.5\}$

d) $A \cup B = \{x \mid x > 0\}$

2.30 a) {*ab, ac, ad, bc, bd, cd, ba, ca, da, cb, db, dc*}

b) {*ab, ac, ad, ae, af, ag, bc, bd, be, bf, bg, cd, ce, cf, cg, ef, eg, fg, ba, ca, da, ea, fa, ga, cb, db, eb, fb, gb, dc, ec, fc, gc, fe, ge, gf*}

c) Let d = defective, g = good; S = {*gg, gd, dg, dd*}

d) Let d = defective, g = good; S = {*gd, dg, gg*}

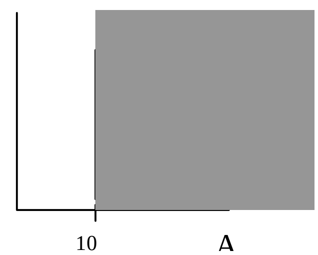
2.31 Let *g* denote a good board, *m* a board with minor defects, and *j* a board with major defects.

a.) S = {*gg, gm, gj, mg, mm, mj, jg, jm, jj*}

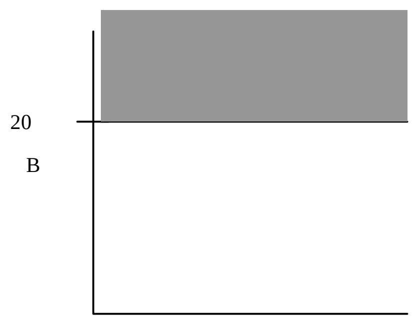
b) *S*={*gg,gm,gj,mg,mm,mj,jg,jm*}

2-32.a.) The sample space contains all points in the positive *X*-*Y* plane.

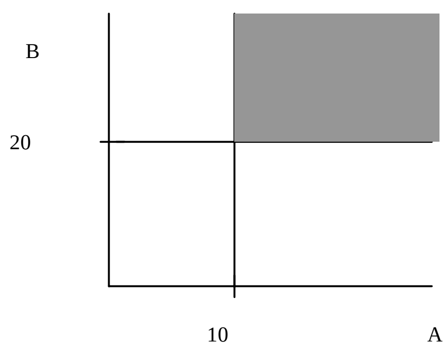
b)



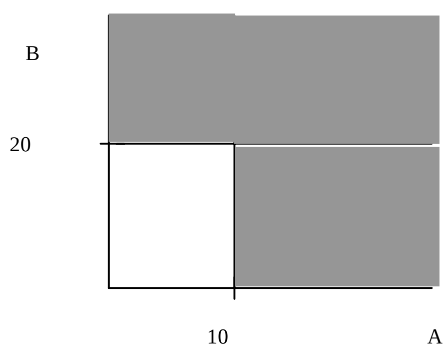
c)



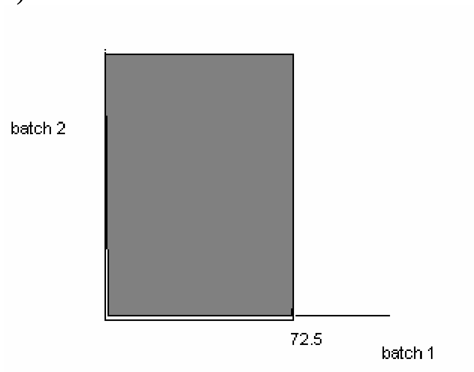
d)



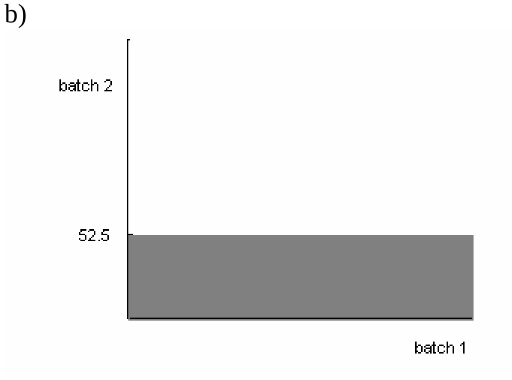
e)



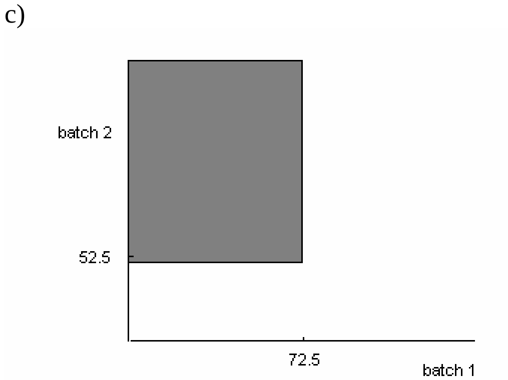
2-33 a)


batch 2
72.5
batch 1

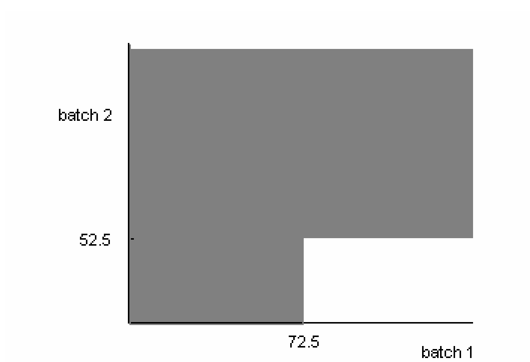
b)


batch 2
52.5
batch 1

c)


batch 2
52.5
72.5
batch 1

d)


batch 2
52.5
72.5
batch 1

Section 2-2

2-34. All outcomes are equally likely
a) P(A) = 2/5
b) P(B) = 3/5
c) P(A') = 3/5
d) $P(A\cup B) = 1$
e) $P(A\cap B) = P(\emptyset) = 0$

2-35. a) P(A) = 0.4
b) P(B) = 0.8
c) P(A') = 0.6
d) $P(A\cup B) = 1$
e) $P(A\cap B) = 0.2$

2-36. a) S = {1, 2, 3, 4, 5, 6}
b) 1/6
c) 2/6
d) 5/6

2-37. a) S = {1,2,3,4,5,6,7,8}
b) 2/8
c) 6/8

2-38. $\frac{x}{20} = 0.3, x = 6$

2-39. a) 0.5 + 0.2 = 0.7
b) 0.3 + 0.5 = 0.8

2-40. a) 1/10
b) 5/10

2-41. a) 0.25
b) 0.75

2-42. Total possible: 10^{16}, Only 10^{8} valid, P(valid) = $10^{8}/10^{16} = 1/10^{8}$

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 $(1/10^3)*(1/26^3) = 5.7 \times 10^{-8}$

2-44. a) 5*5*4 = 100
b) (5*5)/100 = 25/100=1/4

2-45. a) P(A) = 86/100 = 0.86
b) P(B) = 79/100 = 0.79
c) P(A') = 14/100 = 0.14
d) $P(A\cap B)$ = 70/100 = 0.70
e) $P(A\cup B)$ = (70+9+16)/100 = 0.95
f) $P(A'\cup B)$ = (70+9+5)/100 = 0.84

2-46. Let A = excellent surface finish; B = excellent length
a) P(A) = 82/100 = 0.82
b) P(B) = 90/100 = 0.90
c) P(A') = 1 – 0.82 = 0.18
d) $P(A\cap B)$ = 80/100 = 0.80
e) $P(A\cup B)$ = 0.92
f) $P(A'\cup B)$ = 0.98

2-47. a) P(A) = 30/100 = 0.30
b) P(B) = 77/100 = 0.77
c) P(A') = 1 – 0.30 = 0.70
d) $P(A\cap B)$ = 22/100 = 0.22
e) $P(A\cup B)$ = 85/100 = 0.85
f) $P(A'\cup B)$ =92/100 = 0.92

2-48. a) Because E and E' are mutually exclusive events and $E\cup E'$ = S
1 = P(S) = $P(E\cup E')$ = P(E) + P(E'). Therefore, P(E') = 1 - P(E)
b) Because S and $\varnothing$ are mutually exclusive events with S = $S\cup\varnothing$
P(S) = P(S) + P($\varnothing$). Therefore, P($\varnothing$) = 0
c) Now, B = $A\cup(A'\cap B)$ and the events A and $A'\cap B$ are mutually exclusive. Therefore,
P(B) = P(A) + $P(A'\cap B)$. Because $P(A'\cap B) \geq 0$, P(B) $\geq$ P(A).

Section 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

2-50. a) $P(A\cup B\cup C)$ = P(A) + P(B) + P(C), because the events are mutually exclusive. Therefore,
$P(A\cup B\cup C)$ = 0.2+0.3+0.4 = 0.9
b) $P(A\cap B\cap C)$ = 0, because $A\cap B\cap C = \varnothing$
c) $P(A\cap B)$ = 0 , because $A\cap B = \varnothing$
d) $P((A\cup B)\cap C)$ = 0, because $(A\cup B)\cap C = (A\cap C)\cup(B\cap C) = \varnothing$
e) $P(A'\cup B'\cup C')$ =1-[P(A) + P(B) + P(C)] = 1-(0.2+0.3+0.4) = 0.1

2-51. If A,B,C are mutually exclusive, then $P(A\cup B\cup C)$ = P(A) + P(B) + P(C) = 0.3 + 0.4 + 0.5 = 1.2, which greater than 1. Therefore, P(A), P(B),and P(C) cannot equal the given values.

2-52. a) 70/100 = 0.70
b) (79+86-70)/100 = 0.95
c) No, $P(A\cap B) \neq 0$

2-53. a) 350/370
b) $\frac{345+5+12}{370} = \frac{362}{370}$
c) $\frac{345+5+8}{370} = \frac{358}{370}$
d) 345/370

2-54. a) 170/190 = 17/19
b) 7/190

2-55. a) P(unsatisfactory) = (5+10-2)/130 = 13/130
b) P(both criteria satisfactory) = 117/130 = 0.90, No

2-56. a) (207+350+357-201-204-345+200)/370 = 0.9838
b) 366/370 = 0.989
c) (200+145)/370 = 363/370 = 0.981
d) (201+149)/370 = 350/370 = 0.946

Section 2-4

2-57. a) P(A) = 86/100 b) P(B) = 79/100

c) $P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{70/100}{79/100} = \frac{70}{79}$

d) $P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{70/100}{86/100} = \frac{70}{86}$

2-58.a) 0.82

b) 0.90

c) 8/9 = 0.889

d) 80/82 = 0.9756

e) 80/82 = 0.9756

f) 2/10 = 0.20

2-59. a) 345/357 b) 5/13

2-60. a) 12/100 b) 12/28 c) 34/122

2-61. Need data from Table 2-2 on page 34

a) P(A) = 0.05 + 0.10 = 0.15

b) $P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{0.04 + 0.07}{0.72} = 0.153$

c) P(B) = 0.72

d) $P(B|A) = \frac{P(A \cap B)}{P(B)} = \frac{0.04 + 0.07}{0.15} = 0.733$

e) $P(A \cap B) = 0.04 + 0.07 = 0.11$

f) $P(A \cup B) = 0.15 + 0.72 - 0.11 = 0.76$

2-62. a) 20/100

b) 19/99

c) (20/100)(19/99) = 0.038

d) If the chips are replaced, the probability would be (20/100) = 0.2

2-63. a) P(A) = 15/40

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

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

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

2-64. A = first is local, B = second is local, C = third is local

a) $P(A \cap B \cap C) = (15/40)(14/39)(13/38) = 0.046$

b) $P(A \cap B \cap C') = (15/40)(14/39)(25/39) = 0.085$

2-65. a) 4/499 = 0.0080

b) (5/500)(4/499) = 0.000080

c) (495/500)(494/499) = 0.98

2-66. a) 3/498 = 0.0060

b) 4/498 = 0.0080

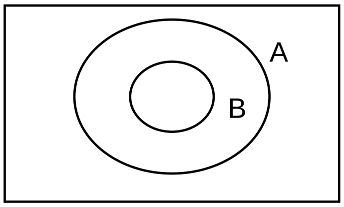
c) $\left(\frac{5}{500}\right)\left(\frac{4}{499}\right)\left(\frac{3}{498}\right) = 4.82\text{x}10^{-7}$

2-67. a) P(gas leak) = (55 + 32)/107 = 0.813

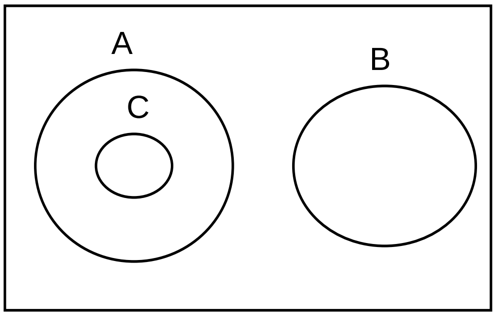
b) P(electric failure|gas leak) = (55/107)/(87/102) = 0.632

c) P(gas leak| electric failure) = (55/107)/(72/107) = 0.764

2-68. No, if $B \subset A$, then P(A/B) = $\frac{P(A \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1$



2-69.



<u>Section 2-5</u>

2-70. a) $P(A \cap B) = P(A|B)P(B) = (0.4)(0.5) = 0.20$

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

2-71.

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

2-72. Let F denote the event that a connector fails.
Let W denote the event that a connector is wet.

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

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

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

2-74. a) P(A) = 0.03
b) P(A') = 0.97
c) P(B|A) = 0.40
d) P(B|A') = 0.05
e) $P(A \cap B) = P(B|A)P(A) = (0.40)(0.03) = 0.012$
f) $P(A \cap B') = P(B'|A)P(A) = (0.60)(0.03) = 0.018$
g) $P(B) = P(B|A)P(A) + P(B|A')P(A') = (0.40)(0.03) + (0.05)(0.97) = 0.0605$

2-75. Let R denote the event that a product exhibits surface roughness. Let N,A, and W denote the events that the blades are new, average, and worn, respectively. Then,

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

2-76. Let B denote the event that a glass breaks.
Let L denote the event that large packaging is used.
P(B)= P(B|L)P(L) + P(B|L')P(L')
= 0.01(0.60) + 0.02(0.40) = 0.014

2-77. Let U denote the event that the user has improperly followed installation instructions.
Let C denote the event that the incoming call is a complaint.
Let P denote the event that the incoming call is a request to purchase more products.
Let R denote the event that the incoming call is a request for information.
a) P(U|C)P(C) = (0.75)(0.03) = 0.0225
b) P(P|R)P(R) = (0.50)(0.25) = 0.125

2-78. a) (0.88)(0.27) = 0.2376
b) (0.12)(0.13+0.52) = 0.0.078

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 third chip selected has excessive shrinkage.

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

2-80. Let A and B denote the events that the first and second chips selected are defective, respectively.
a) P(B) = P(B|A)P(A) + P(B|A')P(A') = (19/99)(20/100) + (20/99)(80/100) = 0.2
b) Let C denote the event that the third chip selected is defective.

$$P(A\cap B\cap C) = P(C|A\cap B)P(A\cap B) = P(C|A\cap B)P(B|A)P(A)$$
$$= \frac{18}{98}\left(\frac{19}{99}\right)\left(\frac{20}{100}\right)$$
$$= 0.00705$$

Section 2-6

2-81. Because P(A|B) $\neq$ P(A), the events are not independent.

2-82. P(A') = 1 - P(A) = 0.7 and P(A'|B) = 1 - P(A|B) = 0.7
Therefore, A' and B are independent events.

2-83. P($A\cap B$) = 70/100, P(A) = 86/100, P(B) = 77/100.
Then, P($A\cap B$) $\neq$ P(A)P(B), so A and B are not independent.

2-84. $P(A \cap B) = 80/100$, $P(A) = 82/100$, $P(B) = 90/100$.
Then, $P(A \cap B) \neq P(A)P(B)$, so A and B are <u>not</u> independent.

2-85. a) $P(A \cap B) = 22/100$, $P(A) = 30/100$, $P(B) = 75/100$, Then $P(A \cap B) \neq P(A)P(B)$, therefore, A and B are not independent.
b) $P(B|A) = P(A \cap B)/P(A) = (22/100)/(30/100) = 0.733$

2-86. If A and B are mutually exclusive, then $P(A \cap B) = 0$ and $P(A)P(B) = 0.04$.
Therefore, A and B are <u>not</u> independent.

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.

2-88. Let A_i denote the event that the ith bit is a one.

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

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

c) The probability of the following sequence is

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

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, $P(B_1 \cup B_2 \cup \ldots \cup B_8) = P(B_1) + P(B_2) + \ldots + P(B_8)$

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

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

2-90. Let A denote the upper devices function. Let B denote the lower devices function.
P(A) = (0.9)(0.8)(0.7) = 0.504
P(B) = (0.95)(0.95)(0.95) = 0.8574
$P(A\cap B)$ = (0.504)(0.8574) = 0.4321
Therefore, the probability that the circuit operates = $P(A\cup B) = P(A) + P(B) - P(A\cap B) = 0.9293$

2-91. [1-(0.1)(0.05)][1-(0.1)(0.05)][1-(0.2)(0.1)] = 0.9702

2-92. Let A_i denote the event that the ith readback is successful. By independence,
$P(A_1' \cap A_2' \cap A_3') = P(A_1')P(A_2')P(A_3') = (0.02)^3 = 0.000008$.

2-93. a) $P(B|A)$ = 4/499 and
$P(B) = P(B|A)P(A) + P(B|A')P(A') = (4/499)(5/500) + (5/499)(495/500) = 5/500$
Therefore, A and B are <u>not</u> independent.
b) A and B are independent.

<u>Section 2-7</u>

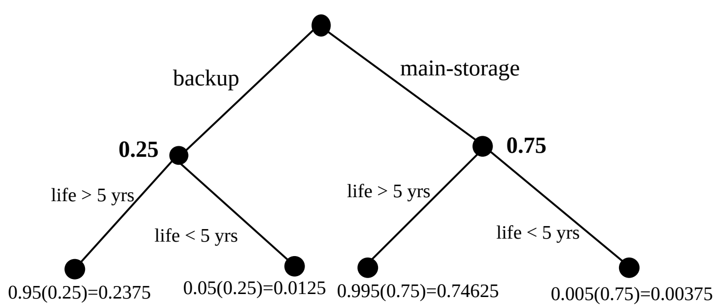
2-94. Because, $P(A|B)\,P(B) = P(A\cap B) = P(B|A)\,P(A)$,

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)} = \frac{0.7(0.2)}{0.5} = 0.28$$

2-95. Let F denote a fraudulent user and let T denote a user that originates calls from two or more metropolitan areas in a day. Then,

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

2-96.



a) P(B) = 0.25
b) $P(A|B) = 0.95$
c) $P(A|B') = 0.995$
d) $P(A\cap B) = P(A|B)P(B) = 0.95(0.25) = 0.2375$
e) $P(A\cap B') = P(A|B')P(B') = 0.995(0.75) = 0.74625$
f) $P(A) = P(A\cap B) + P(A\cap B') = 0.95(0.25) + 0.995(0.75) = 0.98375$
g) 0.95(0.25) + 0.995(0.75) = 0.98375.
h)

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

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)

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

b) Using the result from part a.,

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

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

2-98. a) P(D)=P(D|G)P(G)+P(D|G')P(G')=(.005)(.991)+(.99)(.009)=0.013865

b) $P(G|D')=P(G\cap D')/P(D')=P(D'|G)P(G)/P(D')$=(.995)(.991)/(1-.013865)=0.9999

2-99. a) P(S) = 0.997(0.60) + 0.9995(0.27) + 0.897(0.13) = 0.9847

b) P(Ch|S) =(0.13)(0.897)/0.9847 = 0.1184

Section 2-8

2-100. Continuous: a, c, d, f, h, i; Discrete: b, e, and g

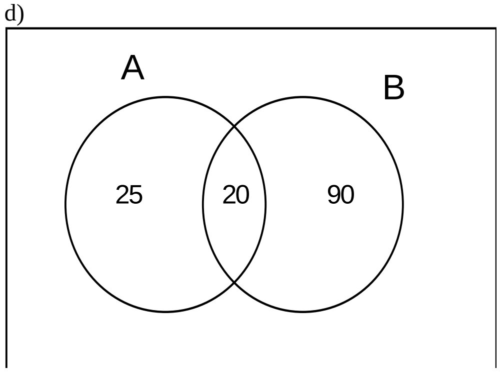
Supplemental Exercises

2-101. Let D_i denote the event that the primary failure mode is type i and let A denote the event that a board passes the test.

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

2-102. a) 20/200 b) 135/200 c) 65/200

d)



2-103. a) $P(A) = 19/100 = 0.19$
b) $P(A \cap B) = 15/100 = 0.15$
c) $P(A \cup B) = (19 + 95 - 15)/100 = 0.99$
d) $P(A' \cap B) = 80/100 = 0.80$
e) $P(A|B) = P(A \cap B)/P(B) = 0.158$

2-104. Let A_i denote the event that the *i*th order is shipped on time.

a) By independence, $P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2)P(A_3) = (0.95)^3 = 0.857$

b) Let

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

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

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

Then, because the B's are mutually exclusive,

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

c) Let

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

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

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

$B_4 = A_1' \cap A_2' \cap A_3'$

Because the B's are mutually exclusive,

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

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) $P(E_1 \cup E_2 \cup E_3) = 1 - P(E_1' \cap E_2' \cap E_3') = 1 - 0 = 1$

2-106. $(0.20)(0.30) + (0.7)(0.9) = 0.69$

2-107. Let A_i denote the event that the *i*th bolt selected is not torqued to the proper limit.

a) Then,

$$P(A_1 \cap A_2 \cap A_3 \cap A_4) = P(A_4 | A_1 \cap A_2 \cap A_3)P(A_1 \cap A_2 \cap A_3)$$
$$= P(A_4 | A_1 \cap A_2 \cap A_3)P(A_3 | A_1 \cap A_2)P(A_2 | A_1)P(A_1)$$
$$= \left(\frac{12}{17}\right)\left(\frac{13}{18}\right)\left(\frac{14}{19}\right)\left(\frac{15}{20}\right) = 0.282$$

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

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

2-108. Let A,B denote the event that the first, second portion of the circuit operates. Then, P(A) = (0.99)(0.99)+0.9-(0.99)(0.99)(0.9) = 0.998
P(B) = 0.9+0.9-(0.9)(0.9) = 0.99 and
P($A \cap B$) = P(A) P(B) = (0.998) (0.99) = 0.988

2-109. A_1 = by telephone, A_2 = website; $P(A_1) = 0.92$, $P(A_2) = 0.95$;
By independence $P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2) = 0.92 + 0.95 - 0.92(0.95) = 0.996$

2-110. P(Possess) = 0.95(0.99) +(0.05)(0.90) = 0.9855

2-111. Let D denote the event that a container is incorrectly filled and let H denote the event that a container is filled under high-speed operation. Then,

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

b) $P(H|D) = \dfrac{P(D|H)P(H)}{P(D)} = \dfrac{0.01(0.30)}{0.0037} = 0.8108$

2-112. a) $P(E' \cap T' \cap D') = (0.995)(0.99)(0.999) = 0.984$
b) $P(E \cup D) = P(E) + P(D) - P(E \cap D) = 0.005995$

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

c) Let A represent the event that the two items NOT inspected are not defective. Then, P(A)=(73/75)(72/74)=0.947.

2-114. The tool fails if any component fails. Let F denote the event that the tool fails. Then, P(F') = 0.99^{10} by independence and P(F) = 1 - 0.99^{10} = 0.0956

2-115. a) (0.3)(0.99)(0.985) + (0.7)(0.98)(0.997) = 0.9764

b) $P(route1|E) = \dfrac{P(E|route1)P(route1)}{P(E)} = \dfrac{0.02485(0.30)}{1-0.9764} = 0.3159$

2-116. a) By independence, $0.15^5 = 7.59 \times 10^{-5}$

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

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

c) As in part b,the probability of 3 of the events is

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

So to get the probability of at least 3, add answer parts a.) and b.) to the above to obtain requested probability. Therefore the answer is

0.0000759 + 0.0022 + 0.0244 = 0.0267

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,

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

2-118. a) If n washers are selected, then the probability they are all less than the target is $\frac{20}{50} \cdot \frac{19}{49} \cdot \ldots \frac{20-n+1}{50-n+1}$.

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

Therefore, the answer is n = 3

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

2-119.

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

$$b)\quad P(A\cap B)=\frac{246}{940}=0.262$$

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

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

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

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

2-120. Let E denote a read error and let S,O,P denote skewed, off-center, and proper alignments, respectively. Then,

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

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

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

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

2-122. a) (0.4)(0.1) + (0.3)(0.1) +(0.2)(0.2) + (0.4)(0.1) = 0.15
b) P(4 or more|provided) = (0.4)(0.1)/0.15 = 0.267

Mind-Expanding Exercises

2-123. Let E denote a read error and let S, O, B, P denote skewed, off-center, both, and proper alignments, respectively.
P(E) = P(E|S)P(S) + P(E|O)P(O) + P(E|B)P(B) + P(E|P)P(P)
= 0.01(0.10) + 0.02(0.05) + 0.06(0.01) + 0.001(0.84) = 0.00344

2-124. Let n denote the number of washers selected.

a) The probability that all are less than the target is 0.4^n, by independence.

n	0.4^n
1	0.4
2	0.16
3	0.064

Therefore, n = 3

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

2-125. Let x denote the number of kits produced.

Revenue at each demand

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

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

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

2-126. Let E denote the probability that none of the bolts are identified as incorrectly torqued. The requested probability is P(E'). Let X denote the number of bolts in the sample that are incorrect. Then,
P(E) = P(E|X=0)P(X=0) + P(E|X=1) P(X=1) + P(E|X=2) P(X=2) + P(E|X=3) P(X=3) + P(E|X=4)P(X=4)
and P(X=0) = (15/20)(14/19)(13/18)(12/17) = 0.2817. The remaining probability for x can be determined from the counting methods in Appendix B-1. Then,

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

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

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

P(X=4) = (5/20)(4/19)(3/18)(2/17) = 0.0010 and P(E|X=0) = 1, P(E|X=1) = 0.05, P(E|X=2) = $0.05^2 = 0.0025$, P(E|X=3) = $0.05^3 = 125 \times 10^{-4}$, P(E|X=4) = $0.05^4 = 6.25 \times 10^{-6}$. Then,

$$P(E) = 1(0.2817) + 0.05(0.4696) + 0.0025(0.2167) + 125 \times 10^{-4}(0.0309)$$
$$+6.25 \times 10^{-6}(0.0010)$$
$$= 0.306$$

and P(E') = 0.694

2-127.

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

2-128. The total sample size is ka + a + kb + b = (k + 1)a + (k +1)b.

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

and

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

Then,

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

<u>Section 2-1.4 on CD</u>

S2-1. From the multiplication rule, the answer is $5 \times 3 \times 4 \times 2 = 120$

S2-2. From the multiplication rule, $3 \times 4 \times 3 = 36$

S2-3. From the multiplication rule, $3 \times 4 \times 3 \times 4 = 144$

S2-4. From equation S2-1, the answer is 10! = 3628800

S2-5. From the multiplication rule and equation S2-1, the answer is 5!5! = 14400

S2-6. From equation S2-3, $\frac{7!}{3!4!} = 35$ sequences are possible

S2-7. a) From equation S2-4, the number of samples of size five is $\binom{140}{5} = \frac{140!}{5!135!} = 416965528$

b) There are 10 ways of selecting one nonconforming chip and there are $\binom{130}{4} = \frac{130!}{4!126!} = 11358880$ ways of selecting four conforming chips. Therefore, the number of samples that contain exactly one nonconforming chip is $10 \times \binom{130}{4} = 113588800$

c) The number of samples that contain at least one nonconforming chip is the total number of samples $\binom{140}{5}$ minus the number of samples that contain no nonconforming chips $\binom{130}{5}$.

That is $\binom{140}{5} - \binom{130}{5} = \frac{140!}{5!135!} - \frac{130!}{5!125!} = 130721752$

S2-8. a) If the chips are of different types, then every arrangement of 5 locations selected from the 12 results in a different layout. Therefore, $P_5^{12} = \frac{12!}{7!} = 95040$ layouts are possible.

b) If the chips are of the same type, then every subset of 5 locations chosen from the 12 results in a different layout. Therefore, $\binom{12}{5} = \frac{12!}{5!7!} = 792$ layouts are possible.

S2-9. a) $\frac{7!}{2!5!} = 21$ sequences are possible.

b) $\frac{7!}{1!1!1!1!1!2!} = 2520$ sequences are possible.

c) 6! = 720 sequences are possible.

S2-10. a) Every arrangement of 7 locations selected from the 12 comprises a different design.

$P_7^{12} = \frac{12!}{5!} = 3991680$ designs are possible.

b) Every subset of 7 locations selected from the 12 comprises a new design. $\frac{12!}{5!7!} = 792$ designs are possible.

c) First the three locations for the first component are selected in $\binom{12}{3} = \frac{12!}{3!9!} = 220$ ways. Then, the four locations for the second component are selected from the nine remaining locations in $\binom{9}{4} = \frac{9!}{4!5!} = 126$ ways. From the multiplication rule, the number of designs is $220 \times 126 = 27720$

S2-11. a) From the multiplication rule, $10^3 = 1000$ prefixes are possible

b) From the multiplication rule, $8 \times 2 \times 10 = 160$ are possible

c) Every arrangement of three digits selected from the 10 digits results in a possible prefix.

$P_3^{10} = \frac{10!}{7!} = 720$ prefixes are possible.

S2-12. a) From the multiplication rule, $2^8 = 256$ bytes are possible

b) From the multiplication rule, $2^7 = 128$ bytes are possible

S2-13. a) The total number of samples possible is $\binom{24}{4} = \frac{24!}{4!20!} = 10626.$ The number of samples in which exactly one tank has high viscosity is $\binom{6}{1}\binom{18}{3} = \frac{6!}{1!5!} \times \frac{18!}{3!15!} = 4896$. Therefore, the probability is

$$\frac{4896}{10626} = 0.461$$

b) The number of samples that contain no tank with high viscosity is $\binom{18}{4} = \frac{18!}{4!14!} = 3060.$ Therefore, the requested probability is $1 - \frac{3060}{10626} = 0.712$.

c) The number of samples that meet the requirements is $\binom{6}{1}\binom{4}{1}\binom{14}{2} = \frac{6!}{1!5!} \times \frac{4!}{1!3!} \times \frac{14!}{2!12!} = 2184$.

Therefore, the probability is $\frac{2184}{10626} = 0.206$

S2-14. a) The total number of samples is $\binom{12}{3} = \frac{12!}{3!9!} = 220$. The number of samples that result in one nonconforming part is $\binom{2}{1}\binom{10}{2} = \frac{2!}{1!1!} \times \frac{10!}{2!8!} = 90.$ Therefore, the requested probability is 90/220 = 0.409.

b) The number of samples with no nonconforming part is $\binom{10}{3} = \frac{10!}{3!7!} = 120.$ The probability of at least one nonconforming part is $1 - \frac{120}{220} = 0.455$.

S2-15. a) The probability that both parts are defective is $\frac{5}{50} \times \frac{4}{49} = 0.0082$

b) The total number of samples is $\binom{50}{2} = \frac{50!}{2!48!} = \frac{50 \times 49}{2}$. The number of samples with two defective parts is $\binom{5}{2} = \frac{5!}{2!3!} = \frac{5 \times 4}{2}$. Therefore, the probability is $\frac{\frac{5 \times 4}{2}}{\frac{50 \times 49}{2}} = \frac{5 \times 4}{50 \times 49} = 0.0082$.

CHAPTER 3

Section 3-1

3-1. The range of X is $\{0,1,2,...,1000\}$

3-2 The range of X is $\{0,1,2,\ldots,50\}$

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

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

3-5. The range of X is $\{1,2,...,491\}$. Because 490 parts are conforming, a nonconforming part must be selected in 491 selections.

3-6 The range of X is $\{0,1,2,\ldots,100\}$. Although the range actually obtained from lots typically might not exceed 10%.

3-7. The range of X is conveniently modeled as all nonnegative integers. That is, the range of X is $\{0,1,2,\ldots\}$

3-8 The range of X is conveniently modeled as all nonnegative integers. That is, the range of X is $\{0,1,2,\ldots\}$

3-9. The range of X is $\{0,1,2,...,15\}$

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

3-11 The range of X is $\{0,1,2,\ldots,10000\}$

3-12 The range of X is $\{0,1,2,...,5000\}$

Section 3-2

3-13.

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

3-14 a) P(X=1.5) = 1/3
b) P(0.5< X < 2.7) = P(X=1.5) + P(X=2) = 1/6 + 1/3 = 1/2
c) P(X > 3) = 0
d) P (0 ≤ X < 2) = P(X=0) + P(X=1.5) = 1/3 + 1/3 = 2/3
e) P(X=0 or X=2) = 1/3 + 1/6 = 1/2

3-15. All probabilities are greater than or equal to zero and sum to one.

a) $P(X \le 2)=1/8 + 2/8 + 2/8 + 2/8 + 1/8 = 1$
b) $P(X > -2) = 2/8 + 2/8 + 2/8 + 1/8 = 7/8$
c) $P(-1 \le X \le 1) = 2/8 + 2/8 + 2/8 =6/8 = 3/4$
d) $P(X \le -1 \text{ or } X=2) = 1/8 + 2/8 +1/8 = 4/8 =1/2$

3-16 All probabilities are greater than or equal to zero and sum to one.
a) $P(X\le 1)=P(X=1)=0.5714$
b) $P(X>1)= 1-P(X=1)=1-0.5714=0.4286$
c) $P(2<X<6)=P(X=3)=0.1429$
d) $P(X\le 1 \text{ or } X>1)= P(X=1)+ P(X=2)+P(X=3)=1$

3-17. Probabilities are nonnegative and sum to one.
a) $P(X = 4) = 9/25$
b) $P(X \le 1) = 1/25 + 3/25 = 4/25$
c) $P(2 \le X < 4) = 5/25 + 7/25 = 12/25$
d) $P(X > -10) = 1$

3-18 Probabilities are nonnegative and sum to one.
a) $P(X = 2) = 3/4(1/4)^2 = 3/64$
b) $P(X \le 2) = 3/4[1+1/4+(1/4)^2] = 63/64$
c) $P(X > 2) = 1 - P(X \le 2) = 1/64$
d) $P(X \ge 1) = 1 - P(X \le 0) = 1 - (3/4) = 1/4$

3-19. P(X = 10 million) = 0.3, P(X = 5 million) = 0.6, P(X = 1 million) = 0.1

3-20 P(X = 50 million) = 0.5, P(X = 25 million) = 0.3, P(X = 10 million) = 0.2

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

3-22 X = number of wafers that pass
$P(X=0) = (0.2)^3 = 0.008$
$P(X=1) = 3(0.2)^2(0.8) = 0.096$
$P(X=2) = 3(0.2)(0.8)^2 = 0.384$
$P(X=3) = (0.8)^3 = 0.512$

3-23 P(X = 15 million) = 0.6, P(X = 5 million) = 0.3, P(X = -0.5 million) = 0.1

3-24 X = number of components that meet specifications
$P(X=0) = (0.05)(0.02) = 0.001$
$P(X=1) = (0.05)(0.98) + (0.95)(0.02) = 0.068$
$P(X=2) = (0.95)(0.98) = 0.931$

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

3-26
$$F(x)=\begin{cases}0, & x<0\\ 1/3 & 0\le x<1.5\\ 2/3 & 1.5\le x<2\\ 5/6 & 2\le x<3\\ 1 & 3\le x\end{cases}$$ where

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

3-27.
$$F(x)=\begin{cases}0, & x<-2\\ 1/8 & -2\le x<-1\\ 3/8 & -1\le x<0\\ 5/8 & 0\le x<1\\ 7/8 & 1\le x<2\\ 1 & 2\le x\end{cases}$$ where

$f_X(-2)=1/8$
$f_X(-1)=2/8$
$f_X(0)=2/8$
$f_X(1)=2/8$
$f_X(2)=1/8$

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

3-28
$$F(x)=\begin{cases}0, & x<0\\ 1/25 & 0\le x<1\\ 4/25 & 1\le x<2\\ 9/25 & 2\le x<3\\ 16/25 & 3\le x<4\\ 1 & 4\le x\end{cases}$$ where

$f_X(0)=1/25$
$f_X(1)=3/25$
$f_X(2)=5/25$
$f_X(3)=7/25$
$f_X(4)=9/25$

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

3-29.
$$F(x)=\begin{cases}0, & x<1\\ 0.1, & 1\le x<5\\ 0.7, & 5\le x<10\\ 1, & 10\le x\end{cases}$$

where P(X = 10 million) = 0.3, P(X = 5 million) = 0.6, P(X = 1 million) = 0.1

3-30

$$F(x) = \begin{cases} 0, & x < 10 \\ 0.2, & 10 \le x < 25 \\ 0.5, & 25 \le x < 50 \\ 1, & 50 \le x \end{cases}$$

where P(X = 50 million) = 0.5, P(X = 25 million) = 0.3, P(X = 10 million) = 0.2

3-31.

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

3-32

$$F(x) = \begin{cases} 0, & x < -0.5 \\ 0.1, & -0.5 \le x < 5 \\ 0.4, & 5 \le x < 15 \\ 1, & 15 \le x \end{cases}$$

where P(X = 15 million) = 0.6, P(X = 5 million) = 0.3, P(X = -0.5 million) = 0.1

3-33. The sum of the probabilities is 1 and all probabilities are greater than or equal to zero;
pmf: $f(1) = 0.5$, $f(3) = 0.5$
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) $P(X>2) = 1 - P(X\le 2) = 0.5$

3-34 The sum of the probabilities is 1 and all probabilities are greater than or equal to zero;
pmf: f(1) = 0.7, f(4) = 0.2, f(7) = 0.1
a) $P(X \le 4) = 0.9$
b) $P(X > 7) = 0$
c) $P(X \le 5) = 0.9$
d) $P(X>4) = 0.1$
e) $P(X\le 2) = 0.7$

3-35. The sum of the probabilities is 1 and all probabilities are greater than or equal to zero;
pmf: f(-10) = 0.25, f(30) = 0.5, f(50) = 0.25
a) $P(X\le 50) = 1$
b) $P(X\le 40) = 0.75$
c) $P(40 \le X \le 60) = P(X=50)=0.25$
d) $P(X<0) = 0.25$
e) $P(0\le X<10) = 0$
f) $P(-10<X<10) = 0$

3-36 The sum of the probabilities is 1 and all probabilities are greater than or equal to zero;
pmf: f1/8) = 0.2, f(1/4) = 0.7, f(3/8) = 0.1
a) P(X≤1/18) = 0
b) P(X≤1/4) = 0.9
c) P(X≤5/16) = 0.9
d) P(X>1/4) = 0.1
e) P(X≤1/2) = 1

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- 38 Mean and Variance

$$\mu = E(X) = 0f(0) + 1.5f(1.5) + 2f(2) + 3f(3)$$
$$= 0(1/3) + 1.5(1/3) + 2(1/6) + 3(1/6) = 1.333$$
$$V(X) = 0^2 f(0) + 1.5^2 f(1) + 2^2 f(2) + 3^2 f(3) - \mu^2$$
$$= 0(1/3) + 2.25(1/3) + 4(1/6) + 9(1/6) - 1.333^2 = 1.139$$

3-39 Determine E(X) and V(X) for random variable in exercise 3-15

$$\mu = E(X) = -2f(-2) - 1f(-1) + 0f(0) + 1f(1) + 2f(2)$$
$$= -2(1/8) - 1(2/8) + 0(2/8) + 1(2/8) + 2(1/8) = 0$$
$$V(X) = -2^2 f(-2) - 1^2 f(-1) + 0^2 f(0) + 1^2 f(1) + 2^2 f(2) - \mu^2$$
$$= 4(1/8) + 1(2/8) + 0(2/8) + 1(2/8) + 4(1/8) - 0^2 = 1.5$$

3-40 Determine E(X) and V(X) for random variable in exercise 3-15

$$\mu = E(X) = 0f(0) + 1f(1) + 2f(2) + 3f(3) + 4f(4)$$
$$= 0(0.04) + 1(0.12) + 2(0.2) + 3(0.28) + 4(0.36) = 2.8$$
$$V(X) = 0^2 f(0) + 1^2 f(1) + 2^2 f(2) + 3^2 f(3) + 4^2 f(4) - \mu^2$$
$$= 0(0.04) + 1(0.12) + 4(0.2) + 9(0.28) + 16(0.36) - 2.8^2 = 1.36$$

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

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

3-42 Mean and variance for exercise 3-20

$$\mu = E(X) = 50f(50) + 25f(25) + 10f(10)$$
$$= 50(0.5) + 25(0.3) + 10(0.2)$$
$$= 34.5 \text{ million}$$

$$V(X) = 50^2 f(50) + 25^2 f(25) + 10^2 f(10) - \mu^2$$
$$= 50^2(0.5) + 25^2(0.3) + 10^2(0.2) - 34.5^2$$
$$= 267.25 \text{ million}^2$$

3-43. Mean and variance for random variable in exercise 3-22

$$\mu = E(X) = 0f(0) + 1f(1) + 2f(2) + 3f(3)$$
$$= 0(0.008) + 1(0.096) + 2(0.384) + 3(0.512) = 2.4$$

$$V(X) = 0^2 f(0) + 1^2 f(1) + 2^2 f(2) + 3^2 f(3) - \mu^2$$
$$= 0^2(0.008) + 1(0.096) + 4(0.384) + 9(0.512) - 2.4^2 = 0.48$$

3-44 Mean and variance for exercise 3-23

$$\mu = E(X) = 15f(15) + 5f(5) - 0.5f(5)$$
$$= 15(0.6) + 5(0.3) - 0.5(0.1)$$
$$= 10.45 \text{ million}$$

$$V(X) = 15^2 f(15) + 5^2 f(5) + (-0.5)^2 f(-0.5) - \mu^2$$
$$= 15^2(0.6) + 5^2(0.3) + (-0.5)^2(0.1) - 10.45^2$$
$$= 33.32 \text{ million}^2$$

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

3-46 E(X) = (0+100)/2 = 50, V(X) = $[(100-0+1)^2-1]/12 = 850$

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

3-48 $$E(X) = \frac{1}{8}\left(\frac{1}{3}\right) + \frac{1}{4}\left(\frac{1}{3}\right) + \frac{3}{8}\left(\frac{1}{3}\right) = \frac{1}{4},$$

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

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

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

3-50 $$E(X) = 2\left(\frac{1}{3}\right)+3\left(\frac{1}{3}\right)+4\left(\frac{1}{3}\right) = 3$$

in 100 codes the expected number of letters is 300

$$V(X) = (2)^2\left(\frac{1}{3}\right)+(3)^2\left(\frac{1}{3}\right)+(4)^2\left(\frac{1}{3}\right)-(3)^2 = \frac{2}{3}$$

in 100 codes the variance is 6666.67

3-51. X = 590 + 0.1Y, Y = 0, 1, 2, ..., 9

E(X) = $590+0.1\left(\frac{0+9}{2}\right) = 590.45$ mm,

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

3-52 The range of Y is 0, 5, 10, ..., 45, E(X) = (0+9)/2 = 4.5

E(Y) = 0(1/10)+5(1/10)+...+45(1/10)
 = 5[0(0.1) +1(0.1)+ ... +9(0.1)]
 = 5E(X)
 = 5(4.5)
 = 22.5

V(X) = 8.25, V(Y) = 5^2(8.25) = 206.25, σ_Y = 14.36

3-53 $$E(cX) = \sum_x cxf(x) = c\sum_x xf(x) = cE(X),$$

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

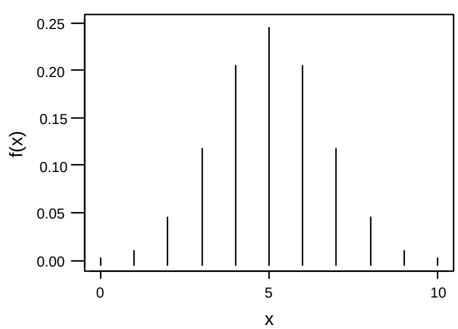
3-54 X is a discrete random variable. X is discrete because it is the number of fields out of 28 that has an error. However, X is not uniform because P(X=0) ≠ P(X=1).

Section 3-6

3-55. A binomial distribution is based on independent trials with two outcomes and a constant probability of success on each trial.

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

3-56



a.) $E(X) = np = 10(0.5) = 5$

b.) Values X=0 and X=10 are the least likely, the extreme values

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

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

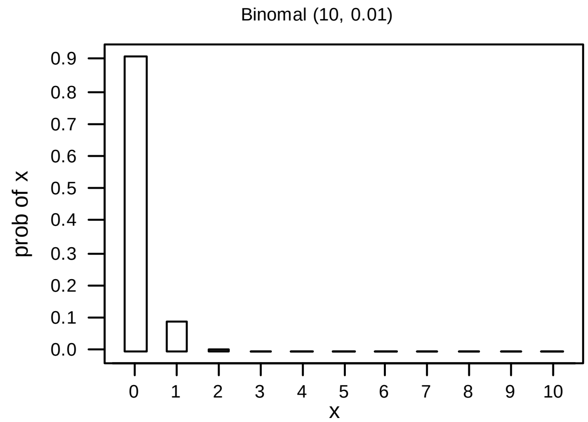
$= 0.5^{10} + 10(0.5)^{10} + 45(0.5)^{10} = 0.0547$

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

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

3-58



a) $E(X) = np = 10(0.01) = 0.1$ The value of X that appears to be most likely is 0. 1.
b) The value of X that appears to be least likely is 10.

3-59. a) $P(X = 5) = \binom{10}{5} 0.01^5 (0.99)^5 = 2.40 \times 10^{-8}$

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

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

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

3-60 n=3 and p=0.5

$$F(x) = \begin{cases} 0 & x < 0 \\ 0.125 & 0 \le x < 1 \\ 0.5 & 1 \le x < 2 \\ 0.875 & 2 \le x < 3 \\ 1 & 3 \le x \end{cases} \quad \text{where} \quad \begin{aligned} f(0) &= \left(\frac{1}{2}\right)^3 = \frac{1}{8} \\ f(1) &= 3\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)^2 = \frac{3}{8} \\ f(2) &= 3\left(\frac{1}{4}\right)^2\left(\frac{3}{4}\right) = \frac{3}{8} \\ f(3) &= \left(\frac{1}{4}\right)^3 = \frac{1}{8} \end{aligned}$$

3-61. n=3 and p=0.25

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

3-62 Let X denote the number of defective circuits. Then, X has a binomial distribution with n = 40 and p = 0.01. Then, P(X = 0) = $\binom{40}{0}0.01^0 0.99^{40}=0.6690$.

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

$$\text{b)}P(X\ge 1)=1-P(X=0)=1-\binom{1000}{1}0.001^1(0.999)^{999}=0.6319$$

$$\text{c)}P(X\le 2)=\binom{1000}{0}0.001^0(0.999)^{1000}+\binom{1000}{1}0.001^1(0.999)^{999}+\binom{1000}{2}0.001^2 0.999^{998}$$
$$=0.9198$$

$$\text{d)}E(X)=1000(0.001)=1$$
$$V(X)=1000(0.001)(0.999)=0.999$$

3-64 Let X denote the number of times the line is occupied. Then, X has a binomial distribution with n = 10 and p = 0.4

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

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

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

3-65. a) n = 50, p = 5/50 = 0.1, since E(X) = 5 = np.

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

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

3-66 E(X) = 20 (0.01) = 0.2
V(X) = 20 (0.01) (0.99) = 0.198
$\mu_X + 3\sigma_X = 0.2 + 3\sqrt{0.198} = 1.53$
a)) X is binomial with n = 20 and p = 0.01

$$P(X > 1.53) = P(X \ge 2) = 1 - P(X \le 1)$$
$$= 1 - \left[\binom{20}{0} 0.01^0 0.99^{20} + \binom{20}{1} 0.01^1 0.99^{19}\right] = 0.0169$$

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

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

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

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

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

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.

$$a) P(X \ge 5) = 1 - P(X \le 4)$$
$$= 1 - \left[\begin{array}{l}\binom{125}{0} 0.1^0 (0.9)^{125} + \binom{125}{1} 0.1^1 (0.9)^{124} + \binom{125}{2} 0.1^2 (0.9)^{123} \\ + \binom{125}{3} 0.1^3 (0.9)^{122} + \binom{125}{4} 0.1^4 (0.9)^{121}\end{array}\right]$$
$$= 0.9961$$

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

3-68 Let X denote the number of defective components among those stocked.

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

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

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

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

$$a) P(X \geq 20) = \binom{25}{20}0.25^{20}(0.75)^5 + \binom{25}{21}0.25^{21}(0.75)^4 + \binom{25}{22}0.25^{22}(0.75)^3$$
$$+ \binom{25}{23}0.25^{23}(0.75)^2 + \binom{25}{24}0.25^{24}(0.75)^1 + \binom{25}{25}0.25^{25}(0.75)^0 = 9.677 \times 10^{-10}$$

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

3-70 Let X denote the number of mornings the light is green.

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

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

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

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

3-72 E(X) = 2.5 = 1/p giving p = 0.4

a.) $P(X = 1) = (1-0.4)^0 0.4 = 0.4$

b.) $P(X = 4) = (1-0.4)^3 0.4 = 0.0864$

c.) $P(X = 5) = (1-0.5)^4 0.5 = 0.05184$

d.) $P(X \leq 3) = P(X = 1) + P(X = 2) + P(X = 3)$
$= (1-0.4)^0 0.4 + (1-0.4)^1 0.4 + (1-0.4)^2 0.4 = 0.7840$

e.) $P(X > 3) = 1 - P(X \leq 3) = 1 - 0.7840 = 0.2160$

3-73. Let X denote the number of trials to obtain the first successful alignment. Then X is a geometric random variable with p = 0.8

a) $P(X = 4) = (1-0.8)^3 0.8 = 0.2^3 0.8 = 0.0064$

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

$$= (1-0.8)^0 0.8 + (1-0.8)^1 0.8 + (1-0.8)^2 0.8 + (1-0.8)^3 0.8$$

$$= 0.8 + 0.2(0.8) + 0.2^2(0.8) + 0.2^3 0.8 = 0.9984$$

c) $P(X \ge 4) = 1 - P(X \le 3) = 1 - [P(X = 1) + P(X = 2) + P(X = 3)]$

$$= 1 - [(1-0.8)^0 0.8 + (1-0.8)^1 0.8 + (1-0.8)^2 0.8]$$

$$= 1 - [0.8 + 0.2(0.8) + 0.2^2(0.8)] = 1 - 0.992 = 0.008$$

3-74 Let X denote the number of people who carry the gene. Then X is a negative binomial random variable with r=2 and p = 0.1

a) $P(X \ge 4) = 1 - P(X < 4) = 1 - [P(X = 2) + P(X = 3)]$

$$= 1 - \left[\binom{1}{1}(1-0.1)^0 0.1^2 + \binom{2}{1}(1-0.1)^1 0.1^2\right] = 1 - (0.01 + 0.018) = 0.972$$

b) $E(X) = r/p = 2/0.1 = 20$

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

$$= 1 - [0.02 + 0.98(0.02) + 0.98^2(0.02) + 0.98^3(0.02)]$$

$$= 1 - 0.0776 = 0.9224$$

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

3-76 Let X denote the number of mornings needed to obtain a green light. Then X is a geometric random variable with p = 0.20.

a) $P(X = 4) = (1-0.2)^3 0.2 = 0.1024$

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

3-77 p = 0.005 , r = 8

a.) $P(X = 8) = 0.005^8 = 3.91x10^{-19}$

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

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

$$\mu = E(Y) = \frac{1}{3.91x10^{-91}} = 2.56x10^{18} \text{ days or } 7.01 \text{ x}10^{15} \text{ years}$$

3-78 Let Y denote the number of samples needed to exceed 1 in Exercise 3-66. Then Y has a geometric distribution with p = 0.0169.

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

b) Y is a geometric random variable with p = 0.1897 from Exercise 3-66.

$P(Y = 10) = (1 - 0.1897)^9(0.1897) = 0.0286$

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

3-79. Let X denote the number of trials to obtain the first success.
a) E(X) = 1/0.2 = 5
b) Because of the lack of memory property, the expected value is still 5.

3-80 Negative binomial random variable: f(x; p, r) = $\binom{x-1}{r-1}(1-p)^{x-r}p^r$.

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

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.

3-82 Let X denote the number of attempts needed to obtain a calibration that conforms to specifications. Then, X is geometric with p = 0.6.
$P(X \le 3) = P(X=1) + P(X=2) + P(X=3) = 0.6 + 0.4(0.6) + 0.4^2(0.6) = 0.936$.

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

3-84 Let X denote the number of transactions until all computers have failed. Then, X is negative binomial random variable with $p = 10^{-8}$ and r = 3.
a) E(X) = 3×10^8
b) V(X) = $[3(1-10^{-8}0]/(10^{-16}) = 3.0 \times 10^{16}$

3-85 Let X denote a geometric random variable with parameter p. Let q = 1-p.

$$E(X)=\sum_{x=1}^{\infty}x(1-p)^{x-1}p$$

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

$$=\frac{p}{(1-q)^2}=\frac{p}{p^2}=\frac{1}{p}$$

$$V(X)=\sum_{x=1}^{\infty}(x-\tfrac{1}{p})^2(1-p)^{x-1}p=\sum_{x=1}^{\infty}\left(px^2-2x+\tfrac{1}{p}\right)(1-p)^{x-1}$$

$$=p\sum_{x=1}^{\infty}x^2q^{x-1}-2\sum_{x=1}^{\infty}xq^{x-1}+\tfrac{1}{p}\sum_{x=1}^{\infty}q^{x-1}$$

$$=p\sum_{x=1}^{\infty}x^2q^{x-1}-\tfrac{2}{p^2}+\tfrac{1}{p^2}$$

$$=p\sum_{x=1}^{\infty}x^2q^{x-1}-\tfrac{1}{p^2}$$

$$=p\tfrac{d}{dq}\left[q+2q^2+3q^3+\ldots\right]-\tfrac{1}{p^2}$$

$$=p\tfrac{d}{dq}\left[q(1+2q+3q^2+\ldots)\right]-\tfrac{1}{p^2}$$

$$=p\tfrac{d}{dq}\left[\tfrac{q}{(1-q)^2}\right]-\tfrac{1}{p^2}=2pq(1-q)^{-3}+p(1-q)^{-2}-\tfrac{1}{p^2}$$

$$=\frac{[2(1-p)+p-1]}{p^2}=\frac{(1-p)}{p^2}=\frac{q}{p^2}$$

Section 3-8

3-86 X has a hypergeometric distribution N=100, n=4, K=20

a.) $P(X=1)=\frac{\binom{20}{1}\binom{80}{3}}{\binom{100}{4}}=\frac{20(82160)}{3921225}=0.4191$

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

c.) $P(X=4)=\frac{\binom{20}{4}\binom{80}{0}}{\binom{100}{4}}=\frac{4845(1)}{3921225}=0.001236$

d.) $E(X)=np=n\frac{K}{N}=4\left(\frac{20}{100}\right)=0.8$

$$V(X)=np(1-p)\left(\frac{N-n}{N-1}\right)=4(0.2)(0.8)\left(\frac{96}{99}\right)=0.6206$$

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)

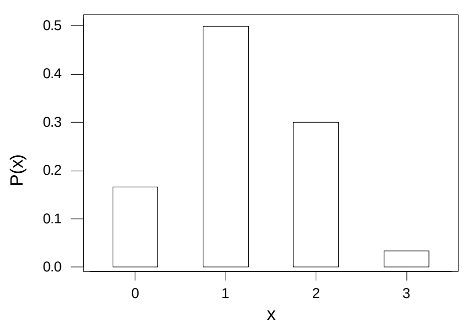
$$P(X\le 2)=P(X=0)+P(X=1)+P(X=2)$$

$$=\frac{\binom{4}{0}\binom{16}{4}}{\binom{20}{4}}+\frac{\binom{4}{1}\binom{16}{3}}{\binom{20}{4}}+\frac{\binom{4}{2}\binom{16}{2}}{\binom{20}{4}}$$

$$=\frac{\left(\frac{16\times15\times14\times13}{24}+\frac{4\times16\times15\times14}{6}+\frac{6\times16\times15}{2}\right)}{\left(\frac{20\times19\times18\times17}{24}\right)}=0.9866$$

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

3-88 N=10, n=3 and K=4



3-89.

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

$$f(0) = \frac{\binom{4}{0}\binom{6}{3}}{\binom{10}{3}} = 0.1667 \quad , \quad f(1) = \frac{\binom{4}{1}\binom{6}{2}}{\binom{10}{3}} = 0.5,$$

$$f(2) = \frac{\binom{4}{2}\binom{6}{1}}{\binom{10}{3}} = 0.3, \quad f(3) = \frac{\binom{4}{3}\binom{6}{0}}{\binom{10}{3}} = 0.0333$$

3-90 Let X denote the number of unacceptable washers in the sample of 10.

$$a.) \quad P(X=0) = \frac{\binom{5}{0}\binom{70}{10}}{\binom{75}{10}} = \frac{\frac{70!}{10!60!}}{\frac{75!}{10!65!}} = \frac{65 \times 64 \times 63 \times 62 \times 61}{75 \times 74 \times 73 \times 72 \times 71} = 0.4786$$

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

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

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

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

$$P(X > 1) = 1 - P(X \le 1) = 1 - [P(X=0) + P(X=1)]$$

$$P(X=0) = \frac{\binom{240}{0}\binom{560}{10}}{\binom{800}{10}} = \frac{\left(\frac{240!}{0!240!}\right)\left(\frac{560!}{10!550!}\right)}{\frac{800!}{10!790!}} = 0.0276$$

$$P(X > 1) = 1 - P(X \le 1) = 1 - [0.0276 + 0.1201] = 0.8523$$

3-92 . Let X denote the number of cards in the sample that are defective.

a)

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

$$P(X = 0) = \frac{\binom{20}{0}\binom{120}{20}}{\binom{140}{20}} = \frac{\frac{120!}{20!100!}}{\frac{140!}{20!120!}} = 0.0356$$

$$P(X \geq 1) = 1 - 0.0356 = 0.9644$$

b)

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

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

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

3-93. Let X denote the number of blades in the sample that are dull.

a)

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

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

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

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

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

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

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

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

3-94 Let X denote the count of the numbers in the state's sample that match those in the player's sample. Then, X has a hypergeometric distribution with N = 40, n = 6, and K = 6.

a) $P(X = 6) = \frac{\binom{6}{6}\binom{34}{0}}{\binom{40}{6}} = \left(\frac{40!}{6!34!}\right)^{-1} = 2.61 \times 10^{-7}$

b) $P(X = 5) = \frac{\binom{6}{5}\binom{34}{1}}{\binom{40}{6}} = \frac{6 \times 34}{\binom{40}{6}} = 5.31 \times 10^{-5}$

c) $P(X = 4) = \frac{\binom{6}{4}\binom{34}{2}}{\binom{40}{6}} = 0.00219$

d) Let Y denote the number of weeks needed to match all six numbers. Then, Y has a geometric distribution with p = $\frac{1}{3{,}838{,}380}$ and E(Y) = 1/p = $3{,}838{,}380$ weeks. This is more than 738 centuries!

3-95. a) For Exercise 3-86, the finite population correction is 96/99.
For Exercise 3-87, the finite population correction is 16/19.
Because the finite population correction for Exercise 3-86 is closer to one, the binomial approximation to the distribution of X should be better in Exercise 3-86.

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

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

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

The results from the binomial approximation are close to the probabilities obtained in Exercise 3-86.

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

3-96 a.) From Exercise 3-92, X is approximately binomial with n = 20 and p = 20/140 = 1/7.

$$P(X \geq 1) = 1 - P(X = 0) = \binom{20}{0}\left(\tfrac{1}{7}\right)^0\left(\tfrac{6}{7}\right)^{20} = 1 - 0.0458 = 0.9542$$

finite population correction is 120/139=0.8633

b) From Exercise 3-92, X is approximately binomial with n = 20 and p = 5/140 =1/28

$$P(X \geq 1) = 1 - P(X = 0) = \binom{20}{0}\left(\tfrac{1}{28}\right)^0\left(\tfrac{27}{28}\right)^{20} = 1 - 0.4832 = 0.5168$$

finite population correction is 120/139=0.8633

Section 3-9

3-97. a) $P(X = 0) = \dfrac{e^{-4}4^0}{0!} = e^{-4} = 0.0183$

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

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

$$= 0.2381$$

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

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

3-98 a) $P(X = 0) = e^{-0.4} = 0.6703$

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

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

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

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

3-100 a) Let X denote the number of calls in one hour. Then, X is a Poisson random variable with $\lambda = 10$.

$$P(X=5)=\frac{e^{-10}10^5}{5!}=0.0378 .$$

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

c) Let Y denote the number of calls in two hours. Then, Y is a Poisson random variable with

$\lambda = 20$. $P(Y=15)=\frac{e^{-20}20^{15}}{15!}=0.0516$

d) Let W denote the number of calls in 30 minutes. Then W is a Poisson random variable with

$\lambda = 5$. $P(W=5)=\frac{e^{-5}5^5}{5!}=0.1755$

3-101. a) Let X denote the number of flaws in one square meter of cloth. Then, X is a Poisson random variable

with $\lambda = 0.1$. $P(X=2)=\frac{e^{-0.1}(0.1)^2}{2!}=0.0045$

b) Let Y denote the number of flaws in 10 square meters of cloth. Then, Y is a Poisson random variable

with $\lambda = 1$. $P(Y=1)=\frac{e^{-1}1^1}{1!}=e^{-1}=0.3679$

c) Let W denote the number of flaws in 20 square meters of cloth. Then, W is a Poisson random variable

with $\lambda = 2$. $P(W=0)=e^{-2}=0.1353$

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

3-102 a) $E(X)=\lambda=0.2$ errors per test area

b.) $P(X \le 2) = e^{-0.2} + \frac{e^{-0.2}0.2}{1!} + \frac{e^{-0.2}(0.2)^2}{2!} = 0.9989$

99.89% of test areas

3-103. a) Let X denote the number of cracks in 5 miles of highway. Then, X is a Poisson random variable with

$\lambda = 10$. $P(X=0)=e^{-10}=4.54\times10^{-5}$

b) Let Y denote the number of cracks in a half mile of highway. Then, Y is a Poisson random variable with

$\lambda = 1$. $P(Y \ge 1) = 1 - P(Y=0) = 1 - e^{-1} = 0.6321$

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

3-104 a.) $E(X) = \lambda = 0.01$ failures per 100 samples. Let Y= the number of failures per day

$E(Y) = E(5X) = 5E(X) = 5\lambda = 0.05$ failures per day.

b.)Let W= the number of failures in 500 participants, now λ=0.05 and $P(W = 0) = e^{-0.05} = 0.9512$

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,

$$P(Y = 10) = \binom{10}{10}(0.6065)^{10}(0.3935)^{0} = 0.0067$$

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.

$$P(W = 0) = \binom{10}{0}(0.3935)^{0}(0.6065)^{10} = 0.0067$$

$$P(W = 1) = \binom{10}{1}(0.3935)^{1}(0.6065)^{9} = 0.0437$$

$$P(W \le 1) = 0.0067 + 0.00437 = 0.00504$$

3-106 a) Let X denote the failures in 8 hours. Then, X has a Poisson distribution with $\lambda = 0.16$.

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

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

<u>Supplemental Exercises</u>

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.

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

3-108 Let X denote the number of calls that are answered in 30 seconds or less. Then, X is a binomial random variable with p = 0.75.

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

b) P(X ≥ 16) = P(X=16) +P(X=17) + P(X=18) + P(X=19) + P(X=20)

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

$$+ \binom{20}{19}(0.75)^{19}(0.25)^{1} + \binom{20}{20}(0.75)^{20}(0.25)^{0} = 0.4148$$

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

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 = 4/3

3-110 Let W denote the number of calls needed to obtain two answers in less than 30 seconds. Then, W has a negative binomial distribution with p = 0.75.

a) P(W=6) = $\binom{5}{1}(0.25)^4(0.75)^2$ = 0.0110

b) E(W) = r/p = 2/0.75 = 8/3

3-111. a) Let X denote the number of messages sent in one hour. $P(X=5)=\dfrac{e^{-5}5^5}{5!}=0.1755$

b) Let Y denote the number of messages sent in 1.5 hours. Then, Y is a Poisson random variable with λ =7.5. $P(Y=10)=\dfrac{e^{-7.5}(7.5)^{10}}{10!}=0.0858$

c) Let W denote the number of messages sent in one-half hour. Then, W is a Poisson random variable with λ = 2.5. $P(W<2)=P(W=0)+P(W=1)=0.2873$

3-112 X is a negative binomial with r=4 and p=0.0001

$E(X)=r/p=4/0.0001=40000$ requests

3-113. X ~ Poisson(λ = 0.01), X ~ Poisson(λ = 1)

$$P(Y\le 3)=e^{-1}+\frac{e^{-1}(1)^1}{1!}+\frac{e^{-1}(1)^2}{2!}+\frac{e^{-1}(1)^3}{3!}=0.9810$$

3-114 Let X denote the number of individuals that recover in one week. Assume the individuals are independent. Then, X is a binomial random variable with n = 20 and p = 0.1. $P(X \ge 4) = 1 - P(X \le 3) = 1 - 0.8670 = 0.1330$.

3-115 a.) P(X=1) = 0 , P(X=2) = 0.0025, P(X=3) = 0.01, P(X=4) = 0.03, P(X=5) = 0.065
P(X=6) = 0.13, P(X=7) = 0.18, P(X=8) = 0.2225, P(X=9) = 0.2, P(X=10) = 0.16

b.) P(X=1) = 0.0025, P(X=1.5) = 0.01, P(X=2) = 0.03, P(X=2.5) = 0.065, P(X=3) = 0.13
P(X=3.5) = 0.18, P(X=4) = 0.2225, P(X=4.5) = 0.2, P(X=5) = 0.16

3-116 Let X denote the number of assemblies needed to obtain 5 defectives. Then, X is a negative binomial random variable with p = 0.01 and r=5.

a) E(X) = r/p = 500.

b) V(X) =(5* 0.99/0.01^2 = 49500 and σ_X = 222.49

3-117. If n assemblies are checked, then let X denote the number of defective assemblies. If $P(X \ge 1) \ge 0.95$, then $P(X=0) \le 0.05$. Now,

P(X=0) = $\binom{n}{0}(0.01)^0(0.99)^n = 99^n$ and $0.99^n \le 0.05$. Therefore,

$$n(\ln(0.99)) \le \ln(0.05)$$

$$n \ge \frac{\ln(0.05)}{\ln(0.95)} = 298.07$$

This would require n = 299.

3-118 Require f(1) + f(2) + f(3) + f(4) = 1. Therefore, $c(1+2+3+4) = 1$. Therefore, $c = 0.1$.

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) $P(X > 2) = 1 - P(X \le 2) = 0.9995$

3-120

$$f_X(0) = (0.1)(0.7) + (0.3)(0.3) = 0.16$$
$$f_X(1) = (0.1)(0.7) + (0.4)(0.3) = 0.19$$
$$f_X(2) = (0.2)(0.7) + (0.2)(0.3) = 0.20$$
$$f_X(3) = (0.4)(0.7) + (0.1)(0.3) = 0.31$$
$$f_X(4) = (0.2)(0.7) + (0)(0.3) = 0.14$$

3-121. a) $P(X \le 3) = 0.2 + 0.4 = 0.6$

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) $V(X) = 2^2(0.2) + 3^2(0.4) + 5^2(0.3) + 8^2(0.1) - (3.9)^2 = 3.09$

3-122

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

3-123. Let X denote the number of bolts in the sample from supplier 1 and let Y denote the number of bolts in the sample from supplier 2. Then, x is a hypergeometric random variable with N = 100, n = 4, and K = 30. Also, Y is a hypergeometric random variable with N = 100, n = 4, and K = 70.

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

$$= \frac{\binom{30}{4}\binom{70}{0}}{\binom{100}{4}} + \frac{\binom{30}{0}\binom{70}{4}}{\binom{100}{4}}$$
$$= 0.2408$$

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

3-124 Let X denote the number of errors in a sector. Then, X is a Poisson random variable with $\lambda = 0.32768$.

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

b) Let Y denote the number of sectors until an error is found. Then, Y is a geometric random variable and
$P = P(X \ge 1) = 1 - P(X=0) = 1 - e^{-0.32768} = 0.2794$
E(Y) = 1/p = 3.58

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 $\lambda = 0.25(8) = 2$.

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

b) Let Y denote the number of orders in 2 weeks. Then, Y is a Poisson random variable with $\lambda = 4$, and $P(Y>2) = 1 - P(Y \le 2) = e^{-4} + (e^{-4}4^1)/1! + (e^{-4}4^2)/2! = 1 - [0.01832+0.07326+0.1465] = 0.7619$.

3-126 a.) hypergeometric random variable with N = 500, n = 5, and K = 125

$$f_X(0) = \frac{\binom{125}{0}\binom{375}{5}}{\binom{500}{5}} = \frac{6.0164E10}{2.5524E11} = 0.2357$$

$$f_X(1) = \frac{\binom{125}{1}\binom{375}{4}}{\binom{500}{5}} = \frac{125(8.10855E8)}{2.5525E11} = 0.3971$$

$$f_X(2) = \frac{\binom{125}{2}\binom{375}{3}}{\binom{500}{5}} = \frac{7750(8718875)}{2.5524E11} = 0.2647$$

$$f_X(3) = \frac{\binom{125}{3}\binom{375}{2}}{\binom{500}{5}} = \frac{317750(70125)}{2.5524E11} = 0.0873$$

$$f_X(4) = \frac{\binom{125}{4}\binom{375}{1}}{\binom{500}{5}} = \frac{9691375(375)}{2.5524E11} = 0.01424$$

$$f_X(5) = \frac{\binom{125}{5}\binom{375}{0}}{\binom{500}{5}} = \frac{2.3453E8}{2.5524E11} = 0.00092$$

b.)

x	0	1	2	3	4	5	6	7	8	9	10
f(x)	0.0546	0.1866	0.2837	0.2528	0.1463	0.0574	0.0155	0.0028	0.0003	0.0000	0.0000

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

which results in p = 0.0739.

3-128 Let t denote an interval of time in hours and let X denote the number of messages that arrive in time t. Then, X is a Poisson random variable with $\lambda = 10t$.
Then, $P(X=0) = 0.9$ and $e^{-10t} = 0.9$, resulting in t = 0.0105 hours = 0.63 seconds

3-129. a) Let X denote the number of flaws in 50 panels. Then, X is a Poisson random variable with $\lambda = 50(0.02) = 1$. $P(X = 0) = e^{-1} = 0.3679$.

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

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

Let V denote the number of panels with 1 or more flaws. Then V is a binomial random variable with n=50 and p=0.0198

$$P(V \le 2) = \binom{50}{0}0.0198^0(.9802)^{50} + \binom{50}{1}0.0198^1(0.9802)^{49}$$
$$+\binom{50}{2}0.0198^2(0.9802)^{48} = 0.9234$$

Mind Expanding Exercises

3-130. Let X follow a hypergeometric distribution with parameters K, n, and N.

To solve this problem, we can find the general expectation:

$$E(X^k) = \sum_{i=0}^{n} i^k P(X = i) = \sum_{i=0}^{n} i^k \frac{\binom{K}{i}\binom{N-K}{n-i}}{\binom{N}{n}}$$

Using the relationships

$$i\binom{K}{i} = K\binom{K-1}{i-1} \quad \text{and} \quad n\binom{N}{n} = N\binom{N-1}{n-1}$$

we can substitute into $E(X^K)$:

$E(X^k) = \sum_{i=0}^{n} i^k P(X=i)$

$$= \sum_{i=0}^{n} i^k \frac{\binom{K}{i}\binom{N-K}{n-i}}{\binom{N}{n}}$$

$$= n\sum_{i=0}^{n} i^{k-1} \frac{K\binom{K-1}{i-1}\binom{N-K}{n-i}}{N\binom{N-1}{n-1}}$$

$$= \frac{nK}{N}\sum_{j=0}^{n-1} (j+1)^{k-1} \frac{\binom{K-1}{j}\binom{N-K}{n-1-j}}{\binom{N-1}{n-1}}$$

$$= \frac{nK}{N} E[(Z+1)^{k-1}]$$

Now, Z is also a hypergeometric random variable with parameters n – 1, N – 1, and K – 1.

To find the mean of X, E(X), set k = 1:

$E(X) = = \frac{nK}{N} E[(Z+1)^{1-1}] = \frac{nK}{N}$

If we let p = K/N, then E(X) = np. In order to find the variance of X using the formula V(X) = $E(X^2) - [E(X)\}^2$, the $E(X^2)$ must be found. Substituting k = 2 into $E(X^k)$ we get

$$E(X^2) = \frac{nK}{N} E[(Z+1)^{2-1}] = \frac{nK}{N} E(Z+1)$$

$$= \frac{nK}{N}\left[E(Z) + E(1)\right] = \frac{nK}{N}\left[\frac{(n-1)(K-1)}{N-1} + 1\right]$$

Therefore, $V(X) = \frac{nK}{N}\left[\frac{(n-1)(K-1)}{N-1} + 1\right] - \left(\frac{nK}{N}\right)^2 = \frac{nK}{N}\left[\frac{(n-1)(K-1)}{N-1} + 1 - \frac{nK}{N}\right]$

If we let p = K/N, the variance reduces to $V(X) = \left(\frac{N-n}{N-1}\right) np(1-p)$

3-131. Show that $\sum_{i=1}^{\infty} (1-p)^{i-1} p = 1$ using an infinite sum.

To begin, $\sum_{i=1}^{\infty} (1-p)^{i-1} p = p\sum_{i=1}^{\infty} (1-p)^{i-1}$, by definition of an infinite sum this can be rewritten as

$$p\sum_{i=1}^{\infty} (1-p)^{i-1} = \frac{p}{1-(1-p)} = \frac{p}{p} = 1$$

3-132

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

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

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

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

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

$$= \frac{\frac{b(b+1)(2b+1)}{6} - \frac{(a-1)a(2a-1)}{6} - (b+a)\left[\frac{b(b+1)-(a-1)a}{2}\right] + \frac{(b-a+1)(b+a)^2}{4}}{b-a+1}$$

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

3-133 Let X denote the number of nonconforming products in the sample. Then, X is approximately binomial with p = 0.01 and n is to be determined.

If $P(X \geq 1) \geq 0.90$, then $P(X = 0) \leq 0.10$.

Now, P(X = 0) = $\binom{n}{0}p^0(1-p)^n = (1-p)^n$. Consequently, $(1-p)^n \leq 0.10$, and

$n \leq \frac{\ln 0.10}{\ln(1-p)} = 229.11$. Therefore, n = 230 is required

3-134 If the lot size is small, 10% of the lot might be insufficient to detect nonconforming product. For example, if the lot size is 10, then a sample of size one has a probability of only 0.2 of detecting a nonconforming product in a lot that is 20% nonconforming.
If the lot size is large, 10% of the lot might be a larger sample size than is practical or necessary. For example, if the lot size is 5000, then a sample of 500 is required. Furthermore, the binomial approximation to the hypergeometric distribution can be used to show the following. If 5% of the lot of size 5000 is nonconforming, then the probability of zero nonconforming product in the sample is approximately 7×10^{-12}. Using a sample of 100, the same probability is still only 0.0059. The sample of size 500 might be much larger than is needed.

3-135 Let X denote the number of panels with flaws. Then, X is a binomial random variable with n =100 and p is the probability of one or more flaws in a panel. That is, p = $1-e^{-0.1}$ = 0.095.

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

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

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

$$=0.034$$

3-136 Let X denote the number of rolls produced.

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

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

3-137 Let X denote the number of acceptable components. Then, X has a binomial distribution with p = 0.98 and n is to be determined such that $P(X \ge 100) \ge 0.95$.

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

Therefore, 105 components are needed.

CHAPTER 4

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

4-2. a) $P(x < X) = \int_x^{\infty} e^{-x} dx = (-e^{-x})\Big|_x^{\infty} = e^{-x} = 0.10$.

Then, x = –ln(0.10) = 2.3

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

Then, x = –ln(0.9) = 0.1054

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

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

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

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

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

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

Then, x = 4 – ln(0.10) = 6.303

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

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

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

d) P(X < –2) = 0

e) P(X < 0 or X > –0.5) = 1

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

Then, x = 0.9655

4-6. a) $P(X > 3000) = \int_{3000}^{\infty} \frac{e^{\frac{-x}{1000}}}{1000}dx = -e^{\frac{-x}{1000}}\Big|_{3000}^{\infty} = e^{-3} = 0.05$

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

c) $P(X < 1000) = \int_{0}^{1000} \frac{e^{\frac{-x}{1000}}}{1000}dx = -e^{\frac{-x}{1000}}\Big|_{0}^{1000} = 1 - e^{-1} = 0.6321$

d) $P(X < x) = \int_0^x \frac{e^{\frac{-x}{1000}}}{1000}dx = -e^{\frac{-x}{1000}}\Big|_0^x = 1 - e^{-x/1000} = 0.10$.

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

4-7 a) $P(X > 50) = \int_{50}^{50.25} 2.0dx = 2x\Big|_{50}^{50.25} = 0.5$

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

Then, 2x = 99.6 and x = 49.8.

4-8. a) $P(X < 74.8) = \int_{74.6}^{74.8} 1.25dx = 1.25x\Big|_{74.6}^{74.8} = 0.25$

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

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

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 2.3 < x < 2.8 and all rods will meet specifications.

4-10. Because the integral $\int_{x_1}^{x_2} f(x)dx$ is not changed whether or not any of the endpoints x_1 and x_2 are included in the integral, all the probabilities listed are equal.

Section 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) $P(X > 6) = 1 - F_X(6) = 0$

4-12. a) $P(X < 1.8) = P(X \le 1.8) = F_X(1.8)$ because X is a continuous random variable.
Then, $F_X(1.8) = 0.25(1.8) + 0.5 = 0.95$
b) $P(X > -1.5) = 1 - P(X \le -1.5) = 1 - .125 = 0.875$
c) P(X < -2) = 0
d) $P(-1 < X < 1) = P(-1 < X \le 1) = F_X(1) - F_X(-1) = .75 - .25 = 0.50$

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

4-14. Now, $f(x) = x/8$ for 3 < x < 5 and $F_X(x) = \int_3^x \frac{x}{8} dx = \frac{x^2}{16}\Big|_3^x = \frac{x^2 - 9}{16}$

for 0 < x. Then, $F_X(x) = \begin{cases} 0, x < 3 \\ \frac{x^2 - 9}{16}, 3 \le x < 5 \\ 1, x \ge 5 \end{cases}$

4-15. Now, $f(x) = e^{-(4-x)}$ for 4 < x and $F_X(x) = \int_4^x e^{-(4-x)} dx = -e^{-(4-x)}\Big|_4^x = 1 - e^{-(4-x)}$

for 4 < x.

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

4-16. Now, $f(x) = \frac{e^{-x/1000}}{1000}$ for 0 < x and

$F_X(x) = 1/1000 \int_4^x e^{-x/1000} dx = -e^{-x/1000)}\Big|_4^x = 1 - e^{-x/1000}$

for 0 < x.

Then, $F_X(x) = \begin{cases} 0, & x \le 0 \\ 1 - e^{-x/1000}, & x > 0 \end{cases}$

P(X>3000) = 1- P(X $\le$ 3000) = 1- F(3000) = $e^{-3000/1000}$ = 0.5

4-17. Now, f(x) = 1.25 for 74.6 < x < 75.4 and $F(x) = \int_{74.6}^x 1.25 dx = 1.25x - 93.25$

for 74.6 < x < 75.4. Then,

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

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

4-18 $f(x) = 2e^{-2x}, \quad x > 0$

4-19. $$f\ (x) = \begin{cases} 0.2, & 0 < x < 4 \\ 0.04, & 4 \le x < 9 \end{cases}$$

4-20. $$f_X(x) = \begin{cases} 0.25, & -2 < x < 1 \\ 0.5, & 1 \le x < 1.5 \end{cases}$$

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

4-22. $$E(X) = \int_0^4 0.25x dx = 0.25 \frac{x^2}{2}\Big|_0^4 = 2$$

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

4-23. $$E(X) = \int_0^4 0.125x^2 dx = 0.125 \frac{x^3}{3}\Big|_0^4 = 2.6667$$

$$V(X) = \int_0^4 0.125x(x - \tfrac{8}{3})^2 dx = 0.125\int_0^4 (x^3 - \tfrac{16}{3}x^2 + \tfrac{64}{9}x)dx$$

$$= 0.125(\tfrac{x^4}{4} - \tfrac{16}{3}\tfrac{x^3}{3} + \tfrac{64}{9} \cdot \tfrac{1}{2}x^2)\Big|_0^4 = 0.88889$$

4-24. $E(X)=\int_{-1}^{1}1.5x^3dx=1.5\frac{x^4}{4}\Big|_{-1}^{1}=0$

$$V(X)=\int_{-1}^{1}1.5x^3(x-0)^2dx=1.5\int_{-1}^{1}x^4dx$$

$$=1.5\frac{x^5}{5}\Big|_{-1}^{1}=0.6$$

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

$$V(X)=\int_{3}^{5}(x-4.083)^2\frac{x}{8}dx=\int_{3}^{5}\left(\frac{x^3}{8}-\frac{8.166x^2}{8}+\frac{16.6709x}{8}\right)dx$$

$$=\frac{1}{8}\left(\frac{x^4}{4}-\frac{8.166x^3}{3}+\frac{16.6709x^2}{2}\right)\Big|_{3}^{5}=0.3264$$

4-26. $E(X)=\int_{49.75}^{50.25}2xdx=x^2\Big|_{49.75}^{50.25}=50$

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

$$=2(\tfrac{x^3}{3}-100\tfrac{x^2}{2}+2500x)\Big|_{49.75}^{50.25}$$

$$=0.0208$$

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

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

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

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

4-28. $E(X)=\int_{1}^{\infty}x2x^{-3}dx=-2x^{-1}\Big|_{1}^{\infty}=2$

4-29. a) $E(X) = \int_5^\infty x10e^{-10(x-5)}dx$.

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

$$E(X) = -xe^{-10(x-5)}\Big|_5^\infty + \int_5^\infty e^{-10(x-5)}dx = 5 - \frac{e^{-10(x-5)}}{10}\Bigg|_5^\infty = 5.1$$

Now, $V(X) = \int_5^\infty (x-5.1)^2 10e^{-10(x-5)}dx$. Using the integration by parts with

$u = (x-5.1)^2$ and

$dv = 10e^{-10(x-5)}$, we obtain $V(X) = -(x-5.1)^2 e^{-10(x-5)}\Big|_5^\infty + 2\int_5^\infty (x-5.1)e^{-10(x-5)}dx$.

From the definition of E(X) the integral above is recognized to equal 0.

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

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

4-30. a)

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

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

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

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

$$P(1195 < X < 1205) = P(1200 < X < 1205) = \int_{1200}^{1205} 0.1dx = 0.1x\Big|_{1200}^{1205} = 0.5$$

Section 4-5

4-31. a) E(X) = (5.5+1.5)/2 = 3.5,

$$V(X) = \frac{(5.5-1.5)^2}{12} = 1.333, \text{ and } \sigma_x = \sqrt{1.333} = 1.155.$$

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

4-32. a) E(X) = (-1+1)/2 = 0,

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

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

Therefore, x should equal 0.90.

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

4-34. a) The distribution of X is $f(x)$ = 10 for 0.95 < x < 1.05. Now,

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

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

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

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

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

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

b) $P(X < 2) = \int_{1.5}^{2} \frac{1}{(2.2-1.5)} dx = \int_{1.5}^{2} (1/0.7) dx = (1/0.7)x\Big|_{1.5}^{2} = (1/0.7)(0.5) = 0.7143$

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

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

4-36 $f(x) = 0.04$ for 50< x <75

a) $P(X > 70) = \int_{70}^{75} 0.04 dx = 0.2x\Big|_{70}^{75} = 0.2$

b) $P(X < 60) = \int_{50}^{60} 0.04 dx = 0.04x\Big|_{50}^{60} = 0.4$

c) $E(X) = \frac{75+50}{2} = 62.5$ seconds

$$V(X) = \frac{(75-50)^2}{12} = 52.0833 \text{ seconds}^2$$

4-37. a) The distribution of X is $f(x)$ = 100 for 0.2050 < x < 0.2150. Therefore,

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

b) $P(X > 0.2125) = 1 - F(0.2125) = 1 - [100(0.2125) - 20.50] = 0.25$

c) If P(X > x)=0.10, then 1 – F(X) = 0.10 and F(X) = 0.90.
Therefore, 100x - 20.50 = 0.90 and x = 0.2140.

d) E(X) = (0.2050 + 0.2150)/2 = 0.2100 μm and

V(X) = $\frac{(0.2150 - 0.2050)^2}{12} = 8.33 \times 10^{-6}\ \mu\text{m}^2$

4-38. a) $P(X > 35) = \int_{35}^{40} 0.1dx = 0.1x \Big|_{35}^{40} = 0.5$

b) P(X > x)=0.90 and $P(X > x) = \int_{x}^{40} 0.1dt = 0.1(40 - x)$.

Now, 0.1(40-x) = 0.90 and x = 31

c) E(X) = (30 + 40)/2 = 35 and V(X) = $\frac{(40-30)^2}{12} = 8.33$

<u>Section 4-6</u>

4-39. a) P(Z<1.32) = 0.90658
b) P(Z<3.0) = 0.99865
c) P(Z>1.45) = 1 – 0.92647 = 0.07353
d) P(Z > –2.15) = p(Z < 2.15) = 0.98422
e) P(–2.34 < Z < 1.76) = P(Z<1.76) – P(Z > 2.34) = 0.95116

4-40. a) P(–1 < Z < 1) = P(Z < 1) – P(Z > 1)
= 0.84134 – (1 – 0.84134)
= 0.68268
b) P(–2 < Z < 2) = P(Z < 2) – [1 – P(Z < 2)]
= 0.9545
c) P(–3 < Z < 3) = P(Z < 3) – [1 – P(Z < 3)]
= 0.9973
d) P(Z > 3) = 1 – P(Z < 3)
= 0.00135
e) P(0 < Z < 1) = P(Z < 1) – P(Z < 0)
= 0.84134 – 0.5
= 0.34134

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

4-42. a) Because of the symmetry of the normal distribution, the area in each tail of the distribution must equal 0.025. Therefore the value in Table II that corresponds to 0.975 is 1.96. Thus, z = 1.96.
b) Find the value in Table II corresponding to 0.995. z = 2.58.
c) Find the value in Table II corresponding to 0.84. z = 1.0
d) Find the value in Table II corresponding to 0.99865. z = 3.0.

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)
= 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)
= 0.00132

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.

4-44. a) P(X > x) = $P\left(Z > \frac{x-10}{2}\right)$ = 0.5. Therefore, $\frac{x-10}{2}$ = 0 and x = 10.

b) P(X > x) = $P\left(Z > \frac{x-10}{2}\right) = 1 - P\left(Z < \frac{x-10}{2}\right)$
= 0.95.

Therefore, $P\left(Z < \frac{x-10}{2}\right)$ = 0.05 and $\frac{x-10}{2}$ = –1.64. Consequently, x = 6.72.

c) P(x < X < 10) = $P\left(\frac{x-10}{2} < Z < 0\right) = P(Z < 0) - P\left(Z < \frac{x-10}{2}\right)$
$= 0.5 - P\left(Z < \frac{x-10}{2}\right) = 0.2.$

Therefore, $P\left(Z < \frac{x-10}{2}\right)$ = 0.3 and $\frac{x-10}{2}$ = –0.52. Consequently, x = 8.96.

d) P(10 – x < X < 10 + x) = P(–x/2 < Z < x/2)= 0.95.
Therefore, x/2 = 1.96 and x = 3.92

e) P(10 – x < X < 10 + x) = P(–x/2 < Z < x/2) = 0.99.
Therefore, x/2 = 2.58 and x = 5.16

4-45. a) P(X < 11) = $P\left(Z < \frac{11-5}{4}\right)$

= P(Z < 1.5)
= 0.93319

b) P(X > 0) = $P\left(Z > \frac{0-5}{4}\right)$

= P(Z > –1.25)
= 1 – P(Z < –1.25)
= 0.89435

c) P(3 < X < 7) = $P\left(\frac{3-5}{4} < Z < \frac{7-5}{4}\right)$

= P(–0.5 < Z < 0.5)
= P(Z < 0.5) – P(Z < –0.5)
= 0.38292

d) P(–2 < X < 9) = $P\left(\frac{-2-5}{4} < Z < \frac{9-5}{4}\right)$

= P(–1.75 < Z < 1)
= P(Z < 1) – P(Z < –1.75)]
= 0.80128

e) P(2 < X < 8) = $P\left(\frac{2-5}{4} < Z < \frac{8-5}{4}\right)$

=P(–0.75 < Z < 0.75)
= P(Z < 0.75) – P(Z < –0.75)
= 0.54674

4-46. a) P(X > x) = $P\left(Z > \frac{x-5}{4}\right)$ = 0.5.

Therefore, x = 5.

b) P(X > x) = $P\left(Z > \frac{x-5}{4}\right)$ = 0.95.

Therefore, $P\left(Z < \frac{x-5}{4}\right)$ = 0.05

Therefore, $\frac{x-5}{4}$ = –1.64, and x = –1.56.

c) P(x < X < 9) = $P\left(\frac{x-5}{4} < Z < 1\right)$ = 0.2.

Therefore, P(Z < 1) – P(Z < $\frac{x-5}{4}$)= 0.2 where P(Z < 1) = 0.84134.

Thus P(Z < $\frac{x-5}{4}$) = 0.64134. Consequently, $\frac{x-5}{4}$ = 0.36 and x = 6.44.

d) P(3 < X < x) = $P\left(\frac{3-5}{4} < Z < \frac{x-5}{4}\right)$ = 0.95.

Therefore, $P\left(Z < \frac{x-5}{4}\right)$ – P(Z < –0.5) = 0.95 and $P\left(Z < \frac{x-5}{4}\right)$ – 0.30854 = 0.95.

Consequently,

$P\left(Z < \frac{x-5}{4}\right)$ = 1.25854. Because a probability can not be greater than one, there is no solution for x. In fact, P(3 < X) = P(–0.5 < Z) = 0.69146. Therefore, even if x is set to infinity the probability requested cannot equal 0.95.

e) P(5 – x < X <5 + x) = $P\left(\frac{5-x-5}{4} < Z < \frac{5+x-5}{4}\right)$

$$= P\left(\frac{-x}{4} < Z < \frac{x}{4}\right) = 0.99$$

Therefore, x/4 = 2.58 and x = 10.32.

4-47. a) P(X < 6250) = $P\left(Z < \frac{6250-6000}{100}\right)$

= P(Z < 2.5)
= 0.99379

b) P(5800 < X < 5900) = $P\left(\frac{5800-6000}{100} < Z < \frac{5900-6000}{100}\right)$

=P(–2 < Z < –1)
= P(Z <– 1) – P(Z < –2)
= 0.13591

c) P(X > x) = $P\left(Z > \frac{x-6000}{100}\right)$ = 0.95.

Therefore, $\frac{x-6000}{100}$ = –1.65 and x = 5835.

4-48. a) P(X < 40) = $P\left(Z < \frac{40-35}{2}\right)$

= P(Z < 2.5)
= 0.99379

b) P(X < 30) = $P\left(Z < \frac{30-35}{2}\right)$

= P(Z < –2.5)
= 0.00621

0.621% are scrapped

4-49. a) P(X > 0.62) = $P\left(Z > \frac{0.62-0.5}{0.05}\right)$

= P(Z > 2.4)
= 1 – P(Z <2.4)
= 0.0082

b) P(0.47 < X < 0.63) = $P\left(\frac{0.47-0.5}{0.05} < Z < \frac{0.63-0.5}{0.05}\right)$

= P(–0.6 < Z < 2.6)
= P(Z < 2.6) – P(Z < –0.6)
= 0.99534 – 0.27425
= 0.72109

c) P(X < x) = $P\left(Z < \frac{x-0.5}{0.05}\right)$ = 0.90.

Therefore, $\frac{x-0.5}{0.05}$ = 1.28 and x = 0.564.

4-50. a) P(X < 12) = P(Z < $\frac{12-12.4}{0.1}$) = P(Z < –4) ≅ 0

b) P(X < 12.1) = $P\left(Z < \frac{12.1-12.4}{0.1}\right)$ = P(Z < –3) = 0.00135

and

P(X > 12.6) = $P\left(Z > \frac{12.6-12.4}{0.1}\right)$ = P(Z > 2) = 0.02275.

Therefore, the proportion of cans scrapped is 0.00135 + 0.02275 = 0.0241, or 2.41%

c) P(12.4 – x < X < 12.4 + x) = 0.99.

Therefore, $P\left(-\frac{x}{0.1} < Z < \frac{x}{0.1}\right)$ = 0.99

Consequently, $P\left(Z < \frac{x}{0.1}\right)$ = 0.995 and x = 0.1(2.58) = 0.258.

The limits are (12.142, 12.658).

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

b) P(X > 65) = $P\left(Z > \frac{65-60}{5}\right)$ = P(Z >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 = 72

4-52. a) If P(X > 12) = 0.999, then $P\left(Z > \frac{12-\mu}{0.1}\right)$ = 0.999.

Therefore, $\frac{12-\mu}{0.1}$ = –3.09 and μ = 12.309.

b) If P(X > 12) = 0.999, then $P\left(Z > \frac{12-\mu}{0.05}\right)$ = 0.999.

Therefore, $\frac{12-\mu}{0.05}$ = -3.09 and μ = 12.1545.

4-53. a) P(X > 0.5) = $P\left(Z > \frac{0.5-0.4}{0.05}\right)$

= P(Z > 2)
= 1 – 0.97725
= 0.02275

b) P(0.4 < X < 0.5) = $P\left(\frac{0.4-0.4}{0.05} < Z < \frac{0.5-0.4}{0.05}\right)$

= P(0 < Z < 2)
= P(Z < 2) – P(Z < 0)
= 0.47725

c) P(X > x) = 0.90, then $P\left(Z > \frac{x-0.4}{0.05}\right)$ = 0.90.

Therefore, $\frac{x-0.4}{0.05}$ = –1.28 and x = 0.336.

4-54 a) P(X > 70) = $P\left(Z > \frac{70-60}{4}\right)$

$= 1 - P(Z < 2.5)$

$= 1 - 0.99379 = 0.00621$

b) P(X < 58) = $P\left(Z < \frac{58-60}{4}\right)$

$= P(Z < -0.5)$

$= 0.308538$

c) $1{,}000{,}000 \text{ bytes} * 8 \text{ bits/byte} = 8{,}000{,}000 \text{ bits}$

$$\frac{8{,}000{,}000 \text{ bits}}{60{,}000 \text{ bits/sec}} = 133.33 \text{ seconds}$$

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 μ = 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%

4-56. a) P(89.7 < X < 90.3) = $P\left(\frac{89.7 - 90}{0.1} < Z < \frac{90.3 - 90}{0.1}\right)$

= P(–3 < Z < 3)

= 0.9973.

P(X=10) = $(0.9973)^{10}$ = 0.9733

b) Let *Y* represent the number of cases out of the sample of 10 that are between 89.7 and 90.3 ml. Then Y follows a binomial distribution with *n*=10 and *p*=0.9973. Thus, E(*Y*)= 9.973 or 10.

4-57. a) P(50 < X < 80) = $P\left(\frac{50 - 100}{20} < Z < \frac{80 - 100}{20}\right)$

= P(–2.5 < Z < -1)

= P(Z < –1) – P(Z < –2.5)

= 0.15245.

b) P(X > x) = 0.10. Therefore, $P\left(Z > \frac{x - 100}{20}\right)$ = 0.10 and $\frac{x-100}{20}$ = 1.28.

Therefore, x = 126. hours

4-58. a) P(X < 5000) = $P\left(Z < \frac{5000-7000}{600}\right)$

= P(Z < –3.33) = 0.00043.

b) P(X > x) = 0.95. Therefore, $P\left(Z > \frac{x-7000}{600}\right)$ = 0.95 and $\frac{x-7000}{600}$ = –1.64.

Consequently, x = 6016.

c) P(X > 7000) = $P\left(Z > \frac{7000-7000}{600}\right) = P(Z > 0) = 0.5$

P(three lasers operating after 7000 hours) = $(1/2)^3$ =1/8

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 σ = 0.000214.

4-60. a) P(X > 13) = $P\left(Z > \frac{13-12}{0.5}\right)$

= P(Z > 2)
= 0.02275

b) If P(X < 13) = 0.999, then $P\left(Z < \frac{13-12}{\sigma}\right)$ = 0.999.

Therefore, 1/ σ = 3.09 and σ = 1/3.09 = 0.324.

c) If P(X < 13) = 0.999, then $P\left(Z < \frac{13-\mu}{0.5}\right)$ = 0.999.

Therefore, $\frac{13-\mu}{0.5}$ = 3.09 and μ = 11.455

Section 4-7

4-61. a) E(X) = 200(0.4) = 80, V(X) = 200(0.4)(0.6) = 48 and $\sigma_X = \sqrt{48}$.

Then, $P(X \le 70) \cong P\left(Z \le \frac{70-80}{\sqrt{48}}\right) = P(Z \le -1.44) = 0.074934$

b)

$$P(70 < X \le 90) \cong P\left(\frac{70-80}{\sqrt{48}} < Z \le \frac{90-80}{\sqrt{48}}\right) = P(-1.44 < Z \le 1.44)$$
$$= 0.925066 - 0.074934 = 0.85013$$

4-62. a)

$$P(X < 4) = \binom{100}{0}0.1^0 0.9^{100} + \binom{100}{1}0.1^1 0.9^{99}$$
$$+ \binom{100}{2}0.1^2 0.9^{98} + \binom{100}{3}0.1^3 0.9^{97} = 0.0078$$

b) E(X) = 10, V(X) = 100(0.1)(0.9) = 9 and $\sigma_X = 3$.

Then, $P(X < 4) \cong P(Z < \frac{4-10}{3}) = P(Z < -2) = 0.023$

c) $P(8 < X < 12) \cong P(\frac{8-10}{3} < Z < \frac{12-10}{3}) = P(-0.67 < Z < 0.67) = 0.497$

4-63. Let X denote the number of defective chips in the lot.
Then, E(X) = 1000(0.02) = 20, V(X) = 1000(0.02)(0.98) = 19.6.

a) $P(X > 25) \cong P\left(Z > \frac{25-20}{\sqrt{19.6}}\right) = P(Z > 1.13) = 1 - P(Z \le 1.13) = 0.129$

b) $P(20 < X < 30) \cong P(0 < Z < \frac{10}{\sqrt{19.6}}) = P(0 < Z < 2.26)$

$$= P(Z \le 2.26) - P(Z < 0) = 0.98809 - 0.5 = 0.488$$

4-64. Let X denote the number of defective electrical connectors.
Then, E(X) = 25(100/1000) = 2.5, V(X) = 25(0.1)(0.9) = 2.25.

a) $P(X=0)=0.9^{25}=0.0718$

b) $P(X \le 0) \cong P(Z < \frac{0-2.5}{\sqrt{2.25}}) = P(Z < -1.67) = 0.047$

The approximation is smaller than the binomial. It is not satisfactory since $np<5$.

c) Then, E(X) = 25(100/500) = 5, V(X) = 25(0.2)(0.8) = 4.

$P(X=0)=0.8^{25}=0.00377$

$P(X \le 0) \cong P(Z < \frac{0-5}{\sqrt{4}}) = P(Z < -2.5) = 0.006$

Normal approximation is now closer to the binomial; however, it is still not satisfactory since $np = 5$ is not > 5.

4-65. Let X denote the number of original components that fail during the useful life of the product. Then, X is a binomial random variable with p = 0.001 and n = 5000. Also, E(X) = 5000 (0.001) = 5 and V(X) = 5000(0.001)(0.999) = 4.995.

$$P(X \geq 10) \cong P\left(Z \geq \frac{10-5}{\sqrt{4.995}}\right) = P(Z \geq 2.24) = 1 - P(Z < 2.24) = 1 - 0.987 = 0.013.$$

4-66. Let X denote the number of particles in 10 cm^2 of dust. Then, X is a Poisson random variable with $\lambda = 10(1000) = 10{,}000$. Also, E(X) = $\lambda = 10{,}000$ = V(X)

$$P(X > 10{,}000) \cong P(Z > \tfrac{10{,}000-10{,}000}{\sqrt{10{,}000}}) = P(Z > 0) = 0.5$$

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

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

4-68. Let X denote the number of particles in 10 cm^2 of dust. Then, X is a Poisson random variable with $\lambda = 10{,}000$. Also, E(X) = $\lambda = 10{,}000$ =V(X)

a)

$$P(X \geq 20{,}000) \cong P(Z \geq \tfrac{20{,}000-10{,}000}{\sqrt{10{,}000}}) = P(Z \geq 100)$$

$$= 1 - P(Z < 100) = 1 - 1 = 0$$

b.) $P(X < 9{,}900) \cong P(Z < \tfrac{9{,}900-10{,}000}{\sqrt{10{,}000}}) = P(Z < -1) = 0.1587$

c.) If P($X > x$) = 0.01, then $P\left(Z > \frac{x-10{,}000}{\sqrt{10{,}000}}\right) = 0.01$.

Therefore, $\frac{x-10{,}000}{100} = 2.33$ and $x = 10{,}233$

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) = λ = 10,000 and V(X) = 10,000

a)

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

$$= 1 - 0.9772 = 0.0228$$

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

4-69 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

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

4-70 E(X) = 1000(0.2) = 200 and V(X) = 1000(0.2)(0.8) = 160

a) $P(X \geq 225) \cong 1 - P(Z \leq \frac{225-200}{\sqrt{160}}) = 1 - P(Z \leq 1.98) = 1 - 0.97615 = 0.02385$

b) $P(175 \leq X \leq 225) \cong P(\frac{175-200}{\sqrt{160}} \leq Z \leq \frac{225-200}{\sqrt{160}}) = P(-1.98 \leq Z \leq 1.98)$
$= 0.97615 - 0.02385 = 0.9523$

c) If P($X > x$) = 0.01, then $P\left(Z > \frac{x-200}{\sqrt{160}} \right) = 0.01$.

Therefore, $\frac{x-200}{\sqrt{160}} = 2.33$ and $x = 229.5$

4-71 X is the number of minor errors on a test pattern of 1000 pages of text. X is a Poisson random variable with a mean of 0.4 per page

a.) The number of errors per page is a random variable because it will be different for each page.

b.) $P(X = 0) = \frac{e^{-0.4} 0.4^0}{0!} = 0.670$

$P(X \geq 1) = 1 - P(X = 0) = 1 - 0.670 = 0.330$

The mean number of pages with one or more errors is 1000(0.330)=330 pages

c.) Let Y be the number of pages with errors.

$$P(Y > 350) \cong P\left(Z \geq \frac{350 - 330}{\sqrt{1000(0.330)(0.670)}} \right) = P(Z \geq 1.35) = 1 - P(Z < 1.35)$$
$$= 1 - 0.9115 = 0.0885$$

Section 4-9

4-72. a) $P(X \le 0) = \int_0^0 \lambda e^{-\lambda x} dx = 0$

b) $P(X \ge 2) = \int_2^\infty 2e^{-2x} dx = -e^{-2x}\Big|_2^\infty = e^{-4} = 0.0183$

c) $P(X \le 1) = \int_0^1 2e^{-2x} dx = -e^{-2x}\Big|_0^1 = 1 - e^{-2} = 0.8647$

d) $P(1 < X < 2) = \int_1^2 2e^{-2x} dx = -e^{-2x}\Big|_1^2 = e^{-2} - e^{-4} = 0.1170$

e) $P(X \le x) = \int_0^x 2e^{-2t} dt = -e^{-2t}\Big|_0^x = 1 - e^{-2x} = 0.05$ and x = 0.0256

4-73. If E(X) = 10, then $\lambda = 0.1$.

a) $P(X > 10) = \int_{10}^\infty 0.1e^{-0.1x} dx = -e^{-0.1x}\Big|_{10}^\infty = e^{-1} = 0.3679$

b) $P(X > 20) = -e^{-0.1x}\Big|_{20}^\infty = e^{-2} = 0.1353$

c) $P(X > 30) = -e^{-0.1x}\Big|_{30}^\infty = e^{-3} = 0.0498$

d) $P(X < x) = \int_0^x 0.1e^{-0.1t} dt = -e^{-0.1t}\Big|_0^x = 1 - e^{-0.1x} = 0.95$ and x = 29.96.

4-74. Let X denote the time until the first count. Then, X is an exponential random variable with $\lambda = 2$ counts per minute.

a) $P(X > 0.5) = \int_{0.5}^\infty 2e^{-2x} dx = -e^{-2x}\Big|_{0.5}^\infty = e^{-1} = 0.3679$

b) $P(X < \frac{10}{60}) = \int_0^{1/6} 2e^{-2x} dx = -e^{-2x}\Big|_0^{1/6} = 1 - e^{-1/3} = 0.2835$

c) $P(1 < X < 2) = -e^{-2x}\Big|_1^2 = e^{-2} - e^{-4} = 0.1170$

4-75. a) E(X) = $1/\lambda = 1/3 = 0.333$ minutes

b) V(X) = $1/\lambda^2 = 1/3^2 = 0.111$, $\sigma = 0.3333$

c) $P(X < x) = \int_0^x 3e^{-3t} dt = -e^{-3t}\Big|_0^x = 1 - e^{-3x} = 0.95$, x = 0.9986

4-76. The time to failure (in hours) for a laser in a cytometry machine is modeled by an exponential distribution with 0.00004.

a) $P(X > 20{,}000) = \int_{20000}^{\infty} 0.00004 e^{-.0.00004x} dx = -e^{-0.00004x} \Big|_{20000}^{\infty} = e^{-0.8} = 0.4493$

b) $P(X < 30{,}000) = \int_{30000}^{\infty} 0.00004 e^{-.0.00004x} dx = -e^{-0.00004x} \Big|_{0}^{30000} = 1 - e^{-1.2} = 0.6988$

c)

$$P(20{,}000 < X < 30{,}000) = \int_{20000}^{30000} 0.00004 e^{-.0.00004x} dx$$

$$= -e^{-0.00004x} \Big|_{20000}^{30000} = e^{-0.8} - e^{-1.2} = 0.1481$$

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.

4-78. Let X be the life of regulator. Then, X is an exponential random variable with $\lambda = 1/E(X) = 1/6$

a) Because the Poisson process from which the exponential distribution is derived is memoryless, this probability is

$$P(X < 6) = \int_{0}^{6} \tfrac{1}{6} e^{-x/6} dx = -e^{-x/6} \Big|_{0}^{6} = 1 - e^{-1} = 0.6321$$

b) Because the failure times are memoryless, the mean time until the next failure is E(X) = 6 years.

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

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

4-80. Let X denote the time until a message is received. Then, X is an exponential random variable and $\lambda = 1/E(X) = 1/2$.

a) P(X > 2) = $\int_{2}^{\infty} \frac{1}{2}e^{-x/2}dx = -e^{-x/2}\Big|_{2}^{\infty} = e^{-1} = 0.3679$

b) The same as part a.

c) E(X) = 2 hours.

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

4-82. a) P(X > x) = $\int_{x}^{\infty} 0.1e^{-0.1t}dt = -e^{-0.1t}\Big|_{x}^{\infty} = e^{-0.1x} = 0.1$ and x = 23.03 minutes.

b) P(X < x) = 0.9 implies that P(X > x) = 0.1. Therefore, this answer is the same as part a.

c) P(X < x) = $-e^{-0.1t}\Big|_{0}^{x} = 1 - e^{-0.1x} = 0.5$ and x = 6.93 minutes.

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.

4-84. a) $P(12 < X < 15) = \int_{12}^{15} 0.2e^{-0.2x}dx = -e^{-0.2x}\Big|_{12}^{15} = e^{-2.4} - e^{-3} = 0.0409$

b) P(X > 5) = $-e^{-0.2x}\Big|_{5}^{\infty} = e^{-1} = 0.3679$. By independence of the intervals in a Poisson process, the answer is $0.3679^2 = 0.1353$. Alternatively, the answer is P(X > 10) = $e^{-2} = 0.1353$. The probability does depend on whether or not the lengths of highway are consecutive.

c) By the memoryless property, this answer is P(X > 10) = 0.1353 from part b.

4-85. Let X denote the lifetime of an assembly. Then, X is an exponential random variable with $\lambda = 1/E(X) = 1/400$ failures per hour.

a) P(X < 100) = $\int_{0}^{100} \frac{1}{400}e^{-x/400}dx = -e^{-x/400}\Big|_{0}^{100} = 1 - e^{-0.25} = 0.2212$

b) P(X > 500) = $-e^{-x/400}\Big|_{500}^{\infty} = e^{-5/4} = 0.2865$

c) From the memoryless property of the exponential, this answer is the same as part a., P(X < 100) = 0.2212.

4-86. a) Let U denote the number of assemblies out of 10 that fail before 100 hours. By the memoryless property of a Poisson process, U has a binomial distribution with n = 10 and p =0.2212 (from Exercise 4-85a). Then,

$P(U \geq 1) = 1 - P(U = 0) = 1 - \binom{10}{0}0.2212^0(1-0.2212)^{10} = 0.9179$

b) Let V denote the number of assemblies out of 10 that fail before 800 hours. Then, V is a binomial random variable with n = 10 and p = P(X < 800), where X denotes the lifetime of an assembly.

Now, P(X < 800) = $\int_{0}^{800} \frac{1}{400}e^{-x/400}dx = -e^{-x/400}\Big|_{0}^{800} = 1 - e^{-2} = 0.8647$.

Therefore, P(V = 10) = $\binom{10}{10}0.8647^{10}(1-0.8647)^0 = 0.2337$.

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.

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

4-88. Let Y denote the number of arrivals in one hour. If the time between arrivals is exponential, then the count of arrivals is a Poisson random variable and $\lambda = 1$ arrival per hour.

P(Y > 3) = $1 - P(Y \leq 3) = 1 - \left[\frac{e^{-1}1^0}{0!} + \frac{e^{-1}1^1}{1!} + \frac{e^{-1}1^2}{2!} + \frac{e^{-1}1^3}{3!}\right] = 0.01899$

4-89. a) From Exercise 4-88, P(Y > 3) = 0.01899. Let W denote the number of one-hour intervals out of 30 that contain more than 3 arrivals. By the memoryless property of a Poisson process, W is a binomial random variable with n = 30 and p = 0.01899.

P(W = 0) = $\binom{30}{0}0.01899^{0}(1-0.01899)^{30} = 0.5626$

b) Let X denote the time between arrivals. Then, X is an exponential random variable with $\lambda = 1$ arrivals per hour. P(X > x) = 0.1 and $P(X > x) = \int_x^\infty 1e^{-1t}dt = -e^{-1t}\Big|_x^\infty = e^{-1x} = 0.1$.

Therefore, x = 2.3 hours.

4-90. Let X denote the number of calls in 30 minutes. Because the time between calls is an exponential random variable, X is a Poisson random variable with $\lambda = 1/E(X) = 0.1$ calls per minute = 3 calls per 30 minutes.

a) P(X > 3) = $1 - P(X \le 3) = 1 - \left[\frac{e^{-3}3^0}{0!} + \frac{e^{-3}3^1}{1!} + \frac{e^{-3}3^2}{2!} + \frac{e^{-3}3^3}{3!}\right] = 0.3528$

b) P(X = 0) = $\frac{e^{-3}3^0}{0!} = 0.04979$

c) Let Y denote the time between calls in minutes. Then, $P(Y \ge x) = 0.01$ and

$P(Y \ge x) = \int_x^\infty 0.1e^{-0.1t}dt = -e^{-0.1t}\Big|_x^\infty = e^{-0.1x}$. Therefore, $e^{-0.1x} = 0.01$ and x = 46.05 minutes.

4-91. a) From Exercise 4-90, P(Y > 120) = $\int_{120}^\infty 0.1e^{-0.1y}dy = -e^{-0.1y}\Big|_{120}^\infty = e^{-12} = 6.14\times10^{-6}$.

b) Because the calls are a Poisson process, the number of calls in disjoint intervals are independent. From Exercise 4-90 part b., the probability of no calls in one-half hour is $e^{-3} = 0.04979$. Therefore, the answer is $\left[e^{-3}\right]^4 = e^{-12} = 6.14\times10^{-6}$. Alternatively, the answer is the probability of no calls in two hours. From part a. of this exercise, this is e^{-12}.

c) Because a Poisson process is memoryless, probabilities do not depend on whether or not intervals are consecutive. Therefore, parts a. and b. have the same answer.

4-92. a) $P(X > \theta) = \int_\theta^\infty \frac{1}{\theta}e^{-x/\theta}dx = -e^{-x/\theta}\Big|_\theta^\infty = e^{-1} = 0.3679$

b) $P(X > 2\theta) = -e^{-x/\theta}\Big|_{2\theta}^\infty = e^{-2} = 0.1353$

c) $P(X > 3\theta) = -e^{-x/\theta}\Big|_{3\theta}^\infty = e^{-3} = 0.0498$

d) The results do not depend on θ.

4-93. X is an exponential random variable with $\lambda = 0.2$ flaws per meter.

a) E(X) = $1/\lambda = 5$ meters.

b) P(X > 10) = $\int_{10}^{\infty} 0.2e^{-0.2x}dx = -e^{-0.2x}\Big|_{10}^{\infty} = e^{-2} = 0.1353$

c) No, see Exercise 4-91 part c.

d) P(X < x) = 0.90. Then, P(X < x) = $-e^{-0.2t}\Big|_{0}^{x} = 1 - e^{-0.2x}$. Therefore, $1 - e^{-0.2x} = 0.9$ and x = 11.51.

4-94. P(X > 8) = $\int_{8}^{\infty} 0.2e^{-0.2x}dx = -e^{-8/5} = 0.2019$

The distance between successive flaws is either less than 8 meters or not. The distances are independent and P(X > 8) = 0.2019. Let Y denote the number of flaws until the distance exceeds 8 meters. Then, Y is a geometric random variable with p = 0.2019.

a) P(Y = 5) = $(1 - 0.2019)^4 0.2019 = 0.0819$.

b) E(Y) = 1/0.2019 = 4.95.

4-95. $E(X) = \int_{0}^{\infty} x\lambda e^{-\lambda x}dx.$ Use integration by parts with u = x and dv = $\lambda e^{-\lambda x}$.

Then, $E(X) = -xe^{-\lambda x}\Big|_{0}^{\infty} + \int_{0}^{\infty} e^{-\lambda x}dx = \frac{-e^{-\lambda x}}{\lambda}\Big|_{0}^{\infty} = 1/\lambda$

V(X) = $\int_{0}^{\infty} (x - \frac{1}{\lambda})^2 \lambda e^{-\lambda x}dx.$ Use integration by parts with $u = (x - \frac{1}{\lambda})^2$ and $dv = \lambda e^{-\lambda x}$. Then,

$$V(X) = -(x - \tfrac{1}{\lambda})^2 e^{-\lambda x}\Big|_{0}^{\infty} + 2\int_{0}^{\infty} (x - \tfrac{1}{\lambda})e^{-\lambda x}dx = (\tfrac{1}{\lambda})^2 + \tfrac{2}{\lambda}\int_{0}^{\infty} (x - \tfrac{1}{\lambda})\lambda e^{-\lambda x}dx$$

The last integral is seen to be zero from the definition of E(X). Therefore, V(X) = $\left(\frac{1}{\lambda}\right)^2$.

<u>Section 4-10</u>

4-96 a) The time until the tenth call is an Erlang random variable with $\lambda = 5$ calls per minute and r = 10.

b) E(X) = 10/5 = 2 minutes. V(X) = 10/25 = 0.4 minutes.

c) Because a Poisson process is memoryless, the mean time is 1/5=0.2 minutes or 12 seconds

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

Let W denote the number of one minute intervals out of 10 that contain more than 2 calls. Because the calls are a Poisson process, W is a binomial random variable with n = 10 and p = 0.8754.

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

4-98. Let X denote the pounds of material to obtain 15 particles. Then, X has an Erlang distribution with r = 15 and $\lambda = 0.01$.

a) E(X) = $\frac{r}{\lambda} = \frac{15}{0.01} = 1500$ pounds.

b) V(X) = $\frac{15}{0.01^2} = 150{,}000$ and $\sigma_X = \sqrt{150{,}000} = 387.3$ pounds.

4-99 Let X denote the time between failures of a laser. X is exponential with a mean of 25,000.

a.) Expected time until the second failure $E(X) = r/\lambda = 2/0.00004 = 50{,}000$ hours

b.) N=no of failures in 50000 hours

$$E(N) = \frac{50000}{25000} = 2$$

$$P(N \le 2) = \sum_{k=0}^{2} \frac{e^{-2}(2)^k}{k!} = 0.6767$$

4-100 Let X denote the time until 5 messages arrive at a node. Then, X has an Erlang distribution with r = 5 and $\lambda = 30$ messages per minute.

a) E(X) = 5/30 = 1/6 minute = 10 seconds.

b)V(X) = $\frac{5}{30^2} = 1/180$ minute^2 = 1/3 second and $\sigma_X = 0.0745$ minute = 4.472 seconds.

c) Let Y denote the number of messages that arrive in 10 seconds. Then, Y is a Poisson random variable with $\lambda = 30$ messages per minute = 5 messages per 10 seconds.

$$P(Y \ge 5) = 1 - P(Y \le 4) = 1 - \left[\frac{e^{-5}5^0}{0!} + \frac{e^{-5}5^1}{1!} + \frac{e^{-5}5^2}{2!} + \frac{e^{-5}5^3}{3!} + \frac{e^{-5}5^4}{4!}\right]$$

$$= 0.5595$$

d) Let Y denote the number of messages that arrive in 5 seconds. Then, Y is a Poisson random variable with $\lambda = 2.5$ messages per 5 seconds.

$$P(Y \ge 5) = 1 - P(Y \le 4) = 1 - 0.8912 = 0.1088$$

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) = $\frac{r}{\lambda} = 5\times 10^{5}$ bits.

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

c) Let Y denote the number of errors in 10^5 bits. Then, Y is a Poisson random variable with

$\lambda = 1/10^5 = 10^{-5}$ error per bit = 1 error per 10^5 bits.

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

4-102 $\lambda = 1/20 = 0.05 \quad r = 100$

a) $E(X) = r/\lambda = 100/.05 = 5$ minutes

b) 4 min - 2.5 min=1.5 min

c)Let Y be the number of calls before 15 seconds $\lambda = 0.25 * 20 = 5$

$$P(Y > 3) = 1 - P(X \leq 2) = 1 - \left[\frac{e^{-5}5^0}{0!} + \frac{e^{-5}5^1}{1!} + \frac{e^{-5}5^2}{2!}\right] = 1 - .1247 = 0.8753$$

4-103. a) Let X denote the number of customers that arrive in 10 minutes. Then, X is a Poisson random variable with $\lambda = 0.2$ arrivals per minute = 2 arrivals per 10 minutes.

$$P(X > 3) = 1 - P(X \leq 3) = 1 - \left[\frac{e^{-2}2^0}{0!} + \frac{e^{-2}2^1}{1!} + \frac{e^{-2}2^2}{2!} + \frac{e^{-2}2^3}{3!}\right] = 0.1429$$

b) Let Y denote the number of customers that arrive in 15 minutes. Then, Y is a Poisson random variable with $\lambda = 3$ arrivals per 15 minutes.

$$P(X \geq 5) = 1 - P(X \leq 4) = 1 - \left[\frac{e^{-3}3^0}{0!} + \frac{e^{-3}3^1}{1!} + \frac{e^{-3}3^2}{2!} + \frac{e^{-3}3^3}{3!} + \frac{e^{-3}3^4}{4!}\right] = 0.1847$$

4-104. Let X denote the time in days until the fourth problem. Then, X has an Erlang distribution with r = 4 and $\lambda = 1/30$ problem per day.

a) E(X) = $\frac{4}{30^{-1}} = 120$ days.

b) Let Y denote the number of problems in 120 days. Then, Y is a Poisson random variable with $\lambda = 4$ problems per 120 days.

$$P(Y < 4) = \left[\frac{e^{-4}4^0}{0!} + \frac{e^{-4}4^1}{1!} + \frac{e^{-4}4^2}{2!} + \frac{e^{-4}4^3}{3!}\right] = 0.4335$$

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

4-106 $\Gamma(r) = \int_0^\infty x^{r-1}e^{-x}dx$. Use integration by parts with $u = x^{r-1}$ and $dv = e^{-x}$. Then,

$$\Gamma(r) = -x^{r-1}e^{-x}\Big|_0^\infty + (r-1)\int_0^\infty x^{r-2}e^{-x}dx = (r-1)\Gamma(r-1).$$

4-107 $\int_0^\infty f(x;\lambda,r)dx = \int_0^\infty \frac{\lambda^r x^{r-1} e^{-\lambda x}}{\Gamma(r)} dx$. Let y = λx, then the integral is $\int_0^\infty \frac{\lambda\, y^{r-1} e^{-y}}{\Gamma(r)} \frac{dy}{\lambda}$. From the definition of $\Gamma(r)$, this integral is recognized to equal 1.

4-108. If X is a chi-square random variable, then X is a special case of a gamma random variable. Now, E(X) = $\frac{r}{\lambda} = \frac{(7/2)}{(1/2)} = 7$ and $V(X) = \frac{r}{\lambda^2} = \frac{(7/2)}{(1/2)^2} = 14$.

Section 4-11

4-109. β=0.2 and δ=100 hours

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

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

4-110. a) $P(X < 10000) = F_X(10000) = 1 - e^{-100^{0.2}} = 1 - e^{-2.512} = 0.9189$

b) $P(X > 5000) = 1 - F_X(5000) = e^{-50^{0.2}} = 0.1123$

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.

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

4-112 a.) $E(X) = \delta\Gamma(1+\frac{1}{\beta}) = 900\Gamma(1+1/3) = 900\Gamma(4/3) = 900(0.89298). = 803.68$ hours

b.)

$$V(X) = \delta^2\Gamma(1+\tfrac{2}{\beta}) - \delta^2\left[\Gamma(1+\tfrac{2}{\beta})\right]^2 = 900^2\Gamma(1+\tfrac{2}{3}) - 900^2\left[\Gamma(1+\tfrac{1}{3})\right]^2$$

$$= 900^2(0.90274) - 900^2(0.89298)^2 = 85314.64 \text{ hours}^2$$

c.) $P(X < 500) = F_X(500) = 1 - e^{-\left(\frac{500}{900}\right)^3} = 0.1576$

4-113. Let *X* denote the lifetime.

a) $E(X) = \delta\Gamma(1+\frac{1}{0.5}) = \delta\Gamma(3) = 2\delta = 600$. Then $\delta = 300$. Now,

P(X > 500) = $e^{-(\frac{500}{300})^{0.5}} = 0.2750$

b) P(X < 400) = $1 - e^{-(\frac{400}{300})^{0.5}} = 0.6848$

4-114 Let X denote the lifetime

a) $E(X) = 700\Gamma(1+\frac{1}{2}) = 620.4$

b)

$$V(X) = 700^2\Gamma(2) - 700^2[\Gamma(1.5)]^2$$
$$= 700^2(1) - 700^2(0.25\pi) = 105{,}154.9$$

c) P(X > 620.4) = $e^{-\left(\frac{620.4}{700}\right)^2} = 0.4559$

4-115 a.) β=2, δ=500

$$E(X) = 500\Gamma(1+\tfrac{1}{2}) = 500\Gamma(1.5)$$
$$= 500(0.5)\Gamma(0.5) = 250\sqrt{\pi} = 443.11$$

= 443.11 hours

b.) $V(X) = 500^2\Gamma(1+ \ 1) - 500^2[\Gamma(1+ \ \frac{1}{2})]^2$

$$= 500^2\Gamma(2) - 500^2[\Gamma(1.5)]^2 = 53650.5$$

c.) P(X < 250) = F(250)=$1 - e^{-(\frac{250}{500})^{25}} = 1 - 0.7788 = 0.2212$

4-116 If X is a Weibull random variable with β=1 and δ=1000, the distribution of X is the exponential distribution with λ=.001.

$$f(x) = \left(\frac{1}{1000}\right)\left(\frac{x}{1000}\right)^0 e^{-\left(\frac{x}{1000}\right)^1} \text{ for } x > 0$$
$$= 0.001e^{-0.001x} \text{ for } x > 0$$

The mean of X is E(X) = 1/λ = 1000.

Section 4-11

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

$$V(X) = e^{2\theta + \omega^2}(e^{\omega^2} - 1) = e^{10+9}(e^9 - 1) = e^{19}(e^9 - 1) = 1.45x10^{12}$$

4-118 a.) X is a lognormal distribution with θ=-2 and ω^2=9

$$P(500 < X < 1000) = P(500 < e^W < 1000) = P(\ln(500) < W < \ln(1000))$$

$$= \Phi\left(\frac{\ln(1000) + 2}{3}\right) - \Phi\left(\frac{\ln(500) + 2}{3}\right) = \Phi(2.97) - \Phi(2.74) = 0.0016$$

b.) $P(X < x) = P(e^W \le x) = P(W < \ln(x)) = \Phi\left(\frac{\ln(x) + 2}{3}\right) = 0.1$

$$\frac{\ln(x) + 2}{3} = -1.28 \qquad x = e^{-1.28(3)-2} = 0.0029$$

c.) $\mu = E(X) = e^{\theta + \omega^2/2} = e^{-2+9/2} = e^{2.5} = 12.1825$

$$V(X) = e^{2\theta + \omega^2}(e^{\omega^2} - 1) = e^{-4+9}(e^9 - 1) = e^5(e^9 - 1) = 1{,}202{,}455.87$$

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.

4-120 X is a lognormal distribution with θ=0.5 and ω^2=1

a)

$$P(X > 10) = P(e^W > 10) = P(W > \ln(10)) = 1-\Phi\left(\frac{\ln(10)-0.5}{1}\right)$$

$$= 1-\Phi(1.80) = 1-0.96407 = 0.03593$$

b.) $$P(X \le x) = P(e^W \le x) = P(W < \ln(x)) = \Phi\left(\frac{\ln(x)-0.5}{1}\right) = 0.50$$

$$\frac{\ln(x)-0.5}{1} = 0 \quad x = e^{0(1)+0.5} = 1.65 \text{ seconds}$$

c.) $$\mu = E(X) = e^{\theta+\omega^2/2} = e^{0.5+1/2} = e^1 = 2.7183$$

$$V(X) = e^{2\theta+\omega^2}(e^{\omega^2}-1) = e^{1+1}(e^1-1) = e^2(e^1-1) = 12.6965$$

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

$$100 = e^{\theta+\omega^2/2} \qquad 85000 = e^{2\theta+\omega^2}(e^{\omega^2}-1)$$

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

$$85000 = 10000\,(y-1)$$

$$y = 9.5$$

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

$$\theta = \ln(32.444) = 3.48 \text{ and } \omega^2 = \ln(9.5) = 2.25$$

4-122 a.) Find the values of θ and ω^2 given that E(X) = 10000 and σ= 20,000

$$10000 = e^{\theta+\omega^2/2} \qquad 20000^2 = e^{2\theta+\omega^2}(e^{\omega^2}-1)$$

let $x = e^{\theta}$ and $y = e^{\omega^2}$ then (1) $10000 = x\sqrt{y}$ and

(2) $20000^2 = x^2y(y-1) = x^2y^2 - x^2y$

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

$$20000^2 = 10000^2\,(y-1)$$

$$y = 5$$

Substitute y into (1) and solve for x $x = \dfrac{10000}{\sqrt{5}} = 4472.1360$

$$\theta = \ln(4472.1360) = 8.4056 \text{ and } \omega^2 = \ln(5) = 1.6094$$

b.)

$$P(X > 10000) = P(e^W > 10000) = P(W > \ln(10000)) = 1 - \Phi\left(\frac{\ln(10000) - 8.4056}{1.2686}\right)$$

$$= 1 - \Phi(0.63) = 1 - 0.7357 = 0.2643$$

c.) $P(X > x) = P(e^W > x) = P(W > \ln(x)) = \Phi\left(\dfrac{\ln(x) - 8.4056}{1.2686}\right) = 0.1$

$$\frac{\ln(x) - 8.4056}{1.2686} = -1.28 \quad x = e^{-1.280\,(1.2686)+8.4056} = 881.65 \text{ hours}$$

4-123 Let X ~N(μ, σ^2), then Y = e^X follows a lognormal distribution with mean μ and variance σ^2. By definition, $F_Y(y) = P(Y \le y) = P(e^X < y) = P(X < \log y) = F_X(\log y) = \Phi\left(\frac{\log y - \mu}{\sigma}\right)$.

Since Y = e^X and X ~ N(μ, σ^2), we can show that $f_Y(Y) = \frac{1}{y} f_X(\log y)$

Finally, $f_Y(y) = \dfrac{\partial F_Y(y)}{\partial y} = \dfrac{\partial F_X(\log y)}{\partial y} = \dfrac{1}{y} f_X(\log y) = \dfrac{1}{y} \cdot \dfrac{1}{\sigma\sqrt{2\pi}} e^{-\left(\frac{\log y - \mu}{2\sigma}\right)^2}$.

Supplemental Exercises

4-124 a) $P(X < 2.5) = \int_2^{2.5} (0.5x - 1)dx = \left(0.5\frac{x^2}{2} - x\right)\Big|_2^{2.5} = 0.0625$

b) $P(X > 3) = \int_3^4 (0.5x - 1)dx = 0.5\frac{x^2}{2} - x\Big|_3^4 = 0.75$

c) $P(2.5 < X < 3.5) = \int_{2.5}^{3.5} (0.5x - 1)dx = 0.5\frac{x^2}{2} - x\Big|_{2.5}^{3.5} = 0.5$

4-125 $F(x) = \int_2^x (0.5t - 1)dt = 0.5\frac{t^2}{2} - t\Big|_2^x = \frac{x^2}{4} - x + 1$. Then,

$$F(x) = \begin{cases} 0, & x < 2 \\ \frac{x^2}{4} - x + 1, & 2 \le x < 4 \\ 1, & 4 \le x \end{cases}$$

4-126 $E(X) = \int_2^4 x(0.5x - 1)dx = 0.5\frac{x^3}{3} - \frac{x^2}{2}\Big|_2^4 = \frac{32}{3} - 8 - (\frac{4}{3} - 2) = \frac{10}{3}$

$$V(X) = \int_2^4 (x - \tfrac{10}{3})^2 (0.5x - 1)dx = \int_2^4 (x^2 - \tfrac{20}{3}x + \tfrac{100}{9})(0.5x - 1)dx$$

$$= \int_2^4 (0.5x^3 - \tfrac{13}{3}x^2 + \tfrac{110}{9}x - \tfrac{100}{9})dx = \tfrac{x^4}{8} - \tfrac{13}{9}x^3 + \tfrac{55}{9}x^2 - \tfrac{100}{9}x\Big|_2^4$$

$$= 0.2222$$

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.

4-128 a) This answer is the same as part a. of Exercise 4-127.

$$P(X < 5) = \int_0^5 0.1e^{-0.1x}dx = -e^{-0.1x}\Big|_0^5 = 1 - e^{-0.5} = 0.3935$$

b) This is the probability that there are no calls over a period of 5 minutes. Because a Poisson process is memoryless, it does not matter whether or not the intervals are consecutive.

$$P(X > 5) = \int_5^\infty 0.1e^{-0.1x}dx = -e^{-0.1x}\Big|_5^\infty = e^{-0.5} = 0.6065$$

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

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.

4-130 Let X denote the lifetime. Then $\lambda = 1/E(X) = 1/6$.

$$P(X < 3) = \int_0^3 \tfrac{1}{6}e^{-x/6}dx = -e^{-x/6}\Big|_0^3 = 1 - e^{-0.5} = 0.3935.$$

4-131. Let W denote the number of CPUs that fail within the next three years. Then, W is a binomial random variable with n = 10 and p = 0.3935 (from Exercise 4-130). Then,

$$P(W \ge 1) = 1 - P(W = 0) = 1 - \binom{10}{0}0.3935^0(1 - 0.3935)^{10} = 0.9933.$$

4-132 X is a lognormal distribution with θ=0 and ω^2=4

a.)

$$P(10 < X < 50) = P(10 < e^W < 50) = P(\ln(10) < W > \ln(50))$$

$$= \Phi\left(\frac{\ln(50)-0}{2}\right) - \Phi\left(\frac{\ln(10)-0}{2}\right)$$

$$= \Phi(1.96) - \Phi(1.15) = 0.975002 - 0.874928 = 0.10007$$

b.) $P(X < x) = P(e^W < x) = P(W < \ln(x)) = \Phi\left(\frac{\ln(x)-0}{2}\right) = 0.05$

$$\frac{\ln(x)-0}{2} = -1.64 \quad x = e^{-1.64(2)} = 0.0376$$

c.) $\mu = E(X) = e^{\theta+\omega^2/2} = e^{0+4/2} = e^2 = 7.389$

$$V(X) = e^{2\theta+\omega^2}(e^{\omega^2}-1) = e^{0+4}(e^4-1) = e^4(e^4-1) = 2926.40$$

4-133 a.) Find the values of θand ω^2 given that E(X) = 50 and V(X) = 4000

$$50 = e^{\theta+\omega^2/2} \qquad 4000 = e^{2\theta+\omega^2}(e^{\omega^2}-1)$$

let $x = e^\theta$ and $y = e^{\omega^2}$ then (1) $50 = x\sqrt{y}$ and (2) $4000 = x^2y(y-1) = x^2y^2 - x^2y$

Square (1) for x $x = \frac{50}{\sqrt{y}}$ and substitute into (2)

$$4000 = \left(\frac{50}{\sqrt{y}}\right)^2 y^2 - \left(\frac{50}{\sqrt{y}}\right)^2 y = 2500(y-1)$$

$$y = 2.6$$

substitute y back in to (1) and solve for x $x = \frac{50}{\sqrt{2.6}} = 31$

$\theta = \ln(31) = 3.43$ and $\omega^2 = \ln(2.6) = 0.96$

b.)

$$P(X < 150) = P(e^W < 150) = P(W < \ln(150)) = \Phi\left(\frac{\ln(150)-3.43}{0.98}\right)$$

$$= \Phi(1.61) = 0.946301$$

4-134 Let X denote the number of fibers visible in a grid cell. Then, X has a Poisson distribution and $\lambda = 100$ fibers per cm^2 = 80,000 fibers per sample = 0.5 fibers per grid cell.

a) $P(X \geq 1) = 1 - P(X) = 1 - \frac{e^{-0.5} 0.5^0}{0!} = 0.3935$.

b) Let W denote the number of grid cells examined until 10 contain fibers. If the number of fibers have a Poisson distribution, then the number of fibers in each grid cell are independent. Therefore, W has a negative binomial distribution with p = 0.3935. Consequently, E(W) = 10/0.3935 = 25.41 cells.

c) V(W) = $\frac{10(1-0.3935)}{0.3935^2}$. Therefore, $\sigma_W = 6.25$ cells.

4-135. Let X denote the height of a plant.

a) P(X>2.25) = $P\left(Z > \frac{2.25-2.5}{0.5}\right)$ = P(Z > -0.5) = 1 - P(Z ≤ -0.5) = 0.6915

b) P(2.0 < X < 3.0) = $P\left(\frac{2.0-2.5}{0.5} < Z < \frac{3.0-2.5}{0.5}\right)$ =P(-1 < Z < 1) = 0.683

c.)P(X > x) = 0.90 = $P\left(Z > \frac{x-2.5}{0.5}\right)$ = 0.90 and $\frac{x-2.5}{0.5}$ = -1.28.

Therefore, x = 1.86.

4-136 a) P(X > 3.5) from part a. of Exercise 4-135 is 0.023.

b) Yes, because the probability of a plant growing to a height of 3.5 centimeters or more without irrigation is low.

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) = 0.95, then $P\left(Z > \frac{x-5}{0.2}\right)$ = 0.95. Therefore, $\frac{x-5}{0.2}$ = 1.65 and x = 5.33.

4-138 Let X denote the dot diameter. If P(0.0014 < X < 0.0026) = 0.9973, then

$$P\left(\frac{0.0014-0.002}{\sigma} < Z < \frac{0.0026-0.002}{\sigma}\right) = P\left(\frac{-0.0006}{\sigma} < Z < \frac{0.0006}{\sigma}\right) = 0.9973.$$

Therefore, $\frac{0.0006}{\sigma}$ = 3 and $\sigma = 0.0002$.

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.

4-140 Let X denote the life.

a) $P(X < 5800) = P(Z < \frac{5800-7000}{600}) = P(Z < -2) = 1 - P(Z \le 2) = 0.023$

d) If P(X > x) = 0.9, then P(Z < $\frac{x-7000}{600}$) = -1.28. Consequently, $\frac{x-7000}{600}$ = -1.28 and x = 6232 hours.

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 $\mu = 11{,}398$.

4-142 The probability a product lasts more than 10000 hours is $[P(X > 10000)]^3$, by independence. If $[P(X > 10000)]^3$ = 0.99, then P(X > 10000) = 0.9967.

Then, P(X > 10000) = $P(Z > \frac{10000-\mu}{600}) = 0.9967$. Therefore, $\frac{10000-\mu}{600}$ = -2.72 and $\mu = 11{,}632$ hours.

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

4-144 Find the values of θ and ω^2 given that E(X) = 7000 and σ= 600

$$7000 = e^{\theta+\omega^2/2} \qquad 600^2 = e^{2\theta+\omega^2}(e^{\omega^2}-1)$$

let $x = e^{\theta}$ and $y = e^{\omega^2}$ then (1) $7000 = x\sqrt{y}$ and

(2) $600^2 = x^2 y(y-1) = x^2y^2 - x^2 y$

Square (1) $7000^2 = x^2 y$ and substitute into (2)

$$600^2 = 7000^2(y-1)$$

$$y = 1.0073$$

Substitute y into (1) and solve for x $x = \dfrac{7000}{\sqrt{1.0073}} = 6974.6$

$\theta = \ln(6974.6) = 8.850$ and $\omega^2 = \ln(1.0073) = 0.0073$

a.)

$$P(X < 5800) = P(e^W < 5800) = P(W < \ln(5800)) = \Phi\left(\frac{\ln(5800)-8.85}{0.0854}\right)$$

$$= \Phi(-2.16) = 0.015$$

b.) $P(X > x) = P(e^W > x) = P(W > \ln(x)) = \Phi\left(\dfrac{\ln(x)-8.85}{0.0854}\right) = 0.9$

$$\frac{\ln(x)-8.85}{0.0854} = -1.28 \quad x = e^{-1.28(0.0854)+8.85} = 6252.20 \text{ hours}$$

4-145. a) Using the normal approximation to the binomial with n = 50*36*36=64,800, and p = 0.0001 we have: E(X)=64800(0.0001)=6.48

$$P(X \ge 1) \cong P\left(\frac{X-np}{\sqrt{np(1-p)}} \ge \frac{1-6.48}{\sqrt{64800(0.0001)(0.9999)}}\right)$$

$$= P(Z > -2.15) = 1 - 0.01578 = 0.98422$$

b)

$$P(X \ge 4) \cong P\left(\frac{X-np}{\sqrt{np(1-p)}} \ge \frac{4-6.48}{\sqrt{64800(0.0001)(0.9999)}}\right)$$

$$= P(Z \ge -0.97) = 1 - 0.166023 = 0.8340$$

4-146 Using the normal approximation to the binomial with X being the number of people who will be seated.

X ~Bin(200, 0.9).

a) P(X ≤ 185) $= P\left(\frac{X-np}{\sqrt{np(1-p)}} \geq \frac{185-180}{\sqrt{200(0.9)(0.1)}}\right) = P(Z \leq 1.18) = 0.8810$

b)

$$P(X < 185)$$
$$= P(X \leq 184) = P\left(\frac{X-np}{\sqrt{np(1-p)}} \geq \frac{184-180}{\sqrt{200(0.9)(0.1)}}\right) = P(Z \leq 0.94) = 0.8264$$

c) P(X ≤ 185) ≅ 0.95, Successively trying various values of n: The number of reservations taken could be reduced to about 198.

n	Z_o	Probability P(Z < Z_0)
190	3.39	0.99965
195	2.27	0.988396
198	**1.61**	**0.946301**

Mind-Expanding Exercises

4-147. a) P(X > x) implies that there are r - 1 or less counts in an interval of length x. Let Y denote the number of counts in an interval of length x. Then, Y is a Poisson random variable with parameter λx. Then,

$$P(X > x) = P(Y \leq r-1) = \sum_{i=0}^{r-1} e^{-\lambda x} \frac{(\lambda x)^i}{i!}.$$

b) $P(X \leq x) = 1 - \sum_{i=0}^{r-1} e^{-\lambda x} \frac{(\lambda x)^i}{i!}$

c) $f_X(x) = \frac{d}{dx} F_X(x) = \lambda e^{-\lambda x} \sum_{i=0}^{r-1} \frac{(\lambda x)^i}{i!} - e^{-\lambda x} \sum_{i=0}^{r-1} \lambda i \frac{(\lambda x)^i}{i!} = \lambda e^{-\lambda x} \frac{(\lambda x)^{r-1}}{(r-1)!}$

4-148. Let X denote the diameter of the maximum diameter bearing. Then, P(X > 1.6) = 1 - $P(X \leq 1.6)$. Also, $X \leq 1.6$ if and only if all the diameters are less than 1.6. Let Y denote the diameter of a bearing. Then, by independence

$$P(X \leq 1.6) = [P(Y \leq 1.6)]^{10} = \left[P(Z \leq \tfrac{1.6-1.5}{0.025})\right]^{10} = 0.999967^{10} = 0.99967$$

Then, P(X > 1.575) = 0.0033.

4-149. a) Quality loss = $Ek(X-m)^2 = kE(X-m)^2 = k\sigma^2$, by the definition of the variance.

b)

$$Quality\ \ loss = Ek(X-m)^2 = kE(X-\mu+\mu-m)^2$$
$$= kE[(X-\mu)^2 + (\mu-m)^2 + 2(\mu-m)(X-\mu)]$$
$$= kE(X-\mu)^2 + k(\mu-m)^2 + 2k(\mu-m)E(X-\mu).$$

The last term equals zero by the definition of the mean.

Therefore, quality loss = $k\sigma^2 + k(\mu-m)^2$.

4-150. Let X denote the event that an amplifier fails before 60,000 hours. Let A denote the event that an amplifier mean is 20,000 hours. Then A' is the event that the mean of an amplifier is 50,000 hours. Now, P(E) = P(E|A)P(A) + P(E|A')P(A') and

$$P(E\mid A) = \int_0^{60,000} \tfrac{1}{20,000} e^{-x/20,000}dx = -e^{-x/20,000}\Big|_0^{60,000} = 1-e^{-3} = 0.9502$$

$$P(E\mid A') = -e^{-x/50,000}\Big|_0^{60,000} = 1-e^{-6/5} = 0.6988.$$

Therefore, P(E) = 0.9502(0.10) + 0.6988(0.90) = 0.7239

4-151. $P(X < t_1 + t_2 \mid X > t_1) = \frac{P(t_1 < X < t_1 + t_2)}{P(X > t_1)}$ from the definition of conditional probability. Now,

$$P(t_1 < X < t_1 + t_2) = \int_{t_1}^{t_1+t_2} \lambda e^{-\lambda x}dx = -e^{-\lambda x}\Big|_{t_1}^{t_1+t_2} = e^{-\lambda t_1} - e^{-\lambda(t_1+t_2)}$$

$$P(X > t_1) = -e^{-\lambda x}\Big|_{t_1}^{\infty} = e^{-\lambda t_1}$$

Therefore, $P(X < t_1 + t_2 \mid X > t_1) = \dfrac{e^{-\lambda t_1}(1-e^{-\lambda t_2})}{e^{-\lambda t_1}} = 1 - e^{-\lambda t_2} = P(X < t_2)$

4-152. a)

$$1-P(\mu_0-6\sigma < X < \mu_0+6\sigma) = 1-P(-6 < Z < 6)$$
$$= 1.97\times 10^{-9} = 0.00197\,ppm$$

b)

$$1-P(\mu_0-6\sigma < X < \mu_0+6\sigma) = 1-P(-7.5 < \tfrac{X-(\mu_0+1.5\sigma)}{\sigma} < 4.5)$$
$$= 3.4\times 10^{-6} = 3.4\,ppm$$

c)

$$1-P(\mu_0-3\sigma < X < \mu_0+3\sigma) = 1-P(-3 < Z < 3)$$
$$= .0027 = 2{,}700\,ppm$$

d)

$$1-P(\mu_0-3\sigma < X < \mu_0+3\sigma) = 1-P(-4.5 < \tfrac{X-(\mu_0+1.5\sigma)}{\sigma} < 1.5)$$
$$= 0.0668106 = 66{,}810.6\,ppm$$

e) If the process is centered six standard deviations away from the specification limits and the process shift, there will be significantly less product loss. If the process is centered only three standard deviations away from the specifications and the process shifts, there could be a great loss of product.

Section 4-8 on CD

S4-1. E(X) = 50(0.1) = 5 and V(X) = 50(0.1)(0.9) = 4.5

a) $P(X \le 2) = P(X \le 2.5) \cong P(Z \le \frac{2.5-5}{\sqrt{4.5}}) = P(Z \le -1.18) = 0.119$

b) $P(X \le 2) \cong P(Z \le \frac{2-2.5}{\sqrt{4.5}}) = P(Z \le -0.24) = 0.206$

c) $P(X \le 2) = \binom{50}{0}0.1^0 0.9^{50} + \binom{50}{1}0.1^1 0.9^{49} + \binom{50}{2}0.1^2 0.9^{48} = 0.118$

The probability computed using the continuity correction is closer.

d) $P(X < 10) = P(X \le 9.5) \cong P\left(Z \le \frac{9.5-5}{\sqrt{4.5}}\right) = P(Z \le 2.12) = 0.983$

S4-2. E(X) = 50(0.1) = 5 and V(X) = 50(0.1)(0.9) = 4.5

a) $P(X \ge 2) = P(X \ge 1.5) \cong P(Z \le \frac{1.5-5}{\sqrt{4.5}}) = P(Z \ge -1.65) = 0.951$

b) $P(X \ge 2) \cong P(Z \ge \frac{2-5}{\sqrt{4.5}}) = P(Z \ge -1.414) = 0.921$

c) $P(X \ge 2) = 1-P(X < 2) = 1-\binom{50}{0}0.1^0 0.9^{50} - \binom{50}{1}0.1^1 0.9^{49} = 0.966$

The probability computed using the continuity correction is closer.

d) $P(X > 6) = P(X \ge 7) = P(X \ge 6.5) \cong P(Z \ge 0.707) = 0.24$

S4-3. E(X) = 50(0.1) = 5 and V(X) = 50(0.1)(0.9) = 4.5

a)

$$P(2 \le X \le 5) = P(1.5 \le X \le 5.5) \cong P(\tfrac{1.5-5}{\sqrt{4.5}} \le Z \le \tfrac{5.5-5}{\sqrt{4.5}})$$

$$= P(-1.65 \le Z \le 0.236) = P(Z \le 0.24) - P(Z \le -1.65)$$

$$= 0.5948 - (1 - 0.95053) = 0.5453$$

b)

$$P(2 \le X \le 5) \cong P(\tfrac{2-5}{\sqrt{4.5}} \le Z \le \tfrac{5-5}{\sqrt{4.5}}) = P(-1.414 \le Z \le 0)$$

$$= 0.5 - P(Z \le -1.414)$$

$$= 0.5 - (1 - 0.921) = 0.421$$

The exact probability is 0.582

S4-4. E(X) = 50(0.1) = 5 and V(X) = 50(0.1)(0.9) = 4.5

a)

$$P(X = 10) = P(9.5 \le X \le 10.5) \cong P(\tfrac{9.5-5}{\sqrt{4.5}} \le Z \le \tfrac{10.5-5}{\sqrt{4.5}})$$

$$= P(2.121 \le Z \le 2.593) = 0.012$$

b)

$$P(X = 5) = P(4.5 \le X \le 5.5) \cong P(\tfrac{4.5-5}{\sqrt{4.5}} \le Z \le \tfrac{5.5-5}{\sqrt{4.5}})$$

$$= P(-0.24 \le Z \le 0.24) = 0.1897$$

S4-5 Let X be the number of chips in the lot that are defective. Then E(X)=1000(0.02)=20 and V(X)=1000(0.02)(0.98)=19.6

a) P(20 ≤ X ≤ 30)=P(19.5 ≤ X ≤ 30.5) =

$$P\left(\frac{19.5-20}{\sqrt{19.6}} \le Z \le \frac{30.5-20}{\sqrt{19.6}}\right) = P(-.11 \le Z \le 2.37) = 0.9911 - 0.4562 = 0.5349$$

b) P(X=20)=P(19.5 ≤ X ≤ 20.5) =P(-0.11 ≤ Z ≤ 0.11)=0.5438-0.4562=0.0876.

c) The answer should be close to the mean. Substituting values close to the mean, we find x=20 gives the maximum probability.

CHAPTER 5

Section 5-1

5-1. First, f(x,y) ≥ 0. Let R denote the range of (X,Y).
Then, $\sum_R f(x,y) = \frac{1}{4} + \frac{1}{8} + \frac{1}{4} + \frac{1}{4} + \frac{1}{8} = 1$

5-2 a) P(X < 2.5, Y < 3) = f(1.5,2)+f(1,1) = 1/8+1/4=3/8
b) P(X < 2.5) = f (1.5, 2) + f (1.5, 3) +f(1,1)= 1/8 + 1/4+1/4 = 5/8
c) P(Y < 3) = f (1.5, 2)+f(1,1) = 1/8+1/4=3/8
d) P(X > 1.8, Y > 4.7) = f (3, 5) = 1/8

5-3. E(X) = 1(1/4)+ 1.5(3/8) + 2.5(1/4) + 3(1/8) = 1.8125
E(Y) = 1(1/4)+2(1/8) + 3(1/4) + 4(1/4) + 5(1/8) = 2.875

5-4 a) marginal distribution of *X*

x	f(x)
1	1/4
1.5	3/8
2.5	1/4
3	1/8

b) $f_{Y|1.5}(y) = \frac{f_{XY}(1.5, y)}{f_X(1.5)}$ and $f_X(1.5)$ = 3/8. Then,

y	$f_{Y\|1.5}(y)$
2	(1/8)/(3/8)=1/3
3	(1/4)/(3/8)=2/3

c) $f_{X|2}(x) = \frac{f_{XY}(x,2)}{f_Y(2)}$ and $f_Y(2)$ = 1/8. Then,

x	$f_{X\|2}(y)$
1.5	(1/8)/(1/8)=1

d) *E*(*Y*|*X*=1.5) = 2(1/3)+3(2/3) =2 1/3
e) Since $f_{Y|1.5}(y) \neq f_Y(y)$, X and Y are not independent

5-5 Let R denote the range of (X,Y). Because
$\sum_R f(x,y) = c(2+3+4+3+4+5+4+5+6) = 1$, 36c = 1, and c = 1/36

5-6. a) $P(X = 1, Y < 4) = f_{XY}(1,1) + f_{XY}(1,2) + f_{XY}(1,3) = \frac{1}{36}(2+3+4) = 1/4$
b) P(X = 1) is the same as part a. = 1/4
c) $P(Y = 2) = f_{XY}(1,2) + f_{XY}(2,2) + f_{XY}(3,2) = \frac{1}{36}(3+4+5) = 1/3$
d) $P(X < 2, Y < 2) = f_{XY}(1,1) = \frac{1}{36}(2) = 1/18$

5-7.

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

5-8 a) marginal distribution of X

x	$f_X(x) = f_{XY}(x,1) + f_{XY}(x,2) + f_{XY}(x,3)$
1	1/4
2	1/3
3	5/12

b) $f_{Y|X}(y) = \dfrac{f_{XY}(1,y)}{f_X(1)}$

y	$f_{Y\|X}(y)$
1	(2/36)/(1/4)=2/9
2	(3/36)/(1/4)=1/3
3	(4/36)/(1/4)=4/9

c) $f_{X|Y}(x) = \dfrac{f_{XY}(x,2)}{f_Y(2)}$ and $f_Y(2) = f_{XY}(1,2) + f_{XY}(2,2) + f_{XY}(3,2) = \frac{12}{36} = 1/3$

x	$f_{X\|Y}(x)$
1	(3/36)/(1/3)=1/4
2	(4/36)/(1/3)=1/3
3	(5/36)/(1/3)=5/12

d) $E(Y|X=1) = 1(2/9)+2(1/3)+3(4/9) = 20/9$

e) Since $f_{XY}(x,y) \neq f_X(x)f_Y(y)$, X and Y are not independent.

5-9. $f(x,y) \geq 0 \ and \ \sum_R f(x,y) = 1$

5-10. a) $P(X < 0.5, Y < 1.5) = f_{XY}(-1,-2) + f_{XY}(-0.5,-1) = \frac{1}{8} + \frac{1}{4} = \frac{3}{8}$

b) $P(X < 0.5) = f_{XY}(-1,-2) + f_{XY}(-0.5,-1) = \frac{3}{8}$

c) $P(Y < 1.5) = f_{XY}(-1,-2) + f_{XY}(-0.5,-1) + f_{XY}(0.5,1) = \frac{7}{8}$

d) $P(X > 0.25, Y < 4.5) = f_{XY}(0.5,1) + f_{XY}(1,2) = \frac{5}{8}$

5-11.

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

5-12 a) marginal distribution of *X*

x	$f_X(x)$
-1	1/8
-0.5	¼
0.5	½
1	1/8

b) $f_{Y|X}(y) = \dfrac{f_{XY}(1,y)}{f_X(1)}$

y	$f_{Y\|X}(y)$
2	1/8/(1/8)=1

c) $f_{X|Y}(x) = \dfrac{f_{XY}(x,1)}{f_Y(1)}$

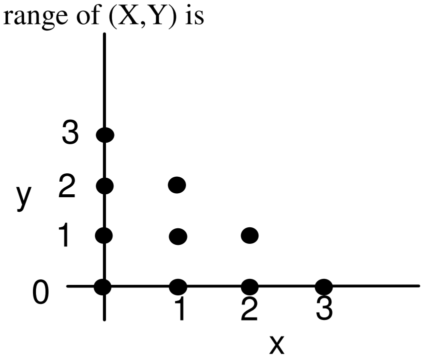
x	$f_{X\|Y}(x)$
0.5	½/(1/2)=1

d) $E(Y|X=1) = 0.5$

e) no, *X* and *Y* are not independent

5-13. Because X and Y denote the number of printers in each category,
$X \geq 0,\ Y \geq 0\ and\ X + Y = 4$

5-14. a) The range of (X,Y) is



Let H = 3, M = 2, and L = 1 denote the events that a bit has high, moderate, and low distortion, respectively. Then,

x,y	$f_{xy}(x,y)$
0,0	0.85738
0,1	0.1083
0,2	0.00456
0,3	0.000064
1,0	0.027075
1,1	0.00228
1,2	0.000048
2,0	0.000285
2,1	0.000012
3,0	0.000001

b)

x	$f_x(x)$
0	0.970299
1	0.029835
2	0.000297
3	0.000001

c) $f_{Y|1}(y) = \dfrac{f_{XY}(1, y)}{f_X(1)}$, $f_x(1) = 0.29835$

y	$f_{Y|1}(x)$
0	0.09075
1	0.00764
2	0.000161

E(X)=0(0.970299)+1(0.029403)+2(0.000297)+3*(0.000001)=0.03 (or np=3*.01).

d) $f_{Y|1}(y) = \dfrac{f_{XY}(1, y)}{f_X(1)}$, $f_x(1) = 0.029403$

y	$f_{Y|1}(x)$
0	0.920824
1	0.077543
2	0.001632

e) E(Y|X=1)=0(.920824)+1(0.077543)+2(0.001632)=0.080807

f) No, X and Y are not independent since, for example, $f_Y(0) \neq f_{Y|1}(0)$.

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	0.00184	0.00092	0	0	0
y=2	0.02031	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
f) V(Y|X=3) = 0^2(0.857)+1^2(0.143)- 0.143^2= 0.123
g) no, *X* and *Y* are not independent

5-16 a) The range of (X,Y) is $X \geq 0,\ Y \geq 0\ and\ X + Y \leq 4$. *X* is the number of defective items found with inspection device 1 and Y is the number of defective items found with inspection device 2.

	x=0	x=1	x=2	x=3	x=4
y=0	1.94×10^{-19}	1.10×10^{-16}	2.35×10^{-14}	2.22×10^{-12}	7.88×10^{-11}
y=1	2.59×10^{-16}	1.47×10^{-13}	3.12×10^{-11}	2.95×10^{-9}	1.05×10^{-7}
y=2	1.29×10^{-13}	7.31×10^{-11}	1.56×10^{-8}	1.47×10^{-6}	5.22×10^{-5}
y=3	2.86×10^{-11}	1.62×10^{-8}	3.45×10^{-6}	3.26×10^{-4}	0.0116
y=4	2.37×10^{-9}	1.35×10^{-6}	2.86×10^{-4}	0.0271	0.961

$$f(x,y) = \left[\binom{4}{x}(0.993)^x (0.007)^{4-x}\right]\left[\binom{4}{y}(0.997)^y (0.003)^{4-y}\right]$$

For *x*=1,2,3,4 and *y*=1,2,3,4

b.)

$$f(x,y) = \left[\binom{4}{x}(0.993)^x (0.007)^{4-x}\right] \text{ for } x = 1,2,3,4$$

	x=0	x=1	x=2	x=3	x=4
f(x)	2.40×10^{-9}	1.36×10^{-6}	2.899×10^{-4}	0.0274	0.972

c.)since x is binomial $E(X) = n(p) = 4*(0.993) = 3.972$

d.) $f_{Y|2}(y) = \dfrac{f_{XY}(2,y)}{f_X(2)} = f(y)$, $f_x(2) = 0.0725$

y	$f_{Y|1}(y)=f(y)$
0	8.1×10^{-11}
1	1.08×10^{-7}
2	5.37×10^{-5}
3	0.0119
4	0.988

e) E(Y|X=2) = E(Y)= *n*(*p*)= 4(0.997)=3.988

f) V(Y|X=2) = V(Y)=*n*(*p*)(1-*p*)=4(0.997)(0.003)=0.0120

g) yes, *X* and *Y* are independent.

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) c) $P(Z<1.5)=f_{XYZ}(1,1,1)+f_{XYZ}(1,2,1)+f_{XYZ}(2,1,1)+f_{XYZ}(2,2,1)=0.5$

d)

$$P(X=1\ or\ Z=2)=P(X=1)+P(Z=2)-P(X=1,Z=2)=0.5+0.5-0.3=0.7$$

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

5-18 a) $P(X=1\,|\,Y=1)=\dfrac{P(X=1,Y=1)}{P(Y=1)}=\dfrac{0.05+0.10}{0.15+0.2+0.1+0.05}=0.3$

b) $P(X=1,Y=1\,|\,Z=2)==\dfrac{P(X=1,Y=1,Z=2)}{P(Z=2)}=\dfrac{0.1}{0.1+0.2+0.15+0.05}=0.2$

c) $P(X=1\,|\,Y=1,Z=2)=\dfrac{P(X=1,Y=1,Z=2)}{P(Y=1,Z=2)}=\dfrac{0.10}{0.10+0.15}=0.4$

5-19. $f_{X|YZ}(x)=\dfrac{f_{XYZ}(x,1,2)}{f_{YZ}(1,2)}$ and $f_{YZ}(1,2)=f_{XYZ}(1,1,2)+f_{XYZ}(2,1,2)=0.25$

x	$f_{X\|YZ}(x)$
1	0.10/0.25=0.4
2	0.15/0.25=0.6

5-20 a.) percentage of slabs classified as high = p_1 = 0.05
percentage of slabs classified as medium = p_2 = 0.85
percentage of slabs classified as low = p_3 = 0.10

b.) X is the number of voids independently classified as high $X \geq 0$
Y is the number of voids independently classified as medium $Y \geq 0$
Z is the number of with a low number of voids and $Z \geq 0$
And $X+Y+Z = 20$

c.) p_1 is the percentage of slabs classified as high.

d) E(X)=np_1 = 20(0.05) = 1
V(X)=np_1 (1-p_1)= 20(0.05)(0.95) = 0.95

5-21. a) $P(X=1,Y=17,Z=3)=0$ Because the point $(1,17,3)\neq 20$ is not in the range of (X,Y,Z).

b)
$$P(X\le 1,Y=17,Z=3)=P(X=0,Y=17,Z=3)+P(X=1,Y=17,Z=3)$$
$$=\frac{20!}{0!17!3!}0.05^{0}0.85^{17}0.10^{3}+0=0.07195$$

Because the point $(1,17,3)\neq 20$ is not in the range of (X,Y,Z).

c) Because X is binomial, $P(X\le 1)=\binom{20}{0}0.05^{0}0.95^{20}+\binom{20}{1}0.05^{1}0.95^{19}=0.7358$

d.) Because X is binomial E(X) = np = 20(0.05) = 1

5-22 a) The probability is 0 since x+y+z>20

$$P(X=2,Z=3\mid Y=17)=\frac{P(X=2,Z=3,Y=17)}{P(Y=17)}.$$

Because Y is binomial, $P(Y=17)=\binom{20}{17}0.85^{17}0.15^{3}=0.2428$.

Then, $P(X=2,Z=3,Y=17)=\dfrac{20!}{2!3!17!}\dfrac{0.05^{2}0.85^{17}0.10^{3}}{0.2054}=0$

b) $P(X=2\mid Y=17)=\dfrac{P(X=2,Y=17)}{P(Y=17)}$. Now, because x+y+z = 20,

P(X=2, Y=17) = P(X=2, Y=17, Z=1) = $\dfrac{20!}{2!17!1!}0.05^{2}0.85^{17}0.10^{1}=0.0540$

$$P(X=2\mid Y=17)=\frac{P(X=2,Y=17)}{P(Y=17)}=\frac{0.0540}{0.2428}=0.2224$$

c)
$$E(X\mid Y=17)=0\left(\frac{P(X=0,Y=17)}{P(Y=17)}\right)+1\left(\frac{P(X=1,Y=17)}{P(Y=17)}\right)$$
$$+2\left(\frac{P(X=2,Y=17)}{P(Y=17)}\right)+3\left(\frac{P(X=3,Y=17)}{P(Y=17)}\right)$$
$$E(X\mid Y=17)=0\left(\frac{0.07195}{0.2428}\right)+1\left(\frac{0.1079}{0.2428}\right)+2\left(\frac{0.05396}{0.2428}\right)+3\left(\frac{0.008994}{0.2428}\right)$$
$$=1$$

5-23. a) The range consists of nonnegative integers with $x+y+z=4$.

b) Because the samples are selected without replacement, the trials are not independent and the joint distribution is not multinomial.

5-24 $$P(X = x \mid Y = 2) = \frac{f_{XY}(x,2)}{f_Y(2)}$$

$$P(Y=2) = \frac{\binom{4}{0}\binom{5}{2}\binom{6}{2}}{\binom{15}{4}} + \frac{\binom{4}{1}\binom{5}{2}\binom{6}{1}}{\binom{15}{4}} + \frac{\binom{4}{2}\binom{5}{2}\binom{6}{0}}{\binom{15}{4}} = 0.1098 + 0.1758 + 0.0440 = 0.3296$$

$$P(X = 0 \text{ and } Y = 2) = \frac{\binom{4}{0}\binom{5}{2}\binom{6}{2}}{\binom{15}{4}} = 0.1098$$

$$P(X = 1 \text{ and } Y = 2) = \frac{\binom{4}{1}\binom{5}{2}\binom{6}{1}}{\binom{15}{4}} = 0.1758$$

$$P(X = 2 \text{ and } Y = 2) = \frac{\binom{4}{1}\binom{5}{2}\binom{6}{1}}{\binom{15}{4}} = 0.0440$$

x	$f_{XY}(x,2)$
0	0.1098/0.3296=0.3331
1	0.1758/0.3296=0.5334
2	0.0440/0.3296=0.1335

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) $P(X = 1, Y = 2, Z = 1) = \frac{\binom{4}{1}\binom{5}{2}\binom{6}{1}}{\binom{15}{4}} = 0.1758$

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

c) The marginal distribution of X is hypergeometric with N = 15, n = 4, K = 4. Therefore, E(X) = nK/N = 16/15 and V(X) = 4(4/15)(11/15)[11/14] = 0.6146.

5-26 a)

$$P(X = 1, Y = 2 \mid Z = 1) = P(X = 1, Y = 2, Z = 1) / P(Z = 1)$$
$$= \left[\frac{\binom{4}{1}\binom{5}{2}\binom{6}{1}}{\binom{15}{4}}\right] \Big/ \left[\frac{\binom{6}{1}\binom{9}{3}}{\binom{15}{4}}\right] = 0.4762$$

b)

$$P(X = 2 \mid Y = 2) = P(X = 2, Y = 2) / P(Y = 2)$$
$$= \left[\frac{\binom{4}{2}\binom{5}{2}\binom{6}{0}}{\binom{15}{4}}\right] \Big/ \left[\frac{\binom{5}{2}\binom{10}{2}}{\binom{15}{4}}\right] = 0.1334$$

c) Because X+Y+Z = 4, if Y = 0 and Z = 3, then X = 1. Because X must equal 1, $f_{X|YZ}(1) = 1$.

5-27. a)The probability distribution is multinomial because the result of each trial (a dropped oven) results in either a major, minor or no defect with probability 0.6, 0.3 and 0.1 respectively. Also, the trials are independent

b.) Let X, Y, and Z denote the number of ovens in the sample of four with major, minor, and no defects, respectively.

$$P(X = 2, Y = 2, Z = 0) = \frac{4!}{2!2!0!} 0.6^2 0.3^2 0.1^0 = 0.1944$$

c.) $$P(X = 0, Y = 0, Z = 4) = \frac{4!}{0!0!4!} 0.6^0 0.3^0 0.1^4 = 0.0001$$

5-28 a.) $f_{XY}(x,y) = \sum_R f_{XYZ}(x,y,z)$ where R is the set of values for z such that x+y+z = 4. That is, R consists of the single value z = 4-x-y and

$$f_{XY}(x, y) = \frac{4!}{x!\,y!(4 - x - y)!} 0.6^x 0.3^y 0.1^{4-x-y} \quad \text{for } x + y \le 4.$$

b.) $E(X) = np_1 = 4(0.6) = 2.4$

c.) $E(Y) = np_2 = 4(0.3) = 1.2$

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

$P(Y = 2) = \binom{4}{2} 0.3^2 0.7^4 = 0.2646$ 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) = 1.7143

5-30 Let X,Y, and Z denote the number of bits with high, moderate, and low distortion. Then, the joint distribution of X, Y, and Z is multinomial with n =3 and $p_1 = 0.01,\ p_2 = 0.04,\ and\ p_3 = 0.95$.

a)

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

$$= \frac{3!}{2!1!0!}0.01^2 0.04^1 0.95^0 = 1.2 \times 10^{-5}$$

b) $P(X=0, Y=0, Z=3) = \frac{3!}{0!0!3!}0.01^0 0.04^0 0.95^3 = 0.8574$

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 $P(X \mid Y=2)$

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

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

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

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

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

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

5-32 a.) Let X,Y, and Z denote the risk of new competitors as no risk, moderate risk, and very high risk. Then, the joint distribution of X, Y, and Z is multinomial with n =12 and $p_1 = 0.13, \quad p_2 = 0.72, \quad and \quad p_3 = 0.15$. X, Y and Z ≥ 0 and x+y+z=12

b.) $P(X = 1, Y = 3, Z = 1) = 0$, not possible since x+y+z≠12

c)

$$P(Z \le 2) = \binom{12}{0} 0.15^0 0.85^{12} + \binom{12}{1} 0.15^1 0.85^{11} + \binom{12}{2} 0.15^2 0.85^{10}$$
$$= 0.1422 + 0.3012 + 0.2924 = 0.7358$$

5-33 a.) $P(Z = 2 \mid Y = 1, X = 10) = 0$

b.) first get

$$P(X = 10) = P(X = 10, Y = 2, Z = 0) + P(X = 10, Y = 1, Z = 1) + P(X = 10, Y = 0, Z = 2)$$
$$= \frac{12!}{10!2!0!} 0.13^{10} 0.72^2 0.15^0 + \frac{12!}{10!1!1!} 0.13^{10} 0.72^1 0.15^1 + \frac{12!}{10!0!2!} 0.13^{10} 0.72^0 0.15^2$$
$$= 4.72x10^{-8} + 1.97x10^{-8} + 2.04x10^{-9} = 6.89x10^{-8}$$
$$P(Z \le 1 \mid X = 10) = \frac{P(Z = 0, Y = 2, X = 10)}{P(X = 10)} + \frac{P(Z = 1, Y = 1, X = 10)}{P(X = 10)}$$
$$= \frac{12!}{10!2!0!} 0.13^{10} 0.72^2 0.15^0 \Big/ 6.89x10^{-8} + \frac{12!}{10!1!1!} 0.13^{10} 0.72^1 0.15^1 \Big/ 6.89x10^{-8}$$
$$= 0.9698$$

c.)

$$P(Y \le 1, Z \le 1 \mid X = 10) = \frac{P(Z = 1, Y = 1, X = 10)}{P(X = 10)}$$
$$= \frac{12!}{10!1!1!} 0.13^{10} 0.72^1 0.15^1 \Big/ 6.89x10^{-8}$$
$$= 0.2852$$

d.

$$E(Z \mid X = 10) = (0(4.72x10^{-8}) + 1(1.97x10^{-8}) + 2(2.04x10^{-9})) / 6.89x10^{-8}$$
$$= 2.378x10^{-8}$$

Section 5-3

5-34 Determine c such that $c\int_0^3\int_0^3 xydxdy = c\int_0^3 y\frac{x^2}{2}\Big|_0^3 dy = c(4.5\frac{y^2}{2}\Big|_0^3) = \frac{81}{4}c.$

Therefore, $c = 4/81$.

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

5-35 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) $P(X<0, Y<4) = \frac{4}{81}\int_0^4\int_0^0 xydxdy = 0\int_0^4 ydy = 0$

5-36 a) $f_X(x) = \int_0^3 f_{XY}(x,y)dy = x\frac{4}{81}\int_0^3 ydy = \frac{4}{81}x(4.5) = \frac{2x}{9}$ *for* $0<x<3$.

b) $f_{Y|1.5}(y) = \dfrac{f_{XY}(1.5,y)}{f_X(1.5)} = \dfrac{\frac{4}{81}y(1.5)}{\frac{2}{9}(1.5)} = \frac{2}{9}y$ for $0 < y < 3$.

c) E(Y|X=1.5) = $\int_0^3 y\left(\frac{2}{9}y\right)dy = \frac{2}{9}\int_0^3 y^2dy = \frac{2y^3}{27}\Big|_0^3 = 6$

d.) $P(Y<2 \mid X=1.5) = f_{Y|1.5}(y) = \int_0^2 \frac{2}{9}ydy = \frac{1}{9}y^2\Big|_0^2 = \frac{4}{9} - 0 = \frac{4}{9}$

e) $f_{X|2}(x) = \dfrac{f_{XY}(x,2)}{f_Y(2)} = \dfrac{\frac{4}{81}x(2)}{\frac{2}{9}(2)} = \frac{2}{9}x$ for 0 < x < 3.

5-37.

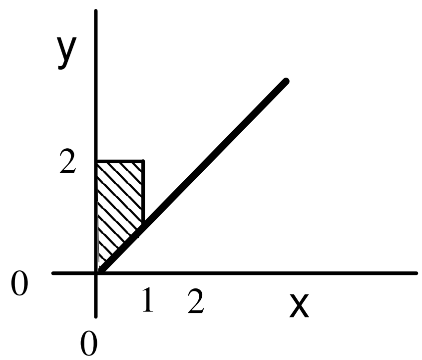
$$c\int_0^3\int_x^{x+2}(x+y)dydx = \int_0^3 xy + \tfrac{y^2}{2}\Big|_x^{x+2} dx$$

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

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

Therefore, c = 1/24.

5-38 a) P(X < 1, Y < 2) equals the integral of $f_{XY}(x, y)$ over the following region.

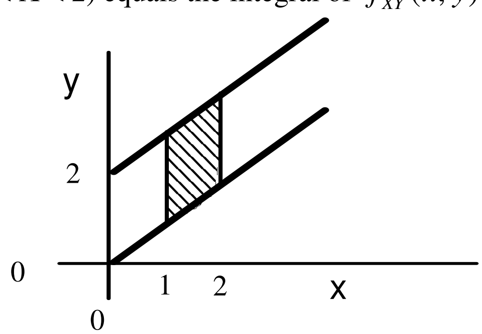


Then,

$$P(X < 1, Y < 2) = \frac{1}{24}\int_0^1\int_x^2(x+y)dydx = \frac{1}{24}\int_0^1 xy + \tfrac{y^2}{2}\Big|_x^2 dx = \frac{1}{24}\int_0^3 2x + 2 - \tfrac{3x^2}{2}dx =$$

$$\frac{1}{24}\left[x^2 + 2x - \tfrac{x^3}{2}\Big|_0^1\right] = 0.10417$$

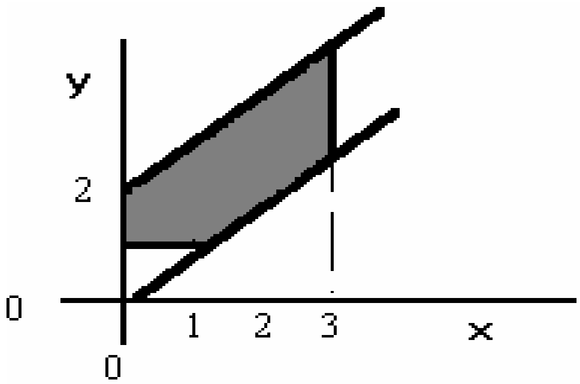
b) P(1 < X < 2) equals the integral of $f_{XY}(x, y)$ over the following region.



$$P(1 < X < 2) = \frac{1}{24}\int_1^2\int_x^{x+2}(x+y)dydx = \frac{1}{24}\int_1^2 xy + \tfrac{y^2}{2}\Big|_x^{x+2} dx$$

$$= \frac{1}{24}\int_0^3 (4x+2)dx = \frac{1}{24}\left[2x^2 + 2x\Big|_1^2\right] = \tfrac{1}{6}.$$

c) P(Y > 1) is the integral of $f_{XY}(x, y)$ over the following region.

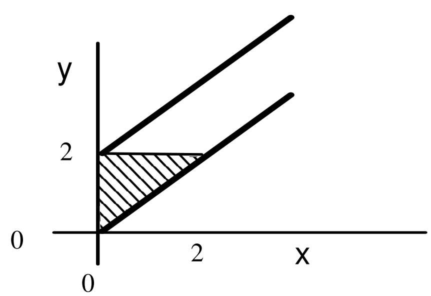


$$P(Y > 1) = 1 - P(Y \le 1) = 1 - \tfrac{1}{24}\int_0^1\int_x^1 (x + y)dydx = 1 - \frac{1}{24}\int_0^1 (xy + \frac{y^2}{2})\Big|_x^1$$

$$= 1 - \frac{1}{24}\int_0^1 x + \frac{1}{2} - \frac{3}{2}x^2 dx = 1 - \frac{1}{24}\left(\frac{x^2}{2} + \frac{1}{2} - \frac{1}{2}x^3\right)\Big|_0^1$$

$$= 1 - 0.02083 = 0.9792$$

d) P(X < 2, Y < 2) is the integral of $f_{XY}(x,y)$ over the following region.



$$E(X) = \frac{1}{24}\int_0^3\int_x^{x+2} x(x + y)dydx = \frac{1}{24}\int_0^3 x^2 y + \tfrac{xy^2}{2}\Big|_x^{x+2} dx$$

$$= \frac{1}{24}\int_0^3 (4x^2 + 2x)dx = \frac{1}{24}\left[\frac{4x^3}{3} + x^2\Big|_0^3\right] = \frac{15}{8}$$

e)

$$E(X) = \frac{1}{24}\int_0^3\int_x^{x+2} x(x + y)dydx = \frac{1}{24}\int_0^3 x^2 y + \tfrac{xy^2}{2}\Big|_x^{x+2} dx$$

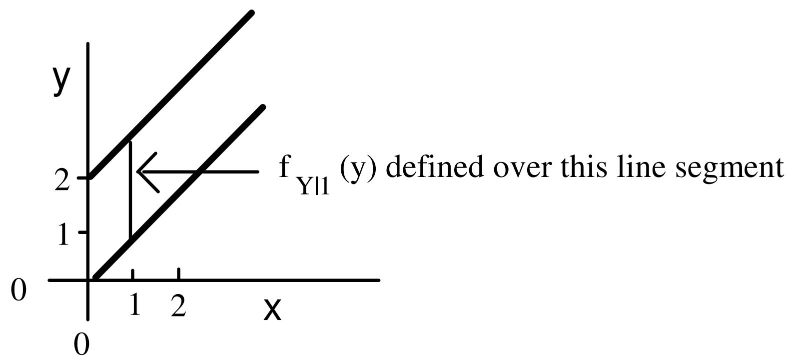
$$= \frac{1}{24}\int_0^3 (3x^2 + 2x)dx = \frac{1}{24}\left[x^3 + x^2\Big|_0^3\right] = \frac{15}{8}$$

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

See the following graph,

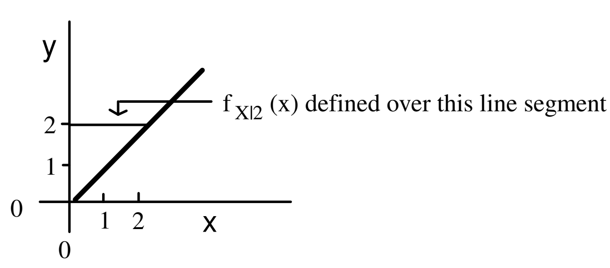


c) E(Y|X=1) = $\int_1^3 y\left(\frac{1+y}{6}\right)dy = \frac{1}{6}\int_1^3 (y+y^2)dy = \frac{1}{6}\left(\frac{y^2}{2}+\frac{y^3}{3}\right)\Bigg|_1^3$ =2.111

d.) $P(Y > 2 \mid X = 1) = \int_2^3\left(\frac{1+y}{6}\right)dy = \frac{1}{6}\int_2^3(1+y)dy = \frac{1}{6}\left(y+\frac{y^2}{2}\right)\Bigg|_2^3$ =0.5833

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 *y*=2, only the first integration is needed.

$$f_Y(y) = \frac{1}{24}\int_0^y (x+y)dx = \frac{1}{24}\left[\frac{x^2}{2} + xy\Big|_0^y\right] = \frac{y^2}{16} \quad \text{for } 0 < y \le 2.$$



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

5-40 $c\int_0^3\int_0^x xydydx = c\int_0^3 x\left.\frac{y^2}{2}\right|_0^x dx = c\int_0^3 \frac{x^3}{2}dx\,\frac{x^4}{8} = \frac{81}{8}c$. Therefore, $c = 8/81$

5-41. a) $P(X<1,Y<2) = \frac{8}{81}\int_0^1\int_0^x xydydx = \frac{8}{81}\int_0^1 \frac{x^3}{2}dx = \frac{8}{81}\left(\frac{1}{8}\right) = \frac{1}{81}.$

b) $P(1<X<2) = \frac{8}{81}\int_1^2\int_0^x xydydx = \frac{8}{81}\int_1^2 x\frac{x^2}{2}dx = \left(\frac{8}{81}\right)\left.\frac{x^4}{8}\right|_1^2 = \left(\frac{8}{81}\right)\frac{(2^4-1)}{8} = \frac{5}{27}.$

c)

$$P(Y>1) = \frac{8}{81}\int_1^3\int_1^x xydydx = \frac{8}{81}\int_1^3 x\left(\frac{x^2-1}{2}\right)dx = \frac{8}{81}\int_1^3 \frac{x^3}{2} - \frac{x}{2}dx = \frac{8}{81}\left.\left(\frac{x^4}{8}-\frac{x^2}{4}\right)\right|_1^3$$

$$= \frac{8}{81}\left[\left(\frac{3^4}{8}-\frac{3^2}{4}\right)-\left(\frac{1^4}{8}-\frac{1^2}{4}\right)\right] = \frac{1}{81} = 0.01235$$

d) $P(X<2,Y<2) = \frac{8}{81}\int_0^2\int_0^x xydydx = \frac{8}{81}\int_0^2 \frac{x^3}{2}dx = \frac{8}{81}\left(\frac{2^4}{8}\right) = \frac{16}{81}.$

e.)

$$E(X) = \frac{8}{81}\int_0^3\int_0^x x(xy)dydx = \frac{8}{81}\int_0^3\int_0^x x^2ydydx = \frac{8}{81}\int_0^3 \frac{x^2}{2}x^2dx = \frac{8}{81}\int_0^3 \frac{x^4}{2}dx$$

$$= \left(\frac{8}{81}\right)\left(\frac{3^5}{10}\right) = \frac{12}{5}$$

f)

$$E(Y) = \frac{8}{81}\int_0^3\int_0^x y(xy)dydx = \frac{8}{81}\int_0^3\int_0^x xy^2dydx = \frac{8}{81}\int_0^3 x\frac{x^3}{3}dx$$

$$= \frac{8}{81}\int_0^3 \frac{x^4}{3}dx = \left(\frac{8}{81}\right)\left(\frac{3^5}{15}\right) = \frac{8}{5}$$

5-42 a.) $f(x) = \frac{8}{81}\int_0^x xydy = \frac{4x^3}{81} \qquad 0 < x < 3$

b.) $f_{Y|x=1}(y) = \frac{f(1,y)}{f(1)} = \frac{\frac{8}{81}(1)y}{\frac{4(1)^3}{81}} = 2y \qquad 0 < y < 1$

c.) $E(Y \mid X=1) = \int_0^1 2ydy = y^2\Big|_0^1 = 1$

d.) $P(Y > 2 \mid X = 1) = 0$ this isn't possible since the values of y are 0< $y < x$.

e.) $f(y) = \frac{8}{81}\int_0^3 xydx = \frac{4y}{9}$, therefore

$$f_{X|Y=2}(x) = \frac{f(x,2)}{f(2)} = \frac{\frac{8}{81}x(2)}{\frac{4(2)}{9}} = \frac{2x}{9} \qquad 0 < x < 2$$

5-43. Solve for c

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

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

5-44 a)

$$P(X<1, Y<2) = 10\int_0^1\int_0^x e^{-2x-3y}dydx = \frac{10}{3}\int_0^1 e^{-2x}(1-e^{-3x})dy = \frac{10}{3}\int_0^1 e^{-2x} - e^{-5x}dy$$

$$= \frac{10}{3}\left(\frac{e^{-5x}}{5} - \frac{e^{-2x}}{2}\right)\Bigg|_0^1 = 0.77893$$

b)

$$P(1 < X < 2) = 10\int_1^2\int_0^x e^{-2x-3y}dydx = \frac{10}{3}\int_1^2 e^{-2x} - e^{-5x}dx$$

$$= \frac{10}{3}\left(\frac{e^{-5x}}{5} - \frac{e^{-2x}}{2}\right)\Bigg|_1^2 = 0.19057$$

c)

$$P(Y > 3) = 10\int_3^\infty\int_3^x e^{-2x-3y}dydx = \frac{10}{3}\int_3^{1\infty} e^{-2x}(e^{-9} - e^{-3x})dy$$

$$= \frac{10}{3}\left(\frac{e^{-5x}}{5} - \frac{e^{-9}e^{-2x}}{2}\right)\Bigg|_3^\infty = 3.059x10^{-7}$$

d)

$$P(X<2,Y<2)=10\int_0^2\int_0^x e^{-2x-3y}dydx=\frac{10}{3}\int_0^2 e^{-2x}(1-e^{-3x})dx=\frac{10}{3}\left(\frac{e^{-10}}{5}-\frac{e^{-4}}{2}\right)\Bigg|_0^2$$

$$=0.9695$$

e) E(X) $=10\int_0^\infty\int_0^x xe^{-2x-3y}dydx=\frac{7}{10}$

f) E(Y) $=10\int_0^\infty\int_0^x ye^{-2x-3y}dydx=\frac{1}{5}$

5-45. a) $f(x)=10\int_0^x e^{-2x-3y}dy=\frac{10e^{-2z}}{3}(1-e^{-3x})=\frac{10}{3}(e^{-2x}-e^{-5x})$ for $0<x$

b) $f_{Y\backslash X=1}(y)=\frac{f_{X,Y}(1,y)}{f_X(1)}=\frac{10e^{-2-3y}}{\frac{10}{3}(e^{-2}-e^{-5})}=3.157e^{-3y}\quad 0<y<1$

c) $E(Y|X=1)=3.157\int_0^1 ye^{-3y}dy=0.2809$

d) $f_{X|Y=2}(x)\ =\ \frac{f_{X,Y}(x,2)}{f_Y(2)}=\frac{10e^{-2x-6}}{5e^{-10}}=2e^{-2x+4}$ for $2<x$,

where $f(y)=5e^{-5y}$ for $0<y$

5-46 $c\int_0^\infty\int_x^\infty e^{-2x}e^{-3y}dydx=\frac{c}{3}\int_0^\infty e^{-2x}(e^{-3x})dx=\frac{c}{3}\int_0^\infty e^{-5x}dx=\frac{1}{15}c$ $\qquad c=15$

5-47. a)

$$P(X<1,Y<2)=15\int_0^1\int_x^2 e^{-2x-3y}dydx=5\int_0^1 e^{-2x}(e^{-3x}-e^{-6})dx$$

$$=5\int_0^1 e^{-5x}dx-5e^{-6}\int_0^1 e^{-2x}dx=1-e^{-5}+\frac{5}{2}e^{-6}(e^{-2}-1)=0.9879$$

b) P(1 < X < 2) = $15\int_1^2\int_x^\infty e^{-2x-3y}dydx=5\int_1^2 e^{-5x}dydx=(e^{-5}-e^{10})=0.0067$

c)

$$P(Y>3)=15\left(\int_0^3\int_3^\infty e^{-2x-3y}dydx+\int_3^\infty\int_x^\infty e^{-2x-3y}dydx\right)=5\int_0^3 e^{-9}e^{-2x}dx+5\int_3^\infty e^{-5x}dx$$

$$=-\frac{3}{2}e^{-15}+\frac{5}{2}e^{-9}=0.000308$$

d)

$$P(X<2,Y<2)=15\int_0^2\int_x^2 e^{-2x-3y}\,dy\,dx=5\int_0^2 e^{-2x}(e^{-3x}-e^{-6})dx=$$

$$=5\int_0^2 e^{-5x}dx-5e^{-6}\int_0^2 e^{-2x}dx=\left(1-e^{-10}\right)+\frac{5}{2}e^{-6}\left(e^{-4}-1\right)=0.9939$$

e) E(X) = $15\int_0^\infty\int_x^\infty xe^{-2x-3y}\,dy\,dx=5\int_0^\infty xe^{-5x}dx=\frac{1}{5^2}=0.04$

f)

$$E(Y)=15\int_0^\infty\int_x^\infty ye^{-2x-3y}\,dy\,dx=\frac{-3}{2}\int_0^\infty 5ye^{-5y}dy+\frac{5}{2}\int_0^\infty 3ye^{-3y}dy$$

$$=-\frac{3}{10}+\frac{5}{6}=\frac{8}{15}$$

5-48 a.) $f(x)=15\int_x^\infty e^{-2x-3y}dy=\frac{15}{3}(e^{-2z-3x})=5e^{-5x}$ for $x>0$

b) $f_X(1)=5e^{-5}$ $f_{XY}(1,y)=15e^{-2-3y}$

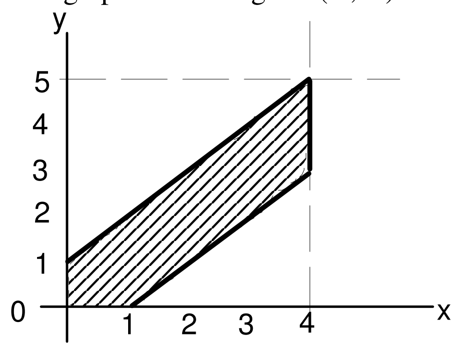
$f_{Y|X=1}(y)=\frac{15e^{-2-3y}}{5e^{-5}}=3e^{3-3y}$ for $1<y$

c) $E(Y\mid X=1)=\int_1^\infty 3ye^{3-3y}dy=-ye^{3-3y}\Big|_1^\infty+\int_1^\infty e^{3-3y}dy=4/3$

d) $\int_1^2 3e^{3-3y}dy=1-e^{-3}=0.9502$ for $0<y$, $f_Y(2)=\frac{15}{2}e^{-6}$

$f_{X|Y=2}(y)=\frac{15e^{-2x-6}}{\frac{15}{2}e^{-6}}=2e^{-2x}$

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



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

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

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

Therefore, c = 1/7.5=2/15

5-50 a) $P(X < 0.5, Y < 0.5) = \int_0^{0.5}\int_0^{0.5} \tfrac{1}{7.5}\,dy\,dx = \tfrac{1}{30}$

b) $P(X < 0.5) = \int_0^{0.5}\int_0^{x+1} \tfrac{1}{7.5}\,dy\,dx = \tfrac{1}{7.5}\int_0^{0.5}(x+1)dx = \tfrac{2}{15}(\tfrac{5}{8}) = \tfrac{1}{12}$

c)

$$E(X) = \int_0^1\int_0^{x+1} \tfrac{x}{7.5}\,dy\,dx + \int_1^4\int_{x-1}^{x+1} \tfrac{x}{7.5}\,dy\,dx$$

$$= \tfrac{1}{7.5}\int_0^1 (x^2 + x)dx + \tfrac{2}{7.5}\int_1^4 (x)dx = \tfrac{12}{15}(\tfrac{5}{6}) + \tfrac{2}{7.5}(7.5) = \tfrac{19}{9}$$

d)

$$E(Y) = \tfrac{1}{7.5}\int_0^1\int_0^{x+1} y\,dy\,dx + \tfrac{1}{7.5}\int_1^4\int_{x-1}^{x+1} y\,dy\,dx$$

$$= \tfrac{1}{7.5}\int_0^1 \tfrac{(x+1)^2}{2}dx + \tfrac{1}{7.5}\int_1^4 \tfrac{(x+1)^2-(x-1)^2}{2}dx$$

$$= \tfrac{1}{15}\int_0^1 (x^2 + 2x + 1)dx + \tfrac{1}{15}\int_1^4 4x\,dx$$

$$= \tfrac{1}{15}\left(\tfrac{7}{3}\right) + \tfrac{1}{15}\left(30\right) = \tfrac{97}{45}$$

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

5-52 Let X, Y, and Z denote the time until a problem on line 1, 2, and 3, respectively.

a)

$$P(X > 40, Y > 40, Z > 40) = [P(X > 40)]^3$$

because the random variables are independent with the same distribution. Now,

$P(X > 40) = \int_{40}^{\infty} \frac{1}{40} e^{-x/40}\,dx = -e^{-x/40}\Big|_{40}^{\infty} = e^{-1}$ and the answer is

$\left(e^{-1}\right)^3 = e^{-3} = 0.0498$.

b) $P(20 < X < 40, 20 < Y < 40, 20 < Z < 40) = [P(20 < X < 40)]^3$ and

$P(20 < X < 40) = -e^{-x/40}\Big|_{20}^{40} = e^{-0.5} - e^{-1} = 0.2387$.

The answer is $0.2387^3 = 0.0136$.

c.) The joint density is not needed because the process is represented by three independent exponential distributions. Therefore, the probabilities may be multiplied.

5-53 a.) μ=3.2 λ=1/3.2

$$P(X>5,Y>5)=(1/10.24)\int_5^\infty\int_5^\infty e^{-\frac{x}{3.2}-\frac{y}{3.2}}dydx = 3.2\int_5^\infty e^{-\frac{x}{3.2}}\left(e^{-\frac{5}{3.2}}\right)dx$$

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

$$P(X>10,Y>10)=(1/10.24)\int_{10}^\infty\int_{10}^\infty e^{-\frac{x}{3.2}-\frac{y}{3.2}}dydx=3.2\int_{10}^\infty e^{-\frac{x}{3.2}}\left(e^{-\frac{10}{3.2}}\right)dx$$

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

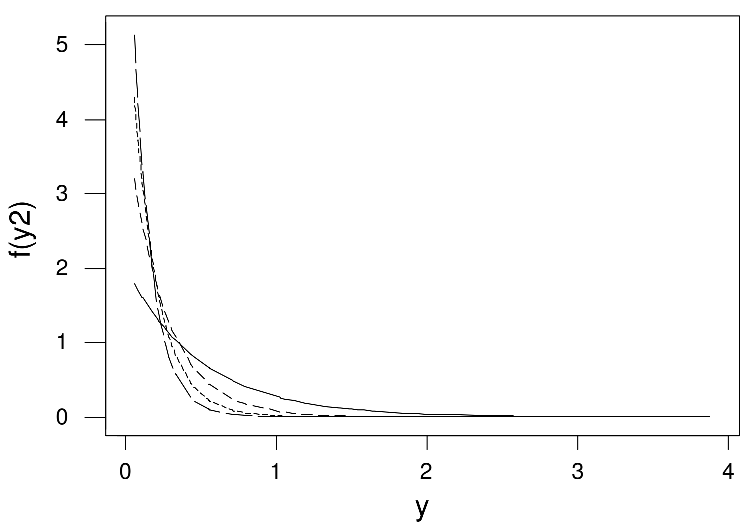
b.) Let *X* denote the number of orders in a 5-minute interval. Then *X* is a Poisson random variable with λ=5/3.2 = 1.5625.

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

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

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

5-54 a) $f_{Y|X=x}(y)$, for x = 2, 4, 6, 8



b) $P(Y < 2 | X = 2) = \int_0^2 2e^{-2y} dy = 0.9817$

c) $E(Y | X = 2) = \int_0^\infty 2ye^{-2y} dy = 1/2$ (using integration by parts)

d) $E(Y | X = x) = \int_0^\infty xye^{-xy} dy = 1/x$ (using integration by parts)

e) Use $f_X(x) = \dfrac{1}{b-a} = \dfrac{1}{10}$, $f_{Y|X}(x, y) = xe^{-xy}$, and the relationship $f_{Y|X}(x, y) = \dfrac{f_{XY}(x, y)}{f_X(x)}$

Therefore, $xe^{-xy} = \dfrac{f_{XY}(x, y)}{1/10}$ and $f_{XY}(x, y) = \dfrac{xe^{-xy}}{10}$

f) $f_Y(y) = \int_0^{10} \dfrac{xe^{-xy}}{10} dx = \dfrac{1 - 10ye^{-10y} - e^{-10y}}{10y^2}$ (using integration by parts)

Section 5-4

5-55. a) $P(X < 0.5) = \int_0^{0.5}\int_0^1\int_0^1 (8xyz) dz dy dx = \int_0^{0.5}\int_0^1 (4xy) dy dx = \int_0^{0.5} (2x) dx = x^2 \Big|_0^{0.5} = 0.25$

b)

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

$$= \int_0^{0.5}\int_0^{0.5} (4xy) dy dx = \int_0^{0.5} (0.5x) dx = \tfrac{x^2}{4} \Big|_0^{0.5} = 0.0625$$

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) $E(X) = \int_0^1\int_0^1\int_0^1 (8x^2 yz) dz dy dx = \int_0^1 (2x^2) dx = \tfrac{2x^3}{3} \Big|_0^1 = 2/3$

5-56 a) $P(X < 0.5 | Y = 0.5)$ is the integral of the conditional density $f_{X|Y}(x)$. Now,

$f_{X|0.5}(x) = \dfrac{f_{XY}(x, 0.5)}{f_Y(0.5)}$ and $f_{XY}(x, 0.5) = \int_0^1 (8xyz) dz = 4xy$ for 0 < x < 1 and 0 < y < 1. Also, $f_Y(y) = \int_0^1\int_0^1 (8xyz) dz dx = 2y$ for 0 < y < 1.

Therefore, $f_{X|0.5}(x) = \dfrac{2x}{1} = 2x$ for 0 < x < 1.

Then, $P(X < 0.5 | Y = 0.5) = \int_0^{0.5} 2x dx = 0.25$.

b) $P(X < 0.5, Y < 0.5 | Z = 0.8)$ is the integral of the conditional density of X and Y.
Now, $f_Z(z) = 2z$ for $0 < z < 1$ as in part a. and

$$f_{XY|Z}(x,y) = \frac{f_{XYZ}(x,y,z)}{f_Z(z)} = \frac{8xy(0.8)}{2(0.8)} = 4xy \text{ for } 0 < x < 1 \text{ and } 0 < y < 1.$$

Then, $P(X < 0.5, Y < 0.5 | Z = 0.8) = \int_0^{0.5}\int_0^{0.5}(4xy)dydx = \int_0^{0.5}(x/2)dx = \frac{1}{16} = 0.0625$

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

5-58 a) $\iint_{x^2+y^2\le 4}\int_0^4 cdzdydx$ = the volume of a cylinder with a base of radius 2 and a height of 4 = $(\pi 2^2)4 = 16\pi$. Therefore, $c = \dfrac{1}{16\pi}$

b) $P(X^2 + Y^2 \le 1)$ equals the volume of a cylinder of radius $\sqrt{2}$ and a height of 4 (=8π) times *c*. Therefore, the answer is $\dfrac{8\pi}{16\pi} = 1/2$.

c) P(Z < 2) equals half the volume of the region where $f_{XYZ}(x,y,z)$ is positive times 1/c. Therefore, the answer is 0.5.

d) $E(X) = \int_{-2}^{2}\int_{-\sqrt{4-x^2}}^{\sqrt{4-x^2}}\int_0^4 \frac{x}{c}dzdydx = \frac{1}{c}\int_{-2}^{2}\left[4xy\Big|_{-\sqrt{4-x^2}}^{\sqrt{4-x^2}}\right]dx = \frac{1}{c}\int_{-2}^{2}(8x\sqrt{4-x^2})dx$. Using substitution, $u = 4 - x^2$, $du = -2x\,dx$, and $E(X) = \frac{1}{c}\int 4\sqrt{u}du = \frac{-4}{c}\frac{2}{3}(4-x^2)^{\frac{3}{2}}\Big|_{-2}^{2} = 0$.

5-59. a) $f_{X|1}(x) = \dfrac{f_{XY}(x,1)}{f_Y(1)}$ *and* $f_{XY}(x,y) = \frac{1}{c}\int_0^4 dz = \frac{4}{c} = \frac{1}{4\pi}$ *for* $x^2 + y^2 < 4$.

Also, $f_Y(y) = \frac{1}{c}\int_{-\sqrt{4-y^2}}^{\sqrt{4-y^2}}\int_0^4 dzdx = \frac{8}{c}\sqrt{4-y^2}$ for $-2 < y < 2$.

Then, $f_{X|y}(x) = \dfrac{4/c}{\frac{8}{c}\sqrt{4-y^2}}$ evaluated at y = 1. That is, $f_{X|1}(x) = \frac{1}{2\sqrt{3}}$ for $-\sqrt{3} < x < \sqrt{3}$.

Therefore, $P(X < 1 | Y < 1) = \int_{-\sqrt{3}}^{1} \frac{1}{2\sqrt{3}}dx = \dfrac{1+\sqrt{3}}{2\sqrt{3}} = 0.7887$

b) $f_{XY|1}(x,y) = \frac{f_{XYZ}(x,y,1)}{f_Z(1)}$ and $f_Z(z) = \int_{-2}^{2} \int_{-\sqrt{4-x^2}}^{\sqrt{4-x^2}} \frac{1}{c} dydx = \int_{-2}^{2} \frac{2}{c}\sqrt{4-x^2}dx$

Because $f_Z(z)$ is a density over the range 0 < z < 4 that does not depend on Z, $f_Z(z)$ =1/4 for

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

Then, $P(x^2 + y^2 < 1 | Z = 1) = \frac{area\ in\ x^2 + y^2 < 1}{4\pi} = 1/4$.

5-60 $f_{Z|xy}(z) = \frac{f_{XYZ}(x,y,z)}{f_{XY}(x,y)}$ and from part 5-59 a., $f_{XY}(x,y) = \frac{1}{4\pi}$ *for* $x^2 + y^2 < 4$.

Therefore, $f_{Z|xy}(z) = \frac{\frac{1}{16\pi}}{\frac{1}{4\pi}} = 1/4$ for $0 < z < 4$.

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

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

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

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

5-62 a.) $P(X < 0.5, Y < 0.5, Z < 0.5) = 6\int_0^1 \int_0^{1-x} \int_0^{1-x-y} dzdydx \Rightarrow$ The conditions x>0.5, y>0.5, z>0.5 and $x+y+z$<1 make a space that is a cube with a volume of 0.125. Therefore the probability of $P(X < 0.5, Y < 0.5, Z < 0.5) = 6(0.125) = 0.75$

b.)

$$P(X < 0.5, Y < 0.5) = \int_0^{0.5}\int_0^{0.5} 6(1-x-y)dydx = \int_0^{0.5} \left(6y - 6xy - 3y^2\right)\Big|_0^{0.5} dx$$

$$= \int_0^{0.5} \left(\frac{9}{4} - 3x \right) dx = \left(\frac{9}{4}x - \frac{3}{2}x^2 \right)\Big|_0^{0.5} = 3/4$$

c.)

$$P(X < 0.5) = 6\int_0^{0.5}\int_0^{1-x}\int_0^{1-x-y} dzdydx = \int_0^{0.5}\int_0^{1-x} 6(1-x-y)dydx = \int_0^{0.5} 6(y - xy - \frac{y^2}{2})\Big|_0^{1-x}$$

$$= \int_0^{0.5} 6(\frac{x^2}{2} - x + \frac{1}{2})dx = \left(x^3 - 3x^2 + 3x\right)\Big|_0^{0.5} = 0.875$$

d.)

$$E(X)=6\int_0^1\int_0^{1-x}\int_0^{1-x-y} xdzdydx=\int_0^1\int_0^{1-x}6x(1-x-y)dydx=\int_0^1 6x(y-xy-\frac{y^2}{2})\Big|_0^{1-x}$$

$$=\int_0^1 6(\frac{x^3}{2}-x^2+\frac{x}{2})dx=\left(\frac{3x^4}{4}-2x^3+\frac{3x^2}{2}\right)\Big|_0^1=0.25$$

5-63 a.)

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

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

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 $0<y<1$.

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

5-64 Let X denote the production yield on a day. Then,

$P(X>1400)=P(Z>\frac{1400-1500}{\sqrt{10000}})=P(Z>-1)=0.84134$.

a) Let Y denote the number of days out of five such that the yield exceeds 1400. Then, by independence, Y has a binomial distribution with $n=5$ and $p=0.8413$. Therefore, the answer is $P(Y=5)=\binom{5}{5}0.8413^5(1-0.8413)^0=0.4215$.

b) As in part a., the answer is

$$P(Y\geq 4)=P(Y=4)+P(Y=5)$$

$$=\binom{5}{4}0.8413^4(1-0.8413)^1+0.4215=0.8190$$

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

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

5-66 a) Let X denote the grams of luminescent ink. Then,

$P(X < 1.14) = P(Z < \frac{1.14-1.2}{0.3}) = P(Z < -2) = 0.022750$.

Let Y denote the number of bulbs in the sample of 25 that have less than 1.14 grams. Then, by independence, Y has a binomial distribution with n = 25 and p = 0.022750. Therefore, the answer is

$P(Y \ge 1) = 1 - P(Y = 0) = \binom{25}{0} 0.02275^0 (0.97725)^{25} = 1 - 0.5625 = 0.4375$.

b)

$$P(Y \le 5) = P(Y = 0) + P(Y = 1) + P(Y = 2) + (P(Y = 3) + P(Y = 4) + P(Y = 5)$$

$$= \binom{25}{0} 0.02275^0 (0.97725)^{25} + \binom{25}{1} 0.02275^1 (0.97725)^{24} + \binom{25}{2} 0.02275^2 (0.97725)^{23}$$

$$+ \binom{25}{3} 0.02275^3 (0.97725)^{22} + \binom{25}{4} 0.02275^4 (0.97725)^{21} + \binom{25}{5} 0.02275^5 (0.97725)^{20}.$$

$$= 0.5625 + 0.3274 + 0.09146 + 0.01632 + 0.002090 + 0.0002043 = 0.99997 \cong 1$$

c.) $P(Y = 0) = \binom{25}{0} 0.02275^0 (0.97725)^{25} = 0.5625$

d.) The lamps are normally and independently distributed, therefore, the probabilities can be multiplied.

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

$$\text{E(XY)} = [1 \times 3 \times (1/8)] + [1 \times 4 \times (1/4)] + [2 \times 5 \times (1/2)] + [4 \times 6 \times (1/8)]$$

$$= 75/8 = 9.375$$

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

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

$V(Y) = 3^2(1/8) + 4^2(1/4) + 5^2(1/2) + 6^2(1/8) - (37/8)^2 = 0.7344$

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

5-68 $E(X) = -1(1/8) + (-0.5)(1/4) + 0.5(1/2) + 1(1/8) = 0.125$

$E(Y) = -2(1/8) + (-1)(1/4) + 1(1/2) + 2(1/8) = 0.25$

$\text{E(XY)} = [-1 \times -2 \times (1/8)] + [-0.5 \times -1 \times (1/4)] + [0.5 \times 1 \times (1/2)] + [1 \times 2 \times (1/8)] = 0.875$

V(X) = 0.4219

V(Y) = 1.6875

$\sigma_{XY} = 0.875 - (0.125)(0.25) = 0.8438$

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} = \frac{0.8438}{\sqrt{0.4219}\sqrt{1.6875}} = 1$$

5-69.

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

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

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

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

5-70 $E(X) = 0(0.01) + 1(0.99) = 0.99$

$E(Y) = 0(0.02) + 1(0.98) = 0.98$

$\text{E(XY)} = [0 \times 0 \times (0.002)] + [0 \times 1 \times (0.0098)] + [1 \times 0 \times (0.0198)] + [1 \times 1 \times (0.9702)] = 0.9702$

V(X) = $0.99 - 0.99^2 = 0.0099$

V(Y) = $0.98 - 0.98^2 = 0.0196$

$\sigma_{XY} = 0.9702 - (0.99)(0.98) = 0$

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} = \frac{0}{\sqrt{0.0099}\sqrt{0.0196}} = 0$$

5-71 $E(X_1) = np_1 = 20(1/3) = 6.67$

$E(X_2) = np_2 = 20(1/3) = 6.67$

$V(X_1) = np_1(1-p_1) = 20(1/3)(2/3) = 4.44$

$V(X_2) = np_2(1-p_2) = 20(1/3)(2/3) = 4.44$

$E(X_1X_2) = n(n-1)p_1p_2 = 20(19)(1/3)(1/3) = 42.22$

$$\sigma_{XY} = 42.22 - 6.67^2 = -2.267 \;\; and \;\; \rho_{XY} = \frac{-2.267}{\sqrt{(4.44)(4.44)}} = -0.51$$

The sign is negative.

5-72 From Exercise 5-40, c=8/81.
From Exercise 5-41, $E(X)$ = 12/5, and $E(Y)$ = 8/5

$$E(XY) = \frac{8}{81}\int_0^3\int_0^x xy(xy)dydx = \frac{8}{81}\int_0^3\int_0^x x^2y^2dydx = \frac{8}{81}\int_0^3 \frac{x^3}{3}x^2dx = \frac{8}{81}\int_0^3 \frac{x^5}{3}dx$$

$$= \left(\frac{8}{81}\right)\left(\frac{3^6}{18}\right) = 4$$

$$\sigma_{xy} = 4 - \left(\frac{12}{5}\right)\left(\frac{8}{5}\right) = 0.16$$

$$E(X^2) = 6 \qquad E(Y^2) = 3$$

$$V(x) = 0.24, \qquad V(Y) = 0.44$$

$$\rho = \frac{0.16}{\sqrt{0.24}\sqrt{0.44}} = 0.4924$$

5-73. Similarly to 5-49, $c = 2/19$

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

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

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

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

$$E(X^2) = 8.7632 \qquad E(Y^2) = 9.11403$$

$$V(x) = 1.930, \qquad V(Y) = 2.0968$$

$$\rho = \frac{1.852}{\sqrt{1.930}\sqrt{2.062}} = 0.9206$$

5-74

$$E(X) = \int_0^1 \int_0^{x+1} \tfrac{x}{7.5}\, dydx + \int_1^4 \int_{x-1}^{x+1} \tfrac{x}{7.5}\, dydx$$

$$= \tfrac{1}{7.5}\int_0^1 (x^2 + x)dx + \tfrac{2}{7.5}\int_1^4 (x)dx = \tfrac{12}{15}(\tfrac{5}{6}) + \tfrac{2}{7.5}(7.5) = \tfrac{19}{9}$$

$E(X^2)$=222,222.2

$V(X)=222222.2-(333.33)^2$=111,113.31

$E(Y^2)$=1,055,556

V(Y)=361,117.11

$$E(XY) = 6\times 10^{-6}\int_0^{\infty}\int_x^{\infty} xye^{-.001x-.002y}\, dydx = 388{,}888.9$$

$$\sigma_{xy} = 388{,}888.9 - (333.33)(833.33) = 111{,}115.01$$

$$\rho = \frac{111{,}115.01}{\sqrt{111113.31}\sqrt{361117.11}} = 0.5547$$

5-75. a) E(X) = 1 E(Y) = 1

$$E(XY) = \int_0^{\infty}\int_0^{\infty} xye^{-x-y}\, dxdy$$

$$= \int_0^{\infty} xe^{-x}\, dx \int_0^{\infty} ye^{-y}\, dy$$

$$= E(X)E(Y)$$

Therefore, $\sigma_{XY} = \rho_{XY} = 0$.

5-76. Suppose the correlation between X and Y is ρ. for constants a, b, c, and d, what is the correlation between the random variables U = aX+b and V = cY+d?

Now, E(U) = a E(X) + b and E(V) = c E(Y) + d.

Also, U - E(U) = a[X - E(X)] and V - E(V) = c[Y - E(Y)]. Then,

$$\sigma_{UV} = E\{[U - E(U)][V - E(V)]\} = acE\{[X - E(X)][Y - E(Y)]\} = ac\sigma_{XY}$$

Also, $\sigma_U^2 = E[U - E(U)]^2 = a^2E[X - E(X)]^2 = a^2\sigma_X^2$ and $\sigma_V^2 = c^2\sigma_Y^2$. Then,

$$\rho_{UV} = \frac{ac\sigma_{XY}}{\sqrt{a^2\sigma_X^2}\sqrt{c^2\sigma_Y^2}} = \begin{cases} \rho_{XY} & \text{if a and c are of the same sign} \\ -\rho_{XY} & \text{if a and c differ in sign} \end{cases}$$

5-77 $E(X) = -1(1/4) + 1(1/4) = 0$

$E(Y) = -1(1/4) + 1(1/4) = 0$

$\mathrm{E(XY)} = [-1\times 0\times (1/4)] + [-1\times 0\times (1/4)] + [1\times 0\times (1/4)] + [0\times 1\times (1/4)] = 0$

V(X) = 1/2

V(Y) = 1/2

$\sigma_{XY} = 0 - (0)(0) = 0$

$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} = \frac{0}{\sqrt{1/2}\sqrt{1/2}} = 0$

The correlation is zero, but *X* and *Y* are not independent, since, for example, if y=0, X must be –1 or 1.

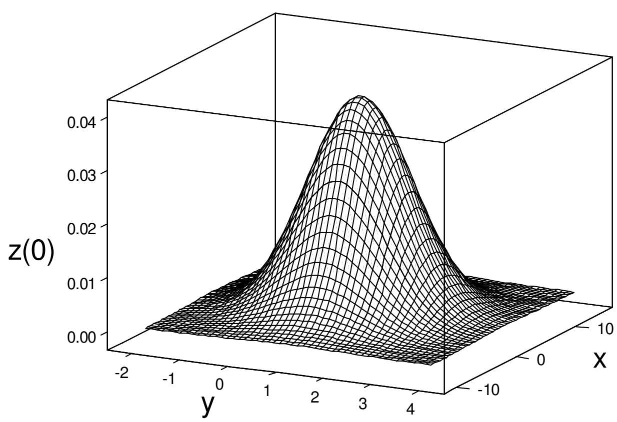
5-78 If X and Y are independent, then $f_{XY}(x,y) = f_X(x) f_Y(y)$ and the range of (X, Y) is rectangular. Therefore,

$E(XY) = \iint xy f_X(x) f_Y(y) dx dy = \int x f_X(x) dx \int y f_Y(y) dy = E(X)E(Y)$

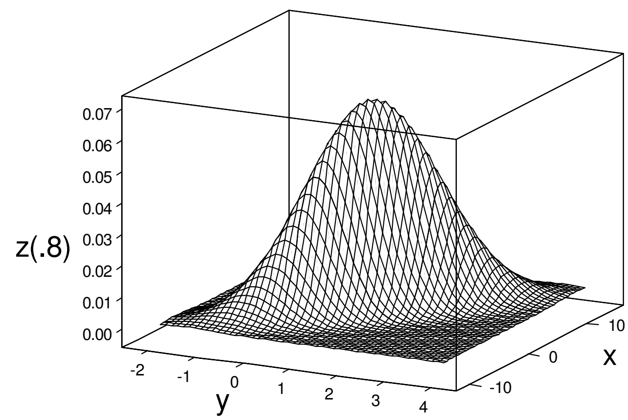
hence $\sigma_{XY}=0$

Section 5-6

5-79 a.)

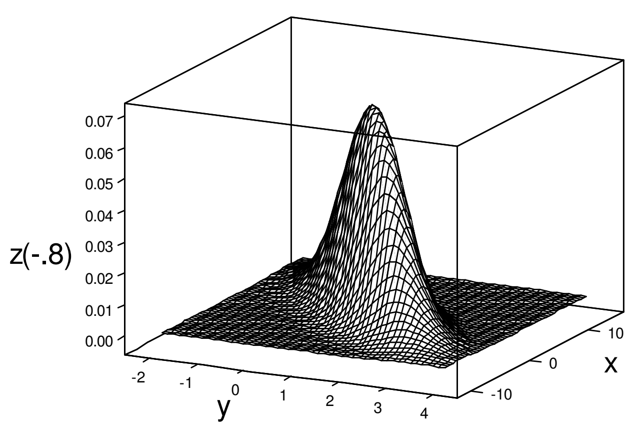


b.)



0.07
0.06
0.05
0.04
0.03
0.02
0.01
0.00
z(.8)
-2
-1
0
1
2
3
4
y
10
0
-10
x

c)



0.07
0.06
0.05
0.04
0.03
0.02
0.01
0.00
z(-.8)
-2
-1
0
1
2
3
4
y
10
0
-10
x

5-80 Because $\rho = 0$ and X and Y are normally distributed, X and Y are independent. Therefore,
P(2.95 < X < 3.05, 7.60 < Y < 7.80) = P(2.95 < X < 3.05) P(7.60 < Y < 7.80) =
$P(\frac{2.95-3}{0.04} < Z < \frac{3.05-3}{0.04})P(\frac{7.60-7.70}{0.08} < Z < \frac{7.80-7.70}{0.08}) = 0.7887^2 = 0.6220$

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

$\mu_X = 0.1$mm $\sigma_X=0.00031$mm $\mu_Y = 0.23$mm $\sigma_Y=0.00017$mm

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

5-82 a) By completing the square in the numerator of the exponent of the bivariate normal PDF, the joint PDF can be written as

$$f_{Y|X=x} = \frac{f_{XY}(x,y)}{f_X(x)} = \frac{\frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} e^{-\frac{\left[\frac{1}{\sigma_Y^2}\left[(y-(\mu_Y+\rho\frac{\sigma_Y}{\sigma_X}(x-\mu_x))\right]^2+(1-\rho^2)\left(\frac{x-\mu_x}{\sigma_x}\right)^2\right]}{2(1-\rho^2)}}}{\frac{1}{\sqrt{2\pi}\sigma_x} e^{-\frac{\left[\frac{x-\mu_x}{\sigma_x}\right]^2}{2}}}$$

Also, $f_x(x) = \frac{1}{\sqrt{2\pi}\sigma_x} e^{-\frac{\left[\frac{x-\mu_x}{\sigma_x}\right]^2}{2}}$ By definition,

$$f_{Y|X=x} = \frac{f_{XY}(x,y)}{f_X(x)} = \frac{\frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} e^{-\frac{\left[\frac{1}{\sigma_Y^2}\left[(y-(\mu_Y+\rho\frac{\sigma_Y}{\sigma_X}(x-\mu_x))\right]^2+(1-\rho^2)\left(\frac{x-\mu_x}{\sigma_x}\right)^2\right]}{2(1-\rho^2)}}}{\frac{1}{\sqrt{2\pi}\sigma_x} e^{-\frac{\left[\frac{x-\mu_x}{\sigma_x}\right]^2}{2(1-\rho^2)}}}$$

$$= \frac{1}{\sqrt{2\pi}\sigma_y\sqrt{1-\rho^2}} e^{-\frac{\frac{1}{\sigma_Y^2}\left[(y-(\mu_Y+\rho\frac{\sigma_Y}{\sigma_X}(x-\mu_x))\right]^2}{2(1-\rho^2)}}$$

Now $f_{Y|X=x}$ is in the form of a normal distribution.

b) E(Y|X=x) = $\mu_y + \rho\frac{\sigma_y}{\sigma_x}(x-\mu_x)$. This answer can be seen from part 5-82a. Since the PDF is in the form of a normal distribution, then the mean can be obtained from the exponent.

c) V(Y|X=x) = $\sigma_y^2(1-\rho^2)$. This answer can be seen from part 5-82a. Since the PDF is in the form of a normal distribution, then the variance can be obtained from the exponent.

5-83

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x,y)dxdy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left[\frac{1}{2\pi\sigma_X\sigma_Y}e^{-\frac{1}{2}\left[\frac{(x-\mu_X)^2}{\sigma_X{}^2}+\frac{(y-\mu_Y)^2}{\sigma_Y{}^2}\right]}\right]dxdy =$$

$$\int_{-\infty}^{\infty}\left[\frac{1}{\sqrt{2\pi}\sigma_X}e^{-\frac{1}{2}\left[\frac{(x-\mu_X)^2}{\sigma_X{}^2}\right]}\right]dx\int_{-\infty}^{\infty}\left[\frac{1}{\sqrt{2\pi}\sigma_Y}e^{-\frac{1}{2}\left[\frac{(y-\mu_Y)^2}{\sigma_Y{}^2}\right]}\right]dy$$

and each of the last two integrals is recognized as the integral of a normal probability density function from $-\infty$ to ∞. That is, each integral equals one. Since $f_{XY}(x,y) = f(x)f(y)$ then *X* and *Y* are independent.

5-84 Let $f_{XY}(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}}e^{-\frac{\left[\left(\frac{X-\mu_X}{\sigma_X}\right)^2-\frac{2\rho(X-\mu_X)(Y-\mu_X)}{\sigma_X\sigma_Y}+\left(\frac{Y-\mu_Y}{\sigma_Y}\right)^2\right]}{2(1-\rho^2)}}$

Completing the square in the numerator of the exponent we get:

$$\left[\left(\frac{X-\mu_X}{\sigma_X}\right)^2-\frac{2\rho(X-\mu_X)(Y-\mu_X)}{\sigma_X\sigma_Y}+\left(\frac{Y-\mu_Y}{\sigma_Y}\right)^2\right] = \left[\left(\frac{Y-\mu_Y}{\sigma_Y}\right)-\rho\left(\frac{X-\mu_X}{\sigma_X}\right)\right]^2 + (1-\rho^2)\left(\frac{X-\mu_X}{\sigma_X}\right)^2$$

But,

$$\left(\frac{Y-\mu_Y}{\sigma_Y}\right)-\rho\left(\frac{X-\mu_X}{\sigma_X}\right) = \frac{1}{\sigma_Y}\left[(Y-\mu_Y)-\rho\frac{\sigma_Y}{\sigma_X}(X-\mu_x)\right] = \frac{1}{\sigma_Y}\left[(Y-(\mu_Y+\rho\frac{\sigma_Y}{\sigma_X}(X-\mu_x))\right]$$

Substituting into $f_{XY}(x,y)$, we get

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}}e^{-\frac{\left[\frac{1}{\sigma_Y^2}\left[y-(\mu_Y+\rho\frac{\sigma_Y}{\sigma_X}(x-\mu_x))\right]^2+(1-\rho^2)\left(\frac{x-\mu_x}{\sigma_x}\right)^2\right]}{2(1-\rho^2)}}dydx$$

$$=\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma_x}e^{-\frac{1}{2}\left(\frac{x-\mu_x}{\sigma_x}\right)^2}dx\times\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma_y\sqrt{1-\rho^2}}e^{-\left[\frac{\left(y-(\mu_y+\rho\frac{\sigma_y}{\sigma_x}(x-\mu_x))\right)^2}{2\sigma_x^2(1-\rho^2)}\right]}dy$$

The integrand in the second integral above is in the form of a normally distributed random variable. By definition of the integral over this function, the second integral is equal to 1:

$$\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma_x}e^{-\frac{1}{2}\left(\frac{x-\mu_x}{\sigma_x}\right)^2}dx\times\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma_y\sqrt{1-\rho^2}}e^{-\left[\frac{\left(y-(\mu_y+\rho\frac{\sigma_y}{\sigma_x}(x-\mu_x))\right)^2}{2\sigma_x^2(1-\rho^2)}\right]}dy$$

$$=\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma_x}e^{-\frac{1}{2}\left(\frac{x-\mu_x}{\sigma_x}\right)^2}dx\times 1$$

The remaining integral is also the integral of a normally distributed random variable and therefore, it also integrates to 1, by definition. Therefore,

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}f_{XY}(x,y)=1$$

5-85

$$f_X(x)=\int_{-\infty}^{\infty}\left[\frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}}e^{\frac{-0.5}{1-\rho^2}\left[\frac{(x-\mu_X)^2}{\sigma_X^2}-\frac{2\rho(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y}+\frac{(y-\mu_Y)^2}{\sigma_Y^2}\right]}\right]dy$$

$$=\frac{1}{\sqrt{2\pi}\sigma_X}e^{\frac{-0.5}{1-\rho^2}\frac{(x-\mu_X)^2}{\sigma_X^2}}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma_Y\sqrt{1-\rho^2}}\;e^{\frac{-0.5}{1-\rho^2}\left\{\left[\frac{(y-\mu_Y)}{\sigma_Y}-\frac{\rho(x-\mu_X)}{\sigma_X}\right]^2-\left[\frac{\rho(x-\mu_X)}{\sigma_X}\right]^2\right\}}dy$$

$$=\frac{1}{\sqrt{2\pi}\sigma_X}e^{-0.5\frac{(x-\mu_X)^2}{\sigma_X^2}}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma_Y\sqrt{1-\rho^2}}\;e^{\frac{-0.5}{1-\rho^2}\left[\frac{(y-\mu_Y)}{\sigma_Y}-\frac{\rho(x-\mu_X)}{\sigma_X}\right]^2}dy$$

The last integral is recognized as the integral of a normal probability density with mean $\mu_Y+\frac{\sigma_Y\rho(x-\mu_X)}{\sigma_X}$ and variance $\sigma_Y{}^2(1-\rho^2)$. Therefore, the last integral equals one and the requested result is obtained.

5-86 $E(X)=\mu_X, E(Y)=\mu_Y, V(X)=\sigma_X^2, \ \ and \ \ V(Y)=\sigma_Y^2$. Also,

$$E(X-\mu_X)(Y-\mu_Y)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{(x-\mu_X)(y-\mu_Y)e^{\frac{-0.5}{1-\rho^2}\left[\frac{(x-\mu_X)^2}{\sigma_X^2}-\frac{2\rho(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y}+\frac{(y-\mu_Y)^2}{\sigma_Y^2}\right]}}{2\pi\sigma_X\sigma_Y(1-\rho^2)^{1/2}}dxdy$$

Let $u=\frac{x-\mu_X}{\sigma_X}$ *and* $v=\frac{y-\mu_Y}{\sigma_Y}$. Then,

$$E(X-\mu_X)(Y-\mu_Y)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{uve^{\frac{-0.5}{1-\rho^2}\left[u^2-2\rho uv+v^2\right]}}{2\pi(1-\rho^2)^{1/2}}\ \sigma_X\sigma_Y dudv$$

$$=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{uve^{\frac{-0.5}{1-\rho^2}\left\{[u-\rho v]^2+(1-\rho^2)v^2\right\}}}{2\pi(1-\rho^2)^{1/2}}\ \sigma_X\sigma_Y dudv$$

The integral with respect to u is recognized as a constant times the mean of a normal random variable with mean ρv and variance $1-\rho^2$. Therefore,

$$E(X-\mu_X)(Y-\mu_Y)=\int_{-\infty}^{\infty}\frac{v}{\sqrt{2\pi}}\ e^{-0.5v^2}\rho v\sigma_X\sigma_Y dv=\rho\sigma_X\sigma_Y\int_{-\infty}^{\infty}\frac{v^2}{\sqrt{2\pi}}\ e^{-0.5v^2}dv.$$

The last integral is recognized as the variance of a normal random variable with mean 0 and variance 1. Therefore, $E(X-\mu_X)(Y-\mu_Y)=\rho\sigma_X\sigma_Y$ and the correlation between X and Y is ρ.

Section 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,
$P(2X+3Y<30)=P(Z<\frac{30-30}{\sqrt{97}})=P(Z<0)=0.5$
d) $P(2X+3Y<40)=P(Z<\frac{40-30}{\sqrt{97}})=P(Z<1.02)=0.8461$

5-88 Y = 10X and E(Y) =10E(X) = 50mm.
$V(Y)=10^2V(X)=25mm^2$

5-89 a) Let T denote the total thickness. Then, T = X + Y and E(T) = 4 mm,
V(T) = $0.1^2 + 0.1^2 = 0.02 mm^2$, and $\sigma_T = 0.1414$ mm.

b)

$$P(T > 4.3) = P\left(Z > \frac{4.3-4}{0.1414}\right) = P(Z > 2.12)$$
$$= 1 - P(Z < 2.12) = 1 - 0.983 = 0.0170$$

5-90 a) X~N(0.1, 0.00031) and Y~N(0.23, 0.00017) Let T denote the total thickness. Then, T = X + Y and E(T) = 0.33 mm,
V(T) = $0.00031^2 + 0.00017^2 = 1.25x10^{-7} mm^2$, and $\sigma_T = 0.000354$ mm.

$$P(T < 0.2337) = P\left(Z < \frac{0.2337 - 0.33}{0.000354}\right) = P(Z < -272) \cong 0$$

b)

$$P(T > 0.2405) = P\left(Z > \frac{0.2405 - 0.33}{0.000345}\right) = P(Z > -253) = 1 - P(Z < 253) \cong 1$$

5-91. Let D denote the width of the casing minus the width of the door. Then, D is normally distributed.

a) E(D) = 1/8 V(D) = $(\frac{1}{8})^2 + (\frac{1}{16})^2 = \frac{5}{256}$

b) $P(D > \frac{1}{4}) = P(Z > \frac{\frac{1}{4}-\frac{1}{8}}{\sqrt{5/256}}) = P(Z > 0.89) = 0.187$

c) $P(D < 0) = P(Z < \frac{0-\frac{1}{8}}{\sqrt{5/256}}) = P(Z < -0.89) = 0.187$

5-92 D = A - B - C

a) E(D) = 10 - 2 - 2 = 6 mm

$V(D) = 0.1^2 + 0.05^2 + 0.05^2 = 0.015mm^2$

$\sigma_D = 0.1225mm$

b) P(D < 5.9) = P(Z < $\frac{5.9-6}{0.1225}$) = P(Z < -0.82) = 0.206.

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 $n = 135.72 \cong 136$.

5-94 Let $\overline{X}$ denote the average thickness of 10 wafers. Then, E($\overline{X}$) = 10 and V($\overline{X}$) = 0.1.

a) $P(9 < \overline{X} < 11) = P(\frac{9-10}{\sqrt{0.1}} < Z < \frac{11-10}{\sqrt{0.1}}) = P(-3.16 < Z < 3.16) = 0.998$.

The answer is 1 − 0.998 = 0.002

b) $P(\overline{X} > 11) = 0.01$ *and* $\sigma_{\overline{X}} = 1/\sqrt{n}$.

Therefore, $P(\overline{X} > 11) = P(Z > \frac{11-10}{1/\sqrt{n}}) = 0.01$, $\frac{11-10}{1/\sqrt{n}}$ = 2.33 and n = 5.43 which is rounded up to 6.

c.) $P(\overline{X} > 11) = 0.0005$ *and* $\sigma_{\overline{X}} = \sigma/\sqrt{10}$.

Therefore, $P(\overline{X} > 11) = P(Z > \frac{11-10}{\sigma/\sqrt{10}}) = 0.0005$, $\frac{11-10}{\sigma/\sqrt{10}}$ = 3.29

$\sigma = \sqrt{10}/3.29 = 0.9612$

5-95. X~N(160, 900)

a) Let Y = 25X, E(Y)=25E(X)=4000, V(Y) = 25^2(900)=562500

P(Y>4300) =

$$P\left(Z > \frac{4300-4000}{\sqrt{562500}}\right) = P(Z > 0.4) = 1 - P(Z < 0.4) = 1 - 0.6554 = 0.3446$$

b.) c) P(Y > x) = 0.0001 implies that $P\left(Z > \frac{x-4000}{\sqrt{562500}}\right) = 0.0001$.

Then $\frac{x-4000}{750}$ = 3.72 and $x = 6790$.

<u>Supplemental Exercises</u>

5-96 The sum of $\sum_x \sum_y f(x, y) = 1$, $\left(1/4\right) + \left(1/8\right) + \left(1/8\right) + \left(1/4\right) + \left(1/4\right) = 1$

and $f_{XY}(x, y) \geq 0$

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

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

5-98 a) $f_X(x) = \sum_y f_{XY}(x,y)$ *and* $f_X(0) = 3/8,\ f_X(1) = 3/8,\ f_X(2) = 1/4$.

b) $f_{Y|1}(y) = \dfrac{f_{XY}(1,y)}{f_X(1)}$ *and* $f_{Y|1}(0) = \frac{1/8}{3/8} = 1/3,\ f_{Y|1}(1) = \frac{1/4}{3/8} = 2/3$.

c) $E(Y \mid X = 1) = \sum_{x=1} y f_{XY}(1,y) = 0(1/3) + 1(2/3) = 2/3$

d) Because the range of (X, Y) is not rectangular, X and Y are not independent.

e.) E(*XY*) = 1.25, E(*X*) = E(*Y*)= 0.875 V(X) = V(Y) = 0.6094
COV(*X*,*Y*)=E(*XY*)-E(*X*)E(*Y*)= $1.25-0.875^2=0.4844$

$$\rho_{XY} = \frac{0.4844}{\sqrt{0.6094}\sqrt{0.6094}} = 0.7949$$

5-99 a.) $P(X = 2, Y = 4, Z = 14) = \dfrac{20!}{2!4!14!} 0.10^2 0.20^4 0.70^{14} = 0.0631$

b.) $P(X = 0) = 0.10^0 0.90^{20} = 0.1216$

c.) $E(X) = np_1 = 20(0.10) = 2$

$V(X) = np_1(1 - p_1) = 20(0.10)(0.9) = 1.8$

d.) $f_{X|Z=z}(X \mid Z = 19) \dfrac{f_{XZ}(x,z)}{f_Z(z)}$

$$f_{XZ}(xz) = \frac{20!}{x!z!(20-x-z)!} 0.1^x 0.2^{20-x-z} 0.7^z$$

$$f_Z(z) = \frac{20!}{z!(20-z)!} 0.3^{20-z} 0.7^z$$

$$f_{X|Z=z}(X \mid Z = 19) \frac{f_{XZ}(x,z)}{f_Z(z)} = \frac{(20-z)!}{x!(20-x-z)!} \frac{0.1^x 0.2^{20-x-z}}{0.3^{20-z}} = \frac{(20-z)!}{x!(20-x-z)!} \left(\frac{1}{3}\right)^x \left(\frac{2}{3}\right)^{20-x-z}$$

Therefore, *X* is a binomial random variable with *n*=20-*z* and *p*=1/3. When *z*=19,

$f_{X|19}(0) = \dfrac{2}{3}$ and $f_{X|19}(1) = \dfrac{1}{3}$.

e.) $E(X \mid Z = 19) = 0\left(\dfrac{2}{3}\right) + 1\left(\dfrac{1}{3}\right) = \dfrac{1}{3}$

5-100 Let X, Y, and Z denote the number of bolts rated high, moderate, and low.
Then, X, Y, and Z have a multinomial distribution.

a) $P(X = 12, Y = 6, Z = 2) = \dfrac{20!}{12!6!2!} 0.6^{12} 0.3^6 0.1^2 = 0.0560$.

b) Because X, Y, and Z are multinomial, the marginal distribution of Z is binomial with
n = 20 and p = 0.1.

c) E(Z) = np = 20(0.1) = 2.

5-101. a) $f_{Z|16}(z) = \frac{f_{XZ}(16,z)}{f_X(16)}$ and $f_{XZ}(x,z) = \frac{20!}{x!z!(20-x-z)!}0.6^x 0.3^{(20-x-z)} 0.1^z$ for

$x + z \le 20$ *and* $0 \le x, 0 \le z$. Then,

$$f_{Z|16}(z) = \frac{\frac{20!}{16!z!(4-z)!}0.6^{16}0.3^{(4-z)}0.1^z}{\frac{20!}{16!4!}0.6^{16}0.4^4} = \frac{4!}{z!(4-z)!}\left(\frac{0.3}{0.4}\right)^{4-z}\left(\frac{0.1}{0.4}\right)^z$$

for $0 \le z \le 4$. That is the distribution of Z given X = 16 is binomial with n = 4 and p = 0.25.

b) From part a., E(Z) = 4 (0.25) = 1.

c) Because the conditional distribution of Z given X = 16 does not equal the marginal distribution of Z, X
and Z are not independent.

5-102 Let X, Y, and Z denote the number of calls answered in two rings or less, three or four rings, and five rings or more, respectively.

a) $P(X = 8, Y = 1, Z = 1) = \frac{10!}{8!1!1!}0.7^8 0.25^1 0.05^1 = 0.0649$

b) Let W denote the number of calls answered in four rings or less. Then, W is a binomial random variable with n = 10 and p = 0.95.

Therefore, P(W = 10) = $\binom{10}{10}0.95^{10}0.05^0 = 0.5987$.

c) E(W) = 10(0.95) = 9.5.

5-103 a) $f_{Z|8}(z) = \frac{f_{XZ}(8,z)}{f_X(8)}$ and $f_{XZ}(x,z) = \frac{10!}{x!z!(10-x-z)!}0.70^x 0.25^{(10-x-z)} 0.05^z$ for

$x + z \le 10$ *and* $0 \le x, 0 \le z$. Then,

$$f_{Z|8}(z) = \frac{\frac{10!}{8!z!(2-z)!}0.70^8 0.25^{(2-z)}0.05^z}{\frac{10!}{8!2!}0.70^8 0.30^2} = \frac{2!}{z!(2-z)!}\left(\frac{0.25}{0.30}\right)^{2-z}\left(\frac{0.05}{0.30}\right)^z$$

for $0 \le z \le 2$. That is Z is binomial with n =2 and p = 0.05/0.30 = 1/6.

b) E(Z) given X = 8 is 2(1/6) = 1/3.

c) Because the conditional distribution of Z given X = 8 does not equal the marginal distribution of Z, X and Z are not independent.

5-104 $\int_0^3\int_0^2 cx^2 y\,dy\,dx = \int_0^3 cx^2 \left.\frac{y^2}{2}\right|_0^2 dx = 2c\left.\frac{x^3}{3}\right|_0^3 = 18c$. Therefore, c = 1/18.

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)

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

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

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

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

5-106 a) $f_X(x)=\int_0^2\frac{1}{18}x^2 y\,dy=\frac{1}{9}x^2 \text{ for } 0<x<3$

b) $f_{Y|X}(y)=\frac{f_{XY}(1,y)}{f_X(1)}=\frac{\frac{1}{18}y}{\frac{1}{9}}=\frac{y}{2}$ for $0<y<2$.

c) $f_{X|1}(x)=\frac{f_{XY}(x,1)}{f_Y(1)}=\frac{\frac{1}{18}x^2}{f_Y(1)}$ and $f_Y(y)=\int_0^3\frac{1}{18}x^2 y\,dx=\frac{y}{2}\ \text{for } 0<y<2$.

Therefore, $f_{X|1}(x)=\frac{\frac{1}{18}x^2}{1/2}=\frac{1}{9}x^2$ for $0<x<3$.

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 $f_{XYZ}(x,y,z)=\frac{1}{4\pi}$ for $x^2+y^2\le 1$ and $0<z<4$.

a) The region $X^2+Y^2\le 0.5$ is a cylinder of radius $\sqrt{0.5}$ and height 4. Therefore,

$P(X^2+Y^2\le 0.5)=\frac{4(0.5\pi)}{4\pi}=1/2$.

b) The region $X^2+Y^2\le 0.5$ and $0<z<2$ is a cylinder of radius $\sqrt{0.5}$ and height 2. Therefore,

$P(X^2+Y^2\le 0.5, Z<2)=\frac{2(0.5\pi)}{4\pi}=1/4$

c) $f_{XY|1}(x,y) = \dfrac{f_{XYZ}(x,y,1)}{f_Z(1)}$ and $f_Z(z) = \iint\limits_{x^2+y^2\le 1} \frac{1}{4\pi} dydx = 1/4$

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

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

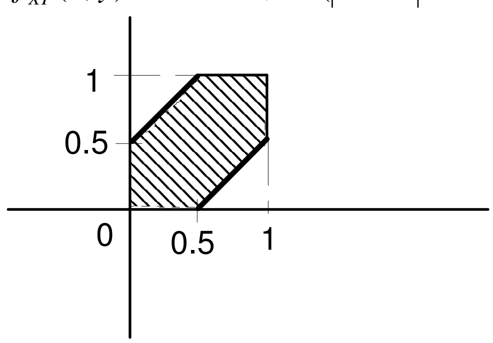
5-108 a) $f_{Z|0,0}(z) = \dfrac{f_{XYZ}(0,0,z)}{f_{XY}(0,0)}$ and $f_{XY}(x,y) = \int_0^4 \frac{1}{4\pi} dz = 1/\pi$ for $x^2 + y^2 \le 1$. Then,

$f_{Z|0,0}(z) = \dfrac{1/4\pi}{1/\pi} = 1/4$ for 0 < z < 4 and $\mu_{Z|0,0} = 2$.

b) $f_{Z|xy}(z) = \dfrac{f_{XYZ}(x,y,z)}{f_{XY}(x,y)} = \dfrac{1/4\pi}{1/\pi} = 1/4$ for 0 < z < 4. Then, E(Z) given X = x and Y = y is

$\int_0^4 \frac{z}{4} dz = 2$.

5-109. $f_{XY}(x,y) = c$ for 0 < x < 1 and 0 < y < 1. Then, $\int_0^1 \int_0^1 c\,dxdy = 1$ and c = 1. Because $f_{XY}(x,y)$ is constant, $P(|X-Y| < 0.5)$ is the area of the shaded region below



That is, $P(|X-Y| < 0.5)$ = 3/4.

5-110 a) Let $X_1, X_2, \ldots, X_6$ denote the lifetimes of the six components, respectively. Because of independence,

$P(X_1 > 5000, X_2 > 5000, \ldots, X_6 > 5000) = P(X_1 > 5000)P(X_2 > 5000)\ldots P(X_6 > 5000)$

If X is exponentially distributed with mean θ, then $\lambda = \frac{1}{\theta}$ and

$$P(X > x) = \int_x^\infty \frac{1}{\theta} e^{-t/\theta} dt = -e^{-t/\theta}\Big|_x^\infty = e^{-x/\theta}$$. Therefore, the answer is

$e^{-5/8} e^{-0.5} e^{-0.5} e^{-0.25} e^{-0.25} e^{-0.2} = e^{-2.325} = 0.0978$.

b) The probability that at least one component lifetime exceeds 25,000 hours is the same as 1 minus the probability that none of the component lifetimes exceed 25,000 hours. Thus,
$1\text{-}P(X_a<25{,}000, X_2<25{,}000, \ldots, X_6<25{,}000)=1\text{-}P(X_1<25{,}000)\ldots P(X_6<25{,}000)$
$=1-(1-e^{-25/8})(1-e^{-2.5})(1-e^{-2.5})(1-e^{-1.25})(1-e^{-1.25})(1-e^{-1})=1-.2592=0.7408$

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.

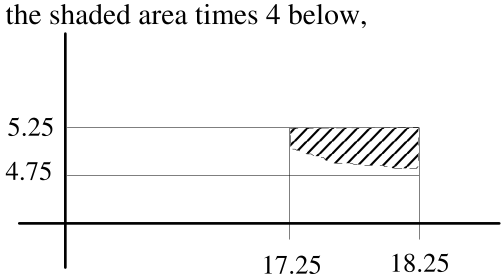
c) E(Z) = 10(0.3) = 3.

5-112 a) Let $\overline{X}$ denote the mean weight of the 25 bricks in the sample. Then, E($\overline{X}$) = 3 and $\sigma_{\overline{X}}=\frac{0.25}{\sqrt{25}}=0.05$. Then, P($\overline{X}$ < 2.95) = P(Z < $\frac{2.95-3}{0.05}$) = P (Z < -1) = 0.159.

b) $P(\overline{X}>x)=P(Z>\frac{x-3}{.05})=0.99$. So, $\frac{x-3}{0.05}=$ -2.33 and x=2.8835.

5-113. a.)

Because $\int_{17.75}^{18.25}\int_{4.75}^{5.25} c\,dy\,dx=0.25c$, c = 4. The area of a panel is XY and P(XY > 90) is the shaded area times 4 below,



That is, $\int_{17.75}^{18.25}\int_{90/x}^{5.25} 4\,dy\,dx=4\int_{17.75}^{18.25} 5.25-\frac{90}{x}\,dx=4(5.25x-90\ln x\Big|_{17.75}^{18.25})=0.499$

b. The perimeter of a panel is 2X+2Y and we want P(2X+2Y >46)

$\int_{17.75}^{18.25}\int_{23-x}^{5.25} 4\,dy\,dx=4\int_{17.75}^{18.25} 5.25-(23-x)\,dx$

$=4\int_{17.75}^{18.25}(-17.75+x)\,dx=4(-17.75x+\frac{x^2}{2}\Big|_{17.75}^{18.25})=0.5$

5-114 a)Let X denote the weight of a piece of candy and X~N(0.1, 0.01). Each package has 16 candies, then P is the total weight of the package with 16 pieces and E(P) = 16(0.1)=1.6 ounces and V(P) = $16^2(0.01^2)$=0.0256 ounces^2

b) $P(P < 1.6) = P(Z < \frac{1.6-1.6}{0.16}) = P(Z < 0) = 0.5$.

c) Let Y equal the total weight of the package with 17 pieces, E(Y) = 17(0.1)=1.7 ounces and V(Y) = $17^2(0.01^2)$=0.0289 ounces^2

$P(Y < 1.6) = P(Z < \frac{1.6-1.7}{\sqrt{0.0289}}) = P(Z < -0.59) = 0.2776$.

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) $P(\overline{X} > 60) = P(Z > \frac{60-45}{30/\sqrt{10}}) = P(Z > 1.58) = 0.057$

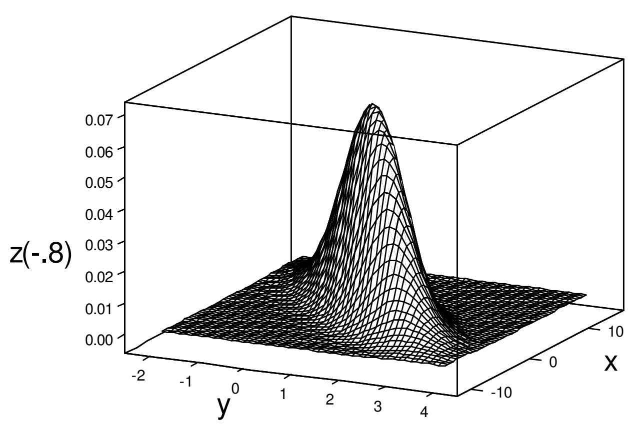
b) Let Y denote the total time to locate 10 parts. Then, Y > 600 if and only if $\overline{X}$ > 60. Therefore, the answer is the same as part a.

5-116 a) Let Y denote the weight of an assembly. Then, E(Y) = 4 + 5.5 + 10 + 8 = 27.5 and V(Y)=$0.4^2 + 0.5^2 + 0.2^2 + 0.5^2 = 0.7$.

$P(Y > 29.5) = P(Z > \frac{29.5-27.5}{\sqrt{0.7}}) = P(Z > 2.39) = 0.0084$

b) Let $\overline{X}$ denote the mean weight of 8 independent assemblies. Then, E($\overline{X}$) = 27.5 and V($\overline{X}$) = 0.7/8 = 0.0875. Also, $P(\overline{X} > 29) = P(Z > \frac{29-27.5}{\sqrt{0.0875}}) = P(Z > 5.07) = 0$.

5-117



5-118

$$f_{XY}(x,y)=\frac{1}{1.2\pi}e^{\left[\frac{-1}{0.72}\{(x-1)^2-1.6(x-1)(y-2)+(y-2)^2\}\right]}$$

$$f_{XY}(x,y)=\frac{1}{2\pi\sqrt{.36}}e^{\left[\frac{-1}{2(0.36)}\{(x-1)^2-1.6(x-1)(y-2)+(y-2)^2\}\right]}$$

$$f_{XY}(x,y)=\frac{1}{2\pi\sqrt{1-.8^2}}e^{\left[\frac{-1}{2(1-.8^2)}\{(x-1)^2-2(.8)(x-1)(y-2)+(y-2)^2\}\right]}$$

$E(X)=1,\ E(Y)=2\ \ V(X)=1\ V(Y)=1$ and $\rho = 0.8$

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

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

$V(T)=V(X_1)+V(X_2)+2Cov(X_1X_2)=0.01+0.04+2(0.014)=0.078mm^2$

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

b.) $P(T<1)=P\left(Z<\frac{1-1.5}{\sqrt{0.078}}\right)=P(Z<-1.79)=0.0367$

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

5-120 Let T denote the total thickness. Then, $T = X_1 + X_2 + X_3$ and

a) E(T) = 0.5+1+1.5 =3 mm

$V(T)=V(X_1)+V(X_2)+V(X_3)+2Cov(X_1X_2)+2Cov(X_2X_3)+$

$2Cov(X_1X_3)=0.01+0.04+0.09+2(0.014)+2(0.03)+2(0.009)=0.246mm^2$

where $Cov(XY)=\rho\sigma_X\sigma_Y$

b.) $P(T<1.5)=P\left(Z<\frac{1.5-3}{0.246}\right)=P(Z<-6.10)\cong 0$

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)= 0.05 (or 5 if given in terms of percent)

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 than investing 2million in the first security.

Mind-Expanding Exercises

5-122. By the independence,

$$P(X_1 \in A_1, X_2 \in A_2, ..., X_p \in A_p) = \int_{A_1} \int_{A_2} ... \int_{A_p} f_{X_1}(x_1) f_{X_2}(x_2) ... f_{X_p}(x_p) dx_1 dx_2 ... dx_p$$

$$= \left[\int_{A_1} f_{X_1}(x_1) dx_1 \right] \left[\int_{A_2} f_{X_2}(x_2) dx_2 \right] ... \left[\int_{A_p} f_{X_p}(x_p) dx_p \right]$$

$$= P(X_1 \in A_1) P(X_2 \in A_2) ... P(X_p \in A_p)$$

5-123 $E(Y) = c_1\mu_1 + c_2\mu_2 + ... + c_p\mu_p.$

Also,

$$V(Y) = \int \left[c_1x_1 + c_2x_2 + ... + c_px_p - (c_1\mu_1 + c_2\mu_2 + ... + c_p\mu_p) \right]^2 f_{X_1}(x_1) f_{X_2}(x_2) ... f_{X_p}(x_p) dx_1 dx_2 ... dx_p$$

$$= \int \left[c_1(x_1 - \mu_1) + ... + c_p(x_p - \mu_p) \right]^2 f_{X_1}(x_1) f_{X_2}(x_2) ... f_{X_p}(x_p) dx_1 dx_2 ... dx_p$$

Now, the cross-term

$$\int c_1 c_2 (x_1 - \mu_1)(x_2 - \mu_2) f_{X_1}(x_1) f_{X_2}(x_2) ... f_{X_p}(x_p) dx_1 dx_2 ... dx_p$$

$$= c_1c_2 \left[\int (x_1 - \mu_1) f_{X_1}(x_1) dx_1 \right] \left[\int (x_2 - \mu_2) f_{X_2}(x_2) dx_2 \right] = 0$$

from the definition of the mean. Therefore, each cross-term in the last integral for V(Y) is zero and

$$V(Y) = \left[\int c_1^2 (x_1 - \mu_1)^2 f_{X_1}(x_1) dx_1 \right] ... \left[\int c_p^2 (x_p - \mu_p)^2 f_{X_p}(x_p) dx_p \right]$$

$$= c_1^2 V(X_1) + ... + c_p^2 V(X_p).$$

5-124 $\int_0^a \int_0^b f_{XY}(x,y)dydx = \int_0^a \int_0^b cdydx = cab$. Therefore, c = 1/ab. Then,

$f_X(x) = \int_0^b cdy = \frac{1}{a}$ for 0 < x < a, and $f_Y(y) = \int_0^a cdx = \frac{1}{b}$ for 0 < y < b. Therefore,

$f_{XY}(x,y)=f_X(x)f_Y(y)$ for all x and y and X and Y are independent.

5-125 $f_X(x) = \int_0^b g(x)h(y)dy = g(x)\int_0^b h(y)dy = kg(x)$ $where\ k = \int_0^b h(y)dy$. Also,

$f_Y(y) = lh(y)$ $where\ l = \int_0^a g(x)dx$. Because $f_{XY}(x,y)$ is a probability density

function, $\int_0^a \int_0^b g(x)h(y)dydx = \left[\int_0^a g(x)dx\right]\left[\int_0^b h(y)dy\right] = 1$. Therefore, kl = 1 and

$f_{XY}(x,y)=f_X(x)f_Y(y)$ for all x and y.

Section 5-8 on CD

S5-1. $f_Y(y) = \frac{1}{4}$ at y = 3, 5, 7, 9 from Theorem S5-1.

S5-2. Because $X \geq 0$, the transformation is one-to-one; that is $y = x^2$ and $x = \sqrt{y}$. From Theorem S5-2,

$f_Y(y) = f_X(\sqrt{y}) = \binom{3}{\sqrt{y}} p^{\sqrt{y}} (1-p)^{3-\sqrt{y}}$ for y = 0, 1, 4, 9.

If p = 0.25, $f_Y(y) = \binom{3}{\sqrt{y}} (0.25)^{\sqrt{y}} (0.75)^{3-\sqrt{y}}$ for y = 0, 1, 4, 9.

S5-3. a) $f_Y(y) = f_X\left(\frac{y-10}{2}\right)\left(\frac{1}{2}\right) = \frac{y-10}{72}$ for $10 \leq y \leq 22$

b) $E(Y) = \int_{10}^{22} \frac{y^2 - 10y}{72} dy = \frac{1}{72}\left(\frac{y^3}{3} - \frac{10y^2}{2}\right)\Big|_{10}^{22} = 18$

S5-4. Because y = -2 ln x, $e^{-\frac{y}{2}} = x$. Then, $f_Y(y) = f_X(e^{\frac{-y}{2}})\left|-\frac{1}{2}e^{\frac{-y}{2}}\right| = \frac{1}{2}e^{\frac{-y}{2}}$ for $0 \leq e^{\frac{-y}{2}} \leq 1$ or

$y \geq 0$, which is an exponential distribution (which equals a chi-square distribution with k = 2 degrees of freedom).

S5-5. a) Let Q = R. Then,

$$p = i^2 r \qquad \text{and} \qquad i = \sqrt{\tfrac{p}{q}}$$
$$q = r \qquad\qquad\qquad r = q$$

$$J = \begin{vmatrix} \frac{\partial i}{\partial p} & \frac{\partial i}{\partial q} \\ \frac{\partial r}{\partial p} & \frac{\partial r}{\partial q} \end{vmatrix} = \begin{vmatrix} \frac{1}{2}(pq)^{-1/2} & -\frac{1}{2}p^{1/2}q^{-3/2} \\ 0 & 1 \end{vmatrix} = \tfrac{1}{2}(pq)^{-1/2}$$

$$f_{PQ}(p,q) = f_{IR}(\sqrt{\tfrac{p}{q}}, q)\tfrac{1}{2}(pq)^{-1/2} = 2\left(\sqrt{\tfrac{p}{q}}\right)\tfrac{1}{2}(pq)^{-1/2} = q^{-1}$$

$$\text{for } 0 \le \sqrt{\tfrac{p}{q}} \le 1,\ 0 \le q \le 1$$

That is, for $0 \le p \le q,\ 0 < q \le 1$.

$$f_P(p) = \int_p^1 q^{-1} dq = -\ln \text{p} \quad \text{for } 0 < p \le 1.$$

b) $E(P) = -\int_0^1 p \ln p \ dp$. Let u = ln p and dv = p dp. Then, du = 1/p and

$v = \dfrac{p^2}{2}$. Therefore, $E(P) = -(\ln p)\frac{p^2}{2}\Big|_0^1 + \int_0^1 \frac{p}{2} dp = \frac{p^2}{4}\Big|_0^1 = \dfrac{1}{4}$

S5-6. a) If $y = x^2$, then $x = \sqrt{y}$ for $x \ge 0$ and $y \ge 0$. Thus, $f_Y(y) = f_X(\sqrt{y})\frac{1}{2}y^{-\frac{1}{2}} = \dfrac{e^{-\sqrt{y}}}{2\sqrt{y}}$ for y > 0.

b) If $y = x^{1/2}$, then $x = y^2$ for $x \ge 0$ and $y \ge 0$. Thus, $f_Y(y) = f_X(y^2)2y = 2ye^{-y^2}$ for y > 0.

c) If y = ln x, then $x = e^y$ for $x \ge 0$. Thus, $f_Y(y) = f_X(e^y)e^y = e^y e^{-e^y} = e^{y - e^y}$ for $-\infty < y < \infty$.

S5-7. a) Now, $\int_0^\infty av^2 e^{-bv} dv$ must equal one. Let u = bv, then $1 = a\int_0^\infty (\frac{u}{b})^2 e^{-u} \dfrac{du}{b} = \dfrac{a}{b^3}\int_0^\infty u^2 e^{-u} du.$ From the definition of the gamma function the last expression is $\dfrac{a}{b^3}\Gamma(3) = \dfrac{2a}{b^3}$. Therefore, $a = \dfrac{b^3}{2}$.

b) If $w = \dfrac{mv^2}{2}$, then $v = \sqrt{\dfrac{2w}{m}}$ for $v \ge 0,\ \ w \ge 0$.

$$f_W(w) = f_V\left(\sqrt{\tfrac{2w}{m}}\right)\frac{dv}{dw} = \frac{b^3 2w}{2m} e^{-b\sqrt{\frac{2w}{m}}}(2mw)^{-1/2}$$
$$= \frac{b^3 m^{-3/2}}{\sqrt{2}} w^{1/2} e^{-b\sqrt{\frac{2w}{m}}}$$

for $w \ge 0$.

S5-8. If $y = e^x$, then x = ln y for $1 \le x \le 2$ and $e^1 \le y \le e^2$. Thus, $f_Y(y) = f_X(\ln y)\frac{1}{y} = \frac{1}{y}$ for $1 \le \ln y \le 2$. That is, $f_Y(y) = \frac{1}{y}$ for $e \le y \le e^2$.

S5-9. Now $P(Y \le a) = P(X \ge u(a)) = \int_{u(a)}^{\infty} f_X(x)dx$. By changing the variable of integration from x to y by using x = u(y), we obtain $P(Y \le a) = \int_{a}^{-\infty} f_X(u(y))u'(y)dy$ because as x tends to ∞, y = h(x) tends to - ∞. Then, $P(Y \le a) = \int_{-\infty}^{a} f_X(u(y))(-u'(y))dy$. Because h(x) is decreasing, $u'(y)$ is negative. Therefore, $|u'(y)| = -u'(y)$ and Theorem S5-1 holds in this case also.

S5-10. If y = $(x-2)^2$, then x = $2-\sqrt{y}$ for $0 \le x \le 2$ and x = $2+\sqrt{y}$ for $2 \le x \le 4$. Thus,

$$f_Y(y) = f_X(2-\sqrt{y})\,|-\tfrac{1}{2}y^{-1/2}| + f_X(2+\sqrt{y})\,|\tfrac{1}{2}y^{-1/2}|$$
$$= \frac{2-\sqrt{y}}{16\sqrt{y}} + \frac{2+\sqrt{y}}{16\sqrt{y}}$$
$$= \left(\tfrac{1}{4}\right)y^{-1/2} \text{ for } 0 \le y \le 4$$

S5-11. a) Let $a = s_1 s_2$ and y = s_1. Then, s_1 = y, $s_2 = \frac{a}{y}$ and

$$J = \begin{vmatrix} \frac{\partial s_1}{\partial a} & \frac{\partial s_1}{\partial y} \\ \frac{\partial s_2}{\partial a} & \frac{\partial s_2}{\partial y} \end{vmatrix} = \begin{vmatrix} 0 & 1 \\ \frac{1}{y} & -\frac{a}{2}y^{-3/2} \end{vmatrix} = -\frac{1}{y}.$$ Then,

$f_{AY}(a, y) = f_{S_1S_2}(y, \frac{a}{y})(\frac{1}{y}) = 2y(\frac{a}{8y})(\frac{1}{y}) = \frac{a}{4y}$ for $0 \le y \le 1$ and $0 \le \frac{a}{y} \le 4$. That is, for $0 \le y \le 1$ and $0 \le a \le 4y$.

b) $f_A(a) = \int_{a/4}^{1} \frac{a}{4y}dy = -\frac{a}{4}\ln(\frac{a}{4})$ for $0 < a \le 4$.

S5-12. Let r = v/i and s = i. Then, i = s and v = rs

$$J = \begin{vmatrix} \frac{\partial i}{\partial r} & \frac{\partial i}{\partial s} \\ \frac{\partial v}{\partial r} & \frac{\partial v}{\partial s} \end{vmatrix} = \begin{vmatrix} 0 & 1 \\ s & r \end{vmatrix} = s$$

$f_{RS}(r,s) = f_{IV}(s, rs)s = e^{-rs}\text{s}$ for rs ≥ 0 and $1 \leq s \leq 2$. That is,

$f_{RS}(r,s) = se^{-rs}$ for $1 \leq s \leq 2$ and $r \geq 0$.

Then, $f_R(r) = \int_1^2 se^{-rs}ds$. Let u = s and dv = e^{-rs} ds. Then, du = ds and v = $\frac{-e^{-rs}}{r}$

Then,

$$f_R(r) = -s\frac{-e^{-rs}}{r}\Bigg|_1^2 + \int_1^2 \frac{-e^{-rs}}{r}ds = \frac{e^{-r} - 2e^{-2r}}{r} - \frac{e^{-rs}}{r^2}\Bigg|_1^2$$

$$= \frac{e^{-r} - 2e^{-2r}}{r} + \frac{e^{-r} - e^{-2r}}{r^2}$$

$$= \frac{e^{-r}(r+1) - e^{-2r}(2r+1)}{r^2}$$

for r > 0.

Section 5-9 on CD

S5-13 . a) $E(e^{tX}) = \sum_{x=1}^{m} \frac{e^{tx}}{m} = \frac{1}{m}\sum_{x=1}^{m}(e^t)^x = \frac{(e^t)^{m+1} - e^t}{m(e^t - 1)} = \frac{e^t(1-e^{tm})}{m(1-e^t)}$

b) $M(t) = \frac{1}{m}e^t(1-e^{tm})(1-e^t)^{-1}$ and

$$\frac{dM(t)}{dt} = \frac{1}{m}\left\{e^t(1-e^{tm})(1-e^t)^{-1} + e^t(-me^{tm})(1-e^t)^{-1} + e^t(1-e^{tm})(-1)(1-e^t)^{-2}(-e^t)\right\}$$

$$\frac{dM(t)}{dt} = \frac{e^t}{m(1-e^t)}\left\{1 - e^{tm} - me^{tm} + \frac{(1-e^{tm})e^t}{1-e^t}\right\}$$

$$= \frac{e^t}{m(1-e^t)^2}\left\{1 - e^{tm} - me^{tm} + me^{(m+1)t}\right\}$$

Using L'Hospital's rule,

$$\lim_{t\to 0}\frac{dM(t)}{dt} = \lim_{t\to 0}\frac{e^t}{m}\lim_{t\to 0}\frac{-me^{tm} - m^2e^{tm} + m(m+1)e^{(m+1)t}}{-2(1-e^t)e^t}$$

$$= \lim_{t\to 0}\frac{e^t}{m}\lim_{t\to 0}\frac{-m^2e^{tm} - m^3e^{tm} + m(m+1)^2e^{(m+1)t}}{-2(1-e^t)e^t - 2e^t(-e^t)}$$

$$= \frac{1}{m}\times\frac{m(m+1)^2 - m^2 - m^3}{2} = \frac{m^2+m}{2m} = \frac{m+1}{2}$$

Therefore, E(X) = $\frac{m+1}{2}$.

$$\frac{d^2M(t)}{dt^2} = \frac{d^2}{dt^2}\sum_{x=1}^{m} e^{tx}\frac{1}{m} = \frac{1}{m}\sum_{x=1}^{m}\frac{d^2}{dt^2}\left(1+tx+\frac{(tx)^2}{2}+\ldots\right)$$

$$= \frac{1}{m}\sum_{x=1}^{m}(x^2 + term\sin volvingpow ersoft)$$

Thus,

$$\left.\frac{d^2M(t)}{dt^2}\right|_{t=0} = \frac{1}{m}\left(\frac{m(m+1)(m+2)}{6}\right) = \frac{(m+1)(2m+2)}{6}$$

Then,

$$V(X) = \frac{2m^2+3m+1}{6} - \frac{(m+1)^2}{4} = \frac{4m^2+6m+2-3m^2-6m-3}{12}$$

$$= \frac{m^2-1}{12}$$

S5-14. a) $E(e^{tX}) = \sum_{x=0}^{\infty} e^{tx}\frac{e^{-\lambda}\lambda^x}{x!} = e^{-\lambda}\sum_{x=0}^{\infty}\frac{(\lambda e^t)^x}{x!} = e^{-\lambda}e^{\lambda e^t} = e^{\lambda(e^t-1)}$

b) $\frac{dM(t)}{dt} = \lambda e^t e^{\lambda(e^t-1)}$

$$\left.\frac{dM(t)}{dt}\right|_{t=0} = \lambda = E(X)$$

$$\frac{d^2M(t)}{dt^2} = \lambda^2 e^{2t} e^{\lambda(e^t-1)} + \lambda e^t e^{\lambda(e^t-1)}$$

$$\left.\frac{d^2M(t)}{dt^2}\right|_{t=0} = \lambda^2 + \lambda$$

$$V(X) = \lambda^2 + \lambda - \lambda^2 = \lambda$$

S5-15 a) $E(e^{tX}) = \sum_{x=1}^{\infty} e^{tx}(1-p)^{x-1}p = \frac{p}{1-p}\sum_{x=1}^{\infty}[e^t(1-p)]^x$

$$= \frac{e^t(1-p)}{1-(1-p)e^t}\left(\frac{p}{1-p}\right) = \frac{pe^t}{1-(1-p)e^t}$$

b) $\frac{dM(t)}{dt} = pe^t(1-(1-p)e^t)^{-2}(1-p)e^t + pe^t(1-(1-p)e^t)^{-1}$

$$= p(1-p)e^{2t}(1-(1-p)e^t)^{-2} + pe^t(1-(1-p)e^t)^{-1}$$

$$\left.\frac{dM(t)}{dt}\right|_{t=0} = \frac{1-p}{p} + 1 = \frac{1}{p} = E(X)$$

$$\frac{d^2M(t)}{dt^2} = p(1-p)e^{2t}2(1-(1-p)e^t)^{-3}(1-p)e^t + p(1-p)(1-(1-p)e^t)^{-2}2e^{2t}$$

$$+ pe^t(1-(1-p)e^t)^{-2}(1-p)e^t + pe^t(1-(1-p)e^t)^{-1}$$

$$\left.\frac{d^2M(t)}{dt^2}\right|_{t=0} = \frac{2(1-p)^2}{p^2} + \frac{2(1-p)}{p} + \frac{1-p}{p} + 1 = \frac{2(1-p)^2 + 3p(1-p) + p^2}{p^2}$$

$$= \frac{2 + 2p^2 - 4p + 3p - 3p^2 + p^2}{p^2} = \frac{2-p}{p^2}$$

$$V(X) = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2}$$

S5-16. $M_Y(t) = Ee^{tY} = Ee^{t(X_1+X_2)} = Ee^{tX_1}Ee^{tX_2}$

$$= (1-2t)^{-k_1/2}(1-2t)^{-k_2/2} = (1-2t)^{-(k_1+k_2)/2}$$

Therefore, Y has a chi-square distribution with $k_1 + k_2$ degrees of freedom.

S5-17. a) $E(e^{tX}) = \int_0^{\infty} e^{tx}4xe^{-2x}dx = 4\int_0^{\infty} xe^{(t-2)x}dx$

Using integration by parts with *u* = x and *dv* = $e^{(t-2)x}dx$ and *du* = dx,

v = $\frac{e^{(t-2)x}}{t-2}$ we obtain

$$4\left(\left.\frac{xe^{(t-2)x}}{t-2}\right|_0^{\infty} - \int_0^{\infty}\frac{e^{(t-2)x}}{t-2}dx\right) = 4\left(\left.\frac{xe^{(t-2)x}}{t-2}\right|_0^{\infty} - \left.\frac{e^{(t-2)x}}{(t-2)^2}\right|_0^{\infty}\right)$$

This integral only exists for t < 2. In that case, $E(e^{tX}) = \frac{4}{(t-2)^2}$ for t < 2

b) $\frac{dM(t)}{dt} = -8(t-2)^{-3}$ and $\left.\frac{dM(t)}{dt}\right|_{t=0} = -8(-2)^{-3} = 1 = E(X)$

c) $\frac{d^2M(t)}{dt^2} = 24(t-2)^{-4}$ and $\left.\frac{d^2M(t)}{dt^2}\right|_{t=0} = \frac{24}{16} = \frac{3}{2}$. Therefore, $V(X) = \frac{3}{2} - 1^2 = \frac{1}{2}$

S5-18. a) $E(e^{tX}) = \int_{\alpha}^{\beta} \frac{e^{tx}}{\beta-\alpha} dx = \left. \frac{e^{tx}}{t(\beta-\alpha)} \right|_{\alpha}^{\beta} = \frac{e^{t\beta} - e^{t\alpha}}{t(\beta-\alpha)}$

b) $\frac{dM(t)}{dt} = \frac{e^{t\beta} - e^{t\alpha}}{-(\beta-\alpha)t^2} + \frac{\beta e^{t\beta} - \alpha e^{t\alpha}}{t(\beta-\alpha)}$

$$= \frac{(\beta t - 1)e^{t\beta} - (\alpha t - 1)e^{t\alpha}}{t^2(\beta-\alpha)}$$

Using L'Hospital's rule,

$$\lim_{t \to 0} \frac{dM(t)}{dt} = \frac{(\beta t - 1)\beta e^{t\beta} + \beta e^{t\beta} - (\alpha t - 1)\alpha e^{t\alpha} - \alpha e^{t\alpha}}{2t(\beta-\alpha)}$$

$$\lim_{t \to 0} \frac{dM(t)}{dt} = \frac{\beta^2(\beta t - 1)e^{t\beta} + \beta^2 e^{t\beta} + \beta^2 e^{t\beta} - \alpha^2(\alpha t - 1)e^{t\alpha} - \alpha^2 e^{t\alpha} - \alpha^2 e^{t\alpha}}{2(\beta-\alpha)}$$

$$\lim_{t \to 0} \frac{dM(t)}{dt} = \frac{\beta^2 - \alpha^2}{2(\beta-\alpha)} = \frac{(\beta+\alpha)}{2} = E(X)$$

$$\frac{d^2M(t)}{dt^2} = \frac{d^2}{dt^2} \int_a^b \frac{1}{b-a} e^{tx} dx = \frac{1}{b-a} \frac{d^2}{dt^2} \left(\frac{e^{tb} - e^{ta}}{t} \right)$$

$$= \frac{1}{b-a} \frac{d^2}{dt^2} \left(\frac{tb + \frac{(tb)^2}{2} + \frac{(tb)^3}{3!} - ta - \frac{(ta)^2}{2} - \frac{(ta)^3}{3!} + \ldots}{t} \right) =$$

$$= \frac{1}{b-a} \frac{d^2}{dt^2} \left(b - a + \frac{(b^2 - a^2)t}{2} + \frac{(b^3 - a^3)t^2}{3!} + \ldots \right)$$

$$= \frac{b^3 - a^3}{3(b-a)} = \frac{b^2 + ba + a^2}{3}$$

Thus,

$$V(X) = \frac{b^2 + ba + a^2}{3} - \frac{(b+a)^2}{4} = \frac{b^2 - 2ab + a^2}{12} = \frac{(b-a)^2}{12}$$

S5-19. a) $M(t) = \int_0^\infty e^{tx} \lambda e^{-\lambda x} dx = \lambda \int_0^\infty e^{(t-\lambda)x} dx$

$$= \lambda \frac{e^{(t-\lambda)x}}{t-\lambda}\Bigg|_0^\infty = \frac{-\lambda}{t-\lambda} = \frac{1}{1-\frac{t}{\lambda}} = \left(1-\tfrac{t}{\lambda}\right)^{-1} \text{ for t < λ}$$

b) $\frac{dM(t)}{dt} = (-1)\left(1-\tfrac{t}{\lambda}\right)^{-2}\left(\tfrac{-1}{\lambda}\right) = \frac{1}{\lambda\left(1-\tfrac{t}{\lambda}\right)^2}$

$$\left.\frac{dM(t)}{dt}\right|_{t=0} = \frac{1}{\lambda}$$

$$\frac{d^2M(t)}{dt^2} = \frac{2}{\lambda^2(1-\tfrac{t}{\lambda})^3}$$

$$\left.\frac{d^2M(t)}{dt^2}\right|_{t=0} = \frac{2}{\lambda^2}$$

$$V(X) = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2}$$

S5-20. a) $M(t) = \int_0^\infty e^{tx} \frac{\lambda}{\Gamma(r)} (\lambda x)^{r-1} e^{-\lambda x} dx = \frac{\lambda^r}{\Gamma(r)} \int_0^\infty x^{r-1} e^{(t-\lambda)x} dx$

Let u = (λ-t)x. Then,

$M(t) = \frac{\lambda^r}{\Gamma(r)} \int_0^\infty \left(\frac{u}{\lambda-t}\right)^{r-1} e^{-u} \frac{du}{\lambda-t} = \frac{\lambda^r \Gamma(r)}{\Gamma(r)(\lambda-t)^r} = \frac{1}{(1-\tfrac{t}{\lambda})^r} = (1-\tfrac{t}{\lambda})^{-r}$ from the definition of the gamma function for t < λ.

b) $M'(t) = -r(1-\frac{t}{\lambda})^{-r-1}(-\frac{1}{\lambda})$

$$M'(t)\big|_{t=0} = \frac{r}{\lambda} = E(X)$$

$$M''(t) = \frac{r(r+1)}{\lambda^2}(1-\tfrac{t}{\lambda})^{r-2}$$

$$M''(t)\big|_{t=0} = \frac{r(r+1)}{\lambda^2}$$

$$V(X) = \frac{r(r+1)}{\lambda^2} - \left(\frac{r}{\lambda}\right)^2 = \frac{r}{\lambda^2}$$

S5-21. a) $E(e^{tY}) = \prod_{i=1}^{n} E(e^{tX_i}) = \left(1 - \frac{t}{\lambda}\right)^{-n}$

b) From Exercise S5-20, Y has a gamma distribution with parameter λ and n.

S5-22. a) $M_Y(t) = e^{\mu_1 t + \sigma_1^2 \frac{t^2}{2} + \mu_2 t + \sigma_2^2 \frac{t^2}{2}} = e^{(\mu_1+\mu_2)t + (\sigma_1^2+\sigma_2^2)\frac{t^2}{2}}$

b) Y has a normal distribution with mean $\mu_1 + \mu_2$ and variances $\sigma_1^2 + \sigma_2^2$

S5-23. Because a chi-square distribution is a special case of the gamma distribution with $\lambda = \frac{1}{2}$ and $r = \frac{k}{2}$, from Exercise S5-20.

$$M(t) = (1-2t)^{-k/2}$$

$$M'(t) = -\frac{k}{2}(1-2t)^{-\frac{k}{2}-1}(-2) = k(1-2t)^{-\frac{k}{2}-1}$$

$$M'(t)\Big|_{t=0} = k = E(X)$$

$$M''(t) = 2k(\tfrac{k}{2}+1)(1-2t)^{-\frac{k}{2}-2}$$

$$M''(t)\Big|_{t=0} = 2k(\tfrac{k}{2}+1) = k^2 + 2k$$

$$V(X) = k^2 + 2k - k^2 = 2k$$

S5-24. a) $M(t) = M(0) + M'(0)t + M'(0)\frac{t^2}{2!} + \ldots + M^{(r)}(0)\frac{t^r}{r!} + \ldots$ by Taylor's expansion. Now, M(0) = 1 and $M^{(r)}(0) = \mu_r'$ and the result is obtained.

b) From Exercise S5-20, $M(t) = 1 + \frac{r}{\lambda}t + \frac{r(r+1)}{\lambda^2}\frac{t^2}{2!} + \ldots$

c) $\mu_1' = \frac{r}{\lambda}$ and $\mu_2' = \frac{r(r+1)}{\lambda^2}$ which agrees with Exercise S5-20.

Section 5-10 on CD

S5-25. Use Chebychev's inequality with c = 4. Then, $P(|X-10| > 4) \le \frac{1}{16}$.

S5-26. E(X) = 5 and $\sigma_X = 2.887$. Then, $P(|X-5| > 2\sigma_X) \le \frac{1}{4}$.

The actual probability is $P(|X-5| > 2\sigma_X) = P(|X-5| > 5.77) = 0$.

S5-27. E(X) = 20 and V(X) = 400. Then, $P(|X-20|>2\sigma)\le\frac{1}{4}$ and $P(|X-20|>3\sigma)\le\frac{1}{9}$ The actual probabilities are

$$P(|X-20|>2\sigma)=1-P(|X-20|<40)$$
$$=1-\int_0^{60}0.05e^{-0.05x}dx=1-\left[-e^{-0.05x}\Big|_0^{60}\right]=0.0498$$
$$P(|X-20|>3\sigma)=1-P(|X-20|<60)$$
$$=1-\int_0^{80}0.05e^{-0.05x}dx=1-\left[-e^{-0.05x}\Big|_0^{80}\right]=0.0183$$

S5-28. E(X) = 4 and $\sigma_X = 2$

$P(|X-4|\ge 4)\le\frac{1}{4}$ *and* $P(|X-4|\ge 6)\le\frac{1}{9}$. The actual probabilities are

$$P(|X-4|\ge 4)=1-P(|X-4|<4)=1-\sum_{x=1}^{7}\frac{e^{-2}2^x}{x!}=1-0.8636=0.1364$$
$$P(|X-4|\ge 6)=1-P(|X-4|<6)=1-\sum_{x=1}^{9}\frac{e^{-2}2^x}{x!}=0.000046$$

S5-29. Let $\overline{X}$ denote the average of 500 diameters. Then, $\sigma_{\overline{X}}=\frac{0.01}{\sqrt{500}}=4.47x10^{-4}$.

a) $P(|\overline{X}-\mu|\ge 4\sigma_{\overline{X}})\le\frac{1}{16}$ *and* $P(|\overline{X}-\mu|<0.0018)\ge\frac{15}{16}$. Therefore, the bound is 0.0018.

If $P(|\overline{X}-\mu|<x)=\frac{15}{16}$, then $P(\frac{-x}{\sigma_{\overline{X}}}<\frac{\overline{X}-\mu}{\sigma_{\overline{X}}}<\frac{x}{\sigma_{\overline{X}}})=0.9375.$ Then,

$P(\frac{-x}{4.47x10^{--4}}<Z<\frac{x}{4.47x10^{-4}})=0.9375.$ and $\frac{x}{4.47x10^{-4}}$ = 1.86. Therefore, x = $8.31x10^{-4}$.

S5-30. a) E(Y) = $P(|X-\mu|\ge c\sigma)$

b) Because $Y\le 1$, $(X-\mu)^2\ge(X-\mu)^2Y$

If $|X-\mu|\ge c\sigma$, then Y = 1 and $(X-\mu)^2Y\ge c^2\sigma^2Y$

If $|X-\mu|<c\sigma$, then Y = 0 and $(X-\mu)^2Y=c^2\sigma^2Y$.

c) Because $(X-\mu)^2\ge c^2\sigma^2Y$, $E[(X-\mu)^2]\ge c^2\sigma^2E(Y)$.

d) From part a., E(Y) = $P(|X-\mu|\ge c\sigma)$. From part c., $\sigma^2\ge c^2\sigma^2P(|X-\mu|\ge c\sigma)$. Therefore,

$\frac{1}{c^2}\ge P(|X-\mu|\ge c\sigma)$.

CHAPTER 6

Section 6-1

6-1. Sample average:

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

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

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

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.

6-2. Sample average:

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

Sample variance:

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

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

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

$$= \frac{6416.49}{18} = 356.47 \text{ (min)}^2$$

Sample standard deviation:

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

The sample standard deviation could also be found using

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

where

$$\sum_{i=1}^{19} (x_i - \bar{x})^2 = 6416.49$$

6-3. Sample average:

$$\bar{x} = \frac{84817}{12} = 7068.1 \quad yards$$

Sample variance:

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

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

$$s^2 = \frac{\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}}{n-1} = \frac{600057949 - \frac{(84817)^2}{12}}{12-1}$$

$$= \frac{564324.92}{11} = 51302.265 \quad (yards)^2$$

Sample standard deviation:

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

The sample standard deviation could also be found using

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

where

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

Dot Diagram: (rounding was used to create the dot diagram)

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

6-4. Sample mean:

$$\bar{x} = \frac{\sum_{i=1}^{18} x_i}{18} = \frac{2272}{18} = 126.22 \text{ kN}$$

Sample variance:

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

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

$$s^2 = \frac{\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}}{n-1} = \frac{298392 - \frac{(2272)^2}{18}}{18-1}$$

$$= \frac{11615.11}{17} = 683.24 \text{ (kN)}^2$$

Sample standard deviation:

$$s = \sqrt{683.24} = 26.14 \text{ kN}$$

The sample standard deviation could also be found using

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

where

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

Dot Diagram:

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

6-5. Sample average:

$$\bar{x} = \frac{351.8}{8} = 43.975$$

Sample variance:

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

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

$$s^2 = \frac{\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}}{n-1} = \frac{16528.043 - \frac{(351.8)^2}{8}}{8-1}$$

$$= \frac{1057.998}{7} = 151.143$$

Sample standard deviation:

$$s = \sqrt{151.143} = 12.294$$

The sample standard deviation could also be found using

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

where

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

Dot Diagram:

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

6-6. Sample average:

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

Sample variance:

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

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

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

$$= 16465.61 \quad (watts/m^2)^2$$

Sample standard deviation:

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

The sample standard deviation could also be found using

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

where

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

6-7. $\mu = \frac{6905}{1270} = 5.44$; The value 5.44 is the population mean since the actual physical population of all flight times during the operation is available.

6-8 a.) Sample average:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{19.56}{9} = 2.173 \quad mm$$

b.) Sample variance:

$$\sum_{i=1}^{9} x_i = 19.56$$

$$\sum_{i=1}^{9} x_i^2 = 45.953$$

$$s^2 = \frac{\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}}{n-1} = \frac{45.953 - \frac{(19.56)^2}{9}}{9-1} = \frac{3.443}{8}$$

$$= 0.4303 \quad (mm)^2$$

Sample standard deviation:

$$s = \sqrt{0.4303} = 0.6560 \quad mm$$

c.) Dot Diagram

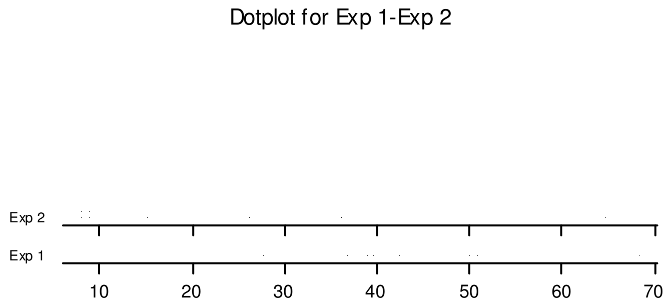
```
.      .             .      .  .          .  .          ..
-------+---------+---------+---------+---------+---------crack length
     1.40      1.75      2.10      2.45      2.80      3.15
```

6-9. Dot Diagram (rounding of the data is used to create the dot diagram)

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

The sample mean is the point at which the data would balance if it were on a scale.

a. Dot Diagram of CRT data in exercise 6-5 (Data were rounded for the plot)



The data are centered a lot lower in the second experiment. The lower CRT resolution reduces the visual accommodation.

6-11. a) $\bar{x} = \dfrac{\sum_{i=1}^{n} x_i}{n} = \dfrac{57.47}{8} = 7.184$

b) $$s^2 = \frac{\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}}{n-1} = \frac{412.853 - \frac{(57.47)^2}{8}}{8-1} = \frac{0.00299}{7} = 0.000427$$

$s = \sqrt{0.000427} = 0.02066$

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.

6-12 sample mean $\bar{x} = \dfrac{\sum_{i=1}^{n} x_i}{n} = \dfrac{748.0}{9} = 83.11$ drag counts

sample variance $$s^2 = \frac{\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}}{n-1} = \frac{62572 - \frac{(748.0)^2}{9}}{9-1}$$

$$= \frac{404.89}{8} = 50.61 \text{ drag counts}^2$$

sample standard deviation $s = \sqrt{50.61} = 7.11$ drag counts

Dot Diagram

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

6-13. a) $\bar{x} = 65.86\ °F$

$s = 12.16\ °F$

b) Dot Diagram

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

c) Removing the smallest observation (31), the sample mean and standard deviation become

$\bar{x} = 66.86\ °F$

$s = 10.74\ °F$

Section 6-3

6-14 Stem-and-leaf display of octane rating N = 83

Leaf Unit = 0.10 83|4 represents 83.4

```
   1   83|4
   3   84|33
   4   85|3
   7   86|777
  13   87|456789
  24   88|23334556679
  34   89|0233678899
 (13)  90|0111344456789
  36   91|00011122256688
  22   92|22236777
  14   93|023347
   8   94|2247
   4   95|
   4   96|15
   2   97|
   2   98|8
   1   99|
   1  100|3
```

6-15 a.) Stem-and-leaf display for cycles to failure: 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.

6-16 Stem-and-leaf display of percentage of cotton N = 64
Leaf Unit = 0.10 32|1 represents 32.1%

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

6-17. Stem-and-leaf display for Problem 2-4.yield: unit = 1 1|2 represents 12

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

6-18 Descriptive Statistics

```
Variable          N  Median    Q1        Q3
Octane Rating   83   90.400 88.600     92.200
```

6-19. Descriptive Statistics

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

6-20 median: $\tilde{x} = 34.700$ %
mode: 34.7 %
sample average: $\bar{x} = 34.798$ %

6-21. Descriptive Statistics

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

6-22 a.) sample mean: $\overline{x} = 65.811$ inches standard deviation $s = 2.106$ inches

b.) Stem-and-leaf display of female engineering student heights N = 37
Leaf Unit = 0.10 61|0 represents 61.0 inches

```
  1    61|0
  3    62|00
  5    63|00
  9    64|0000
 17    65|00000000
 (4)   66|0000
 16    67|00000000
  8    68|00000
  3    69|00
  1    70|0
```

c.) median: $\tilde{x} = 66.000$ inches

6-23 Stem-and-leaf display for Problem 6-23. Strength: unit = 1.0 1|2 represents 12

```
  1   532|9
  1   533|
  2   534|2
  4   535|47
  6   536|6
 11   537|5678
 22   538|12345778888
 28   539|016999
 39   540|11166677889
 47   541|123666688
(12)  542|0011222357899
 41   543|011112556
 32   544|00012455678
 22   545|2334457899
 12   546|23569
  8   547|357
  5   548|11257
```

6-24 Stem-and-leaf of concentration N = 60 Leaf Unit = 1.0 2|2 represents 29

Note: Minitab has dropped the value to the right of the decimal to make this display.

```
   1    2|9
   2    3|1
   3    3|9
   8    4|22223
  12    4|5689
  20    5|01223444
 (13)   5|5666777899999
  27    6|11244
  22    6|556677789
  13    7|022333
   7    7|6777
   3    8|01
   1    8|9
```

The data have a symmetrical bell-shaped distribution, and therefore may be normally distributed.

Sample Mean $\bar{x} = \dfrac{\sum_{i=1}^{n} x_i}{n} = \dfrac{3592.0}{60} = 59.87$

Sample Standard Deviation

$$\sum_{i=1}^{60} x_i = 3592.0 \quad \text{and} \quad \sum_{i=1}^{60} x_i^2 = 224257$$

$$s^2 = \frac{\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}}{n-1} = \frac{224257 - \frac{(3592.0)^2}{60}}{60-1} = \frac{9215.93}{59}$$

$$= 156.20$$

and

$$s = \sqrt{156.20} = 12.50$$

Sample Median $\tilde{x} = 59.45$

```
Variable           N     Median
concentration     60      59.45
```

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.

1	22	6
5	23	2334
8	23	677
16	24	00112444
20	24	5578
33	25	0111122334444
46	25	5555556677899
(15)	26	000011123334444
39	26	56677888
31	27	0000112222233333444
12	27	66788999
4	28	003
1	28	5

Sample Mean $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{\sum_{i=1}^{100} x_i}{100} = \frac{26030.2}{100} = 260.3 \quad yards$

Sample Standard Deviation

$$\sum_{i=1}^{100} x_i = 26030.2 \qquad \text{and} \qquad \sum_{i=1}^{100} x_i^2 = 6793512$$

$$s^2 = \frac{\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}}{n-1} = \frac{6793512 - \frac{((26030.2)^2}{100}}{100-1} = \frac{17798.42}{99}$$

$$= 179.782 \quad yards^2$$

and

$$s = \sqrt{179.782} = 13.41 \quad yards$$

Sample Median

```
Variable              N     Median
yards               100     260.85
```

6-26 Stem-and-leaf of speed (in megahertz) N = 120
Leaf Unit = 1.0 63|4 represents 634 megahertz

```
    2   63|47
    7   64|24899
   16   65|223566899
   35   66|0000001233455788899
   48   67|0022455567899
  (17)  68|00001111233333458
   55   69|0000112345555677889
   36   70|011223444556
   24   71|0057889
   17   72|000012234447
    5   73|59
    3   74|68
    1   75|
    1   76|3
```

35/120= 29% exceed 700 megahertz.

Sample Mean $\bar{x} = \dfrac{\sum_{i=1}^{120} x_i}{120} = \dfrac{82413}{120} = 686.78 \quad mhz$

Sample Standard Deviation

$$\sum_{i=1}^{120} x_i = 82413 \qquad \text{and} \qquad \sum_{i=1}^{120} x_i^2 = 56677591$$

$$s^2 = \frac{\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}}{n-1} = \frac{56677591 - \frac{(82413)^2}{120}}{120-1} = \frac{78402.925}{119}$$

$$= 658.85 \quad mhz^2$$

and

$$s = \sqrt{658.85} = 25.67 \quad mhz$$

Sample Median $\tilde{x} = 683.0\ mhz$

```
Variable             N     Median
speed              120     683.00
```

6-27 a.)Stem-and-leaf display of Problem 6-27. Rating: unit = 0.10 1|2 represents 1.2

```
  1    83|0
  2    84|0
  5    85|000
  7    86|00
  9    87|00
 12    88|000
 18    89|000000
(7)    90|0000000
 15    91|0000000
  8    92|0000
  4    93|0
  3    94|0
  2    95|00
```

b.) Sample Mean

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{\sum_{i=1}^{40} x_i}{40} = \frac{3578}{40} = 89.45$$

Sample Standard Deviation

$$\sum_{i=1}^{40} x_i = 3578 \qquad \text{and} \qquad \sum_{i=1}^{40} x_i^2 = 320366$$

$$s^2 = \frac{\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}}{n-1} = \frac{320366 - \frac{(3578)^2}{40}}{40-1} = \frac{313.9}{39}$$

$$= 8.05$$

and

$$s = \sqrt{8.05} = 2.8$$

Sample Median

```
Variable          N       Median
rating           40       90.000
```

c.) 22/40 or 55% of the taste testers considered this particular Pinot Noir truly exceptional.

6-28 a.) Stem-and-leaf diagram of $NbOCl_3$ N = 27
Leaf Unit = 100 0|4 represents 40 gram-mole/liter x 10^{-3}

```
  6     0|444444
  7     0|5
 (9)    1|001122233
 11     1|5679
  7     2|
  7     2|5677
  3     3|124
```

b.) sample mean $\bar{x} = \frac{\sum_{i=1}^{27} x_i}{27} = \frac{41553}{27} = 1539$ gram − mole/liter x10^{-3}

Sample Standard Deviation

$$\sum_{i=1}^{27} x_i = 41553 \quad \text{and} \quad \sum_{i=1}^{27} x_i^2 = 87792869$$

$$s^2 = \frac{\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}}{n-1} = \frac{87792869 - \frac{(41553)^2}{27}}{27-1} = \frac{23842802}{26} = 917030.85$$

and $s = \sqrt{917030.85} = 957.62$ gram - mole/liter x 10^{-3}

Sample Median $\tilde{x} = 1256$ gram − mole/liter x10^{-3}

```
Variable              N        Median
NbOCl3               40          1256
```

6-29 a.)Stem-and-leaf display for Problem 6-29. Height: unit = 0.10 1|2 represents 1.2

```
   Female Students|| Male Students
         0|61    1  |
        00|62    3  |
        00|63    5  |
      0000|64    9  |
  00000000|65   17  |   2    65|00
      0000|66   (4) |   3    66|0
  00000000|67   16  |   7    67|0000
     00000|68    8  |  17    68|0000000000
        00|69    3  | (15)   69|000000000000000
         0|70    1  |  18    70|0000000
                    |  11     71|00000
                    |   6     72|00
                    |   4    73|00
                    |   2   74|0
                    |   1    75|0
```

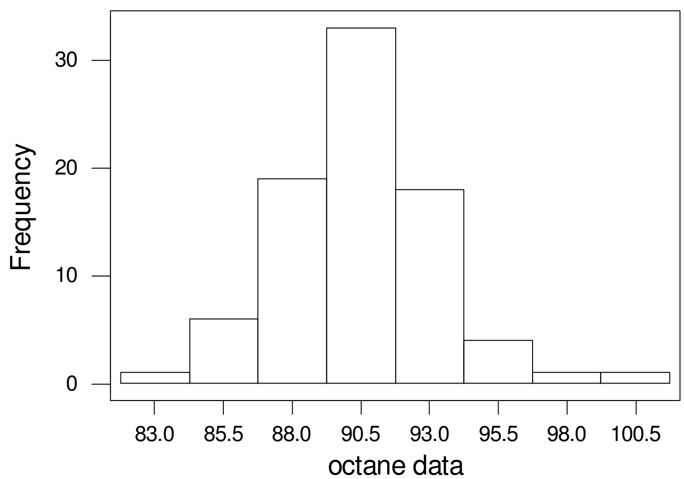
b.) The male engineering students are taller than the female engineering students. Also there is a slightly wider range in the heights of the male students.

Section 6-4

6-30 Frequency Tabulation for Exercise 6-14.Octane Data

Class	Lower Limit	Upper Limit	Midpoint	Frequency	Relative Frequency	Cumulative Frequency	Cum. Rel. Frequency
at or below		81.75		0	.0000	0	.0000
1	81.75	84.25	83.0	1	.0120	1	.0120
2	84.25	86.75	85.5	6	.0723	7	.0843
3	86.75	89.25	88.0	19	.2289	26	.3133
4	89.25	91.75	90.5	33	.3976	59	.7108
5	91.75	94.25	93.0	18	.2169	77	.9277
6	94.25	96.75	95.5	4	.0482	81	.9759
7	96.75	99.25	98.0	1	.0120	82	.9880
8	99.25	101.75	100.5	1	.0120	83	1.0000
above	101.75			0	.0000	83	1.0000

Mean = 90.534 Standard Deviation = 2.888 Median = 90.400

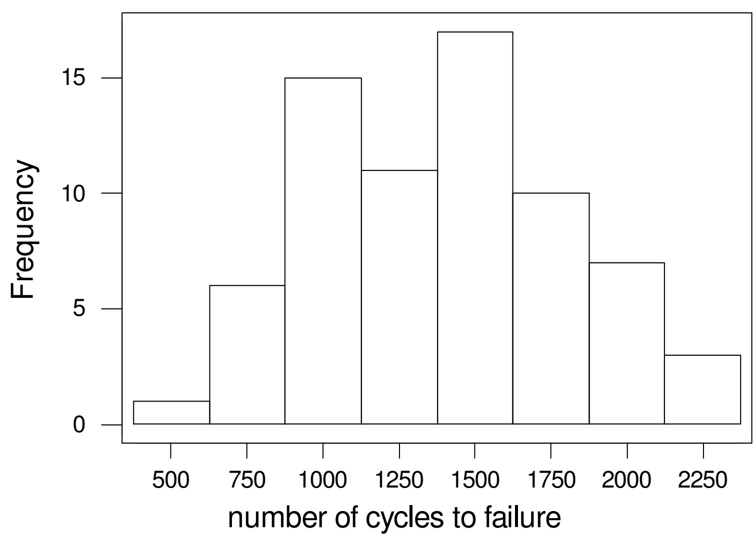


6-31.

Frequency Tabulation for Exercise 6-15.Cycles

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

Mean = 1403.66   Standard Deviation = 402.385   Median = 1436.5

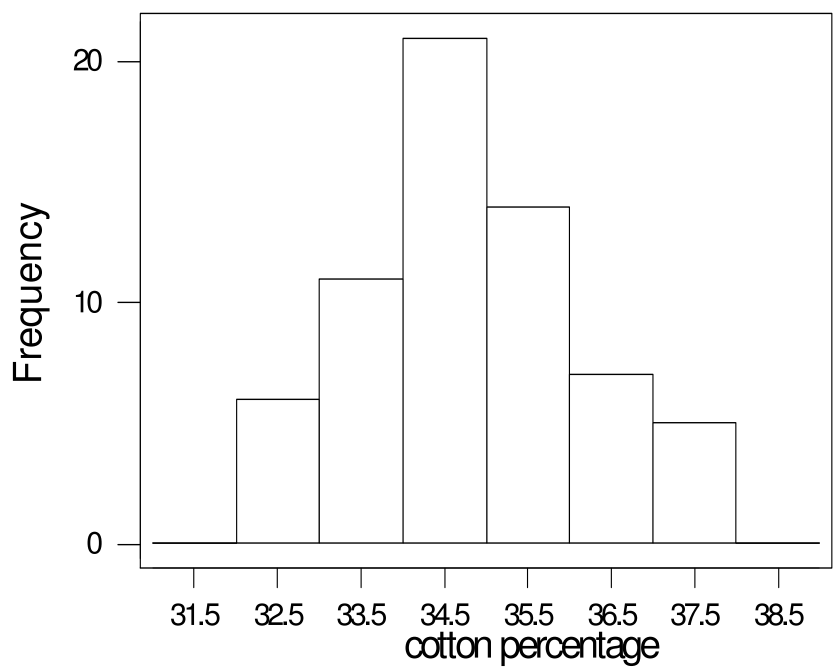


6-32

Frequency Tabulation for Exercise 6-16.Cotton content

Class	Lower Limit	Upper Limit	Midpoint	Frequency	Relative Frequency	Cumulative Frequency	Cum. Rel. Frequency
at or below		31.0		0	.0000	0	.0000
1	31.0	32.0	31.5	0	.0000	0	.0000
2	32.0	33.0	32.5	6	.0938	6	.0938
3	33.0	34.0	33.5	11	.1719	17	.2656
4	34.0	35.0	34.5	21	.3281	38	.5938
5	35.0	36.0	35.5	14	.2188	52	.8125
6	36.0	37.0	36.5	7	.1094	59	.9219
7	37.0	38.0	37.5	5	.0781	64	1.0000
8	38.0	39.0	38.5	0	.0000	64	1.0000
above	39.0			0	.0000	64	1.0000

Mean = 34.798   Standard Deviation = 1.364   Median = 34.700

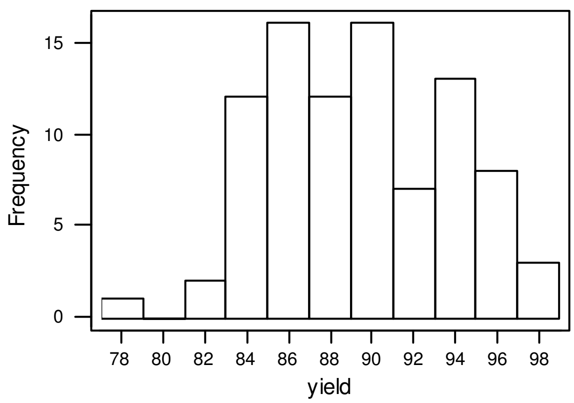


6-33.

Frequency Tabulation for Exercise 6-17.Yield

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

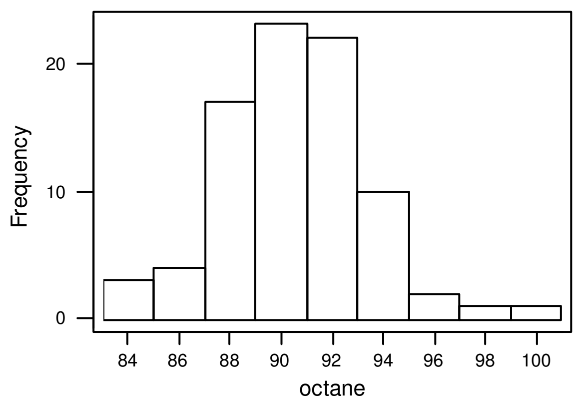
Mean = 89.3756   Standard Deviation = 4.31591   Median = 89.25



6-34 Frequency Tabulation for Exercise 6-14.Octane Data

Class	Lower Limit	Upper Limit	Midpoint	Frequency	Relative Frequency	Cumulative Frequency	Cum. Rel. Frequency
at or below		83.000		0	.0000	0	.0000
1	83.000	84.125	83.5625	1	.0120	1	.0120
2	84.125	85.250	84.6875	2	.0241	3	.0361
3	85.250	86.375	85.8125	1	.0120	4	.0482
4	86.375	87.500	86.9375	5	.0602	9	.1084
5	87.500	88.625	88.0625	13	.1566	22	.2651
6	88.625	89.750	89.1875	8	.0964	30	.3614
7	89.750	90.875	90.3125	16	.1928	46	.5542
8	90.875	92.000	91.4375	15	.1807	61	.7349
9	92.000	93.125	92.5625	9	.1084	70	.8434
10	93.125	94.250	93.6875	7	.0843	77	.9277
11	94.250	95.375	94.8125	2	.0241	79	.9518
12	95.375	96.500	95.9375	2	.0241	81	.9759
13	96.500	97.625	97.0625	0	.0000	81	.9759
14	97.625	98.750	98.1875	0	.0000	81	.9759
15	98.750	99.875	99.3125	1	.0120	82	.9880
16	99.875	101.000	100.4375	1	.0120	83	1.0000
above	101.000			0	.0000	83	1.0000

Mean = 90.534 Standard Deviation = 2.888 Median = 90.400



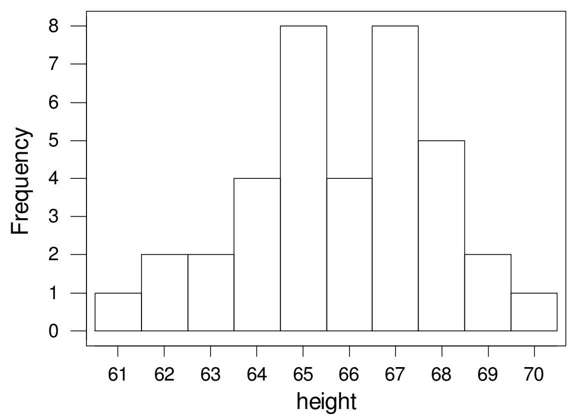
The histograms have the same shape. Not much information is gained by doubling the number of bins.

6-35

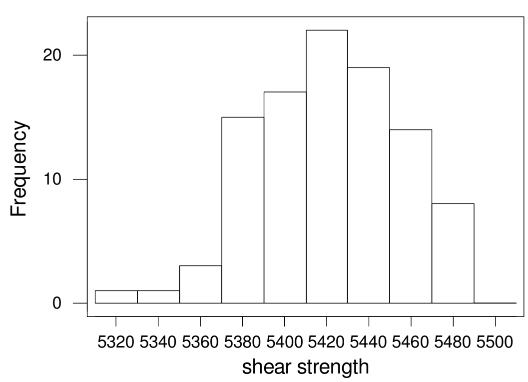
Frequency Tabulation for Problem 6-22. Height Data

Class	Lower Limit	Upper Limit	Midpoint	Frequency	Relative Frequency	Cumulative Frequency	Cum. Rel. Frequency
at or below		60.500		0	.0000	0	.0000
1	60.500	61.500	61.000	1	.0270	1	.0270
2	61.500	62.500	62.000	2	.0541	3	.0811
3	62.500	63.500	63.000	2	.0541	5	.1351
4	63.500	64.500	64.000	4	.1081	9	.2432
5	64.500	65.500	65.000	8	.2162	17	.4595
6	65.500	66.500	66.000	4	.1081	21	.5676
7	66.500	67.500	67.000	8	.2162	29	.7838
8	67.500	68.500	68.000	5	.1351	34	.9189
9	68.500	69.500	69.000	2	.0541	36	.9730
10	69.500	70.500	70.000	1	.0270	37	1.0000
above	70.500			0	.0000	37	1.0000

Mean = 65.811   Standard Deviation = 2.106   Median = 66.0



6-36 The histogram for the spot weld shear strength data shows that the data appear to be normally distributed (the same shape that appears in the stem-leaf-diagram).

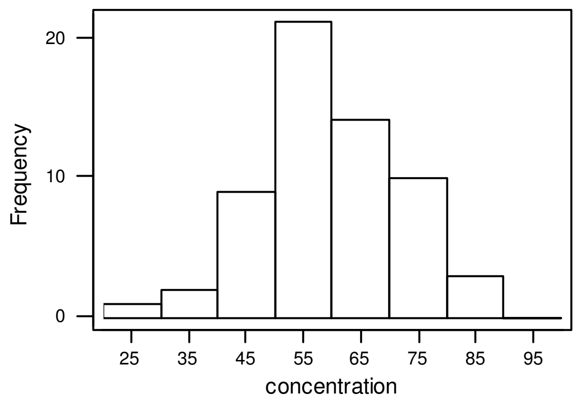


6-37

Frequency Tabulation for exercise 6-24. Concentration data

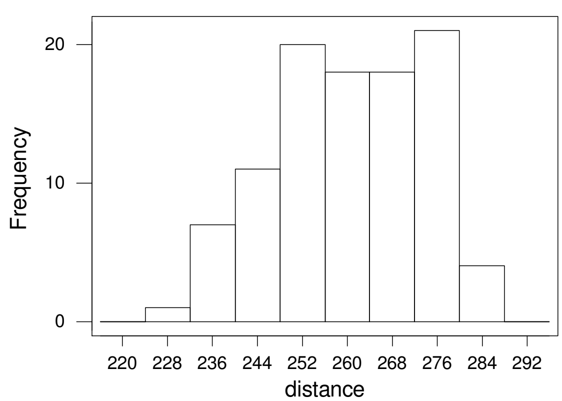
Class	Lower Limit	Upper Limit	Midpoint	Frequency	Relative Frequency	Cumulative Frequency	Cum. Rel. Frequency
at or below		29.000		0	.0000	0	.0000
1	29.0000	37.000	33.000	2	.0333	2	.0333
2	37.0000	45.000	41.000	6	.1000	8	.1333
3	45.0000	53.000	49.000	8	.1333	16	.2667
4	53.0000	61.000	57.000	17	.2833	33	.5500
5	61.0000	69.000	65.000	13	.2167	46	.7667
6	69.0000	77.000	73.000	8	.1333	54	.9000
7	77.0000	85.000	81.000	5	.0833	59	.9833
8	85.0000	93.000	89.000	1	.0167	60	1.0000
above	93.0000			0	.0800	60	1.0000

Mean = 59.87   Standard Deviation = 12.50   Median = 59.45

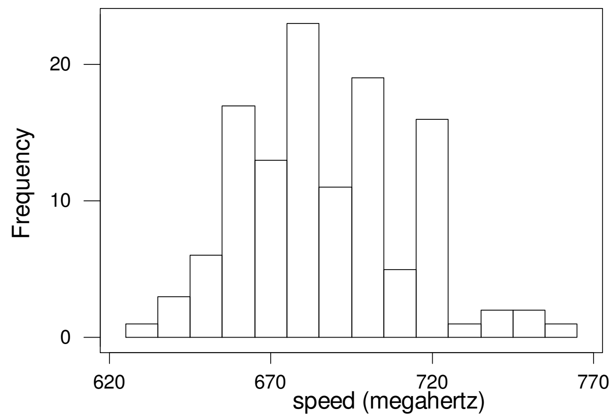


Yes, the histogram shows the same shape as the stem-and-leaf display.

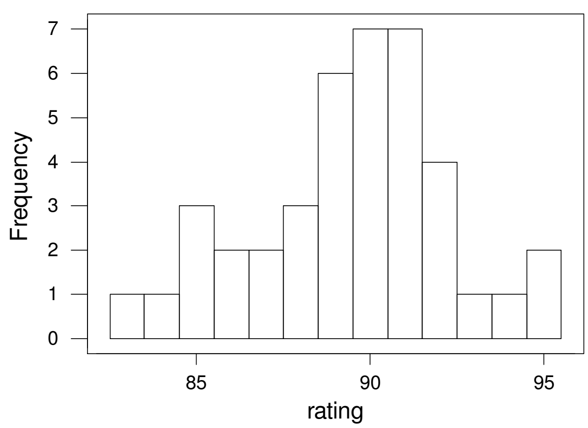
6-38 Yes, the histogram of the distance data shows the same shape as the stem-and-leaf display in exercise 6-25.



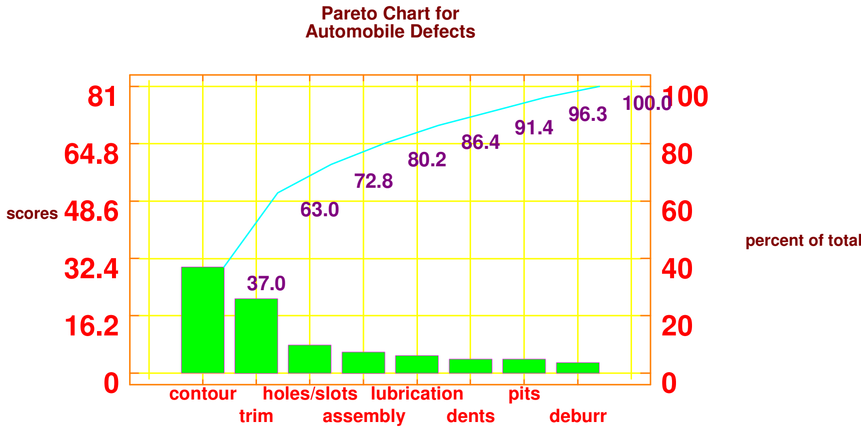
6-39 Histogram for the speed data in exercise 6-26. Yes, the histogram of the speed data shows the same shape as the stem-and-leaf display in exercise 6-26



6-40 Yes, the histogram of the wine rating data shows the same shape as the stem-and-leaf display in exercise 6-27.



6-41



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

<u>Section 6-5</u>

6-42 Descriptive Statistics of O-ring joint temperature data

Variable	N	Mean	Median	TrMean	StDev	SE Mean
Temp	36	65.86	67.50	66.66	12.16	2.03

Variable	Minimum	Maximum	Q1	Q3
Temp	31.00	84.00	58.50	75.00

a.)

Lower Quartile: Q_1=58.50

Upper Quartile: Q_3=75.00

b.) Median = 67.50

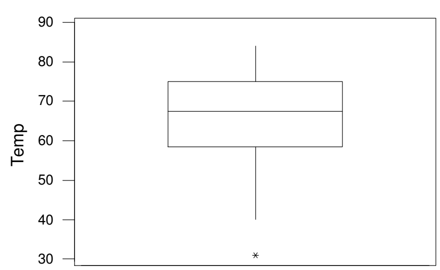
c.) Data with lowest point removed

Variable	N	Mean	Median	TrMean	StDev	SE Mean
Temp	35	66.86	68.00	67.35	10.74	1.82

Variable	Minimum	Maximum	Q1	Q3
Temp	40.00	84.00	60.00	75.00

The mean and median have increased and the standard deviation and difference between the upper and lower quartile has decreased.

d.)Box Plot - The box plot indicates that there is an outlier in the data.

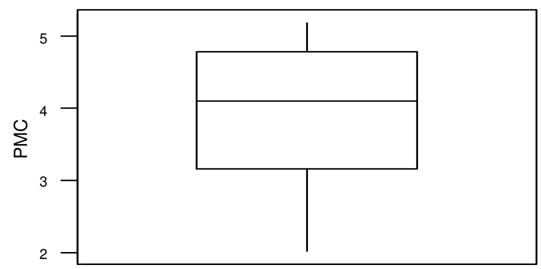


6-43. Descriptive Statistics

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)

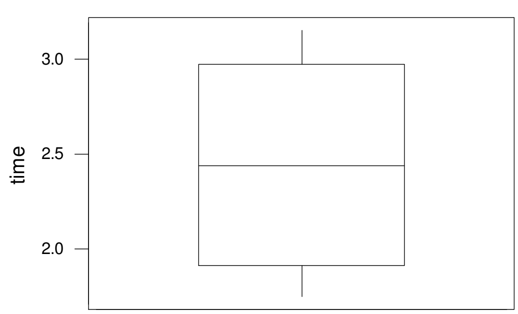


6-44 Descriptive Statistics

Variable	N	Mean	Median	TrMean	StDev	SE Mean
time	8	2.415	2.440	2.415	0.534	0.189

Variable	Minimum	Maximum	Q1	Q3
time	1.750	3.150	1.912	2.973

a.)Sample Mean: 2.415
Sample Standard Deviation: 0.543
b.) Box Plot – There are no outliers in the data.

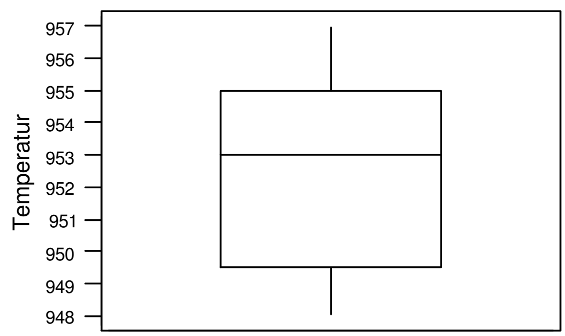


6-45. Descriptive Statistics

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

Variable	Min	Max	Q1	Q3
Temperat	948.00	957.00	949.50	955.00

a) Sample Mean: 952.44
Sample Variance: 9.55
Sample Standard Deviation: 3.09
b) Median: 953; Any increase in the largest temperature measurement will not affect the median.
c)

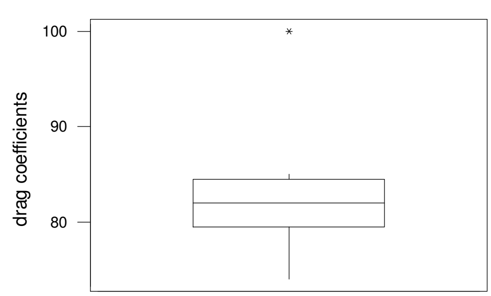


6-46 Descriptive statistics

Variable	N	Mean	Median	TrMean	StDev	SE Mean
drag coefficients	9	83.11	82.00	83.11	7.11	2.37

Variable	Minimum	Maximum	Q1	Q3
drag coefficients	74.00	100.00	79.50	84.50

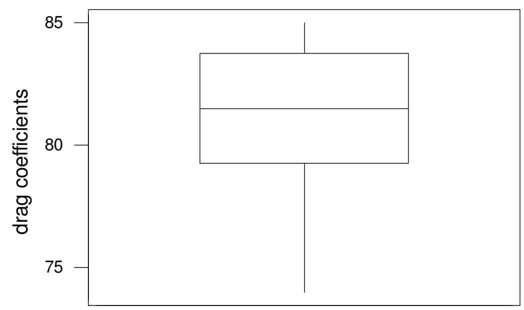
a.) Upper quartile: Q_1 =79.50
Lower Quartile: Q_3=84.50
b.)



c.)

Variable	N	Mean	Median	TrMean	StDev	SE Mean
drag coefficients	8	81.00	81.50	81.00	3.46	1.22

Variable	Minimum	Maximum	Q1	Q3
drag coefficients	74.00	85.00	79.25	83.75



Removing the largest observation (100) lowers the mean and median. Removing this "outlier also greatly reduces the variability as seen by the smaller standard deviation and the smaller difference between the upper and lower quartiles.

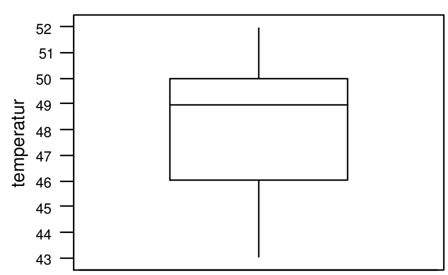
6-47.

Descriptive Statistics

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

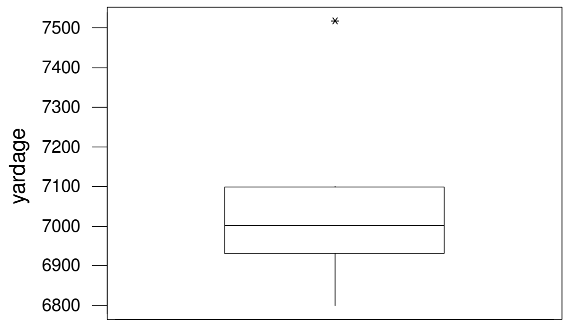
Variable	Min	Max	Q1	Q3
temperat	43.000	52.000	46.000	50.000

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

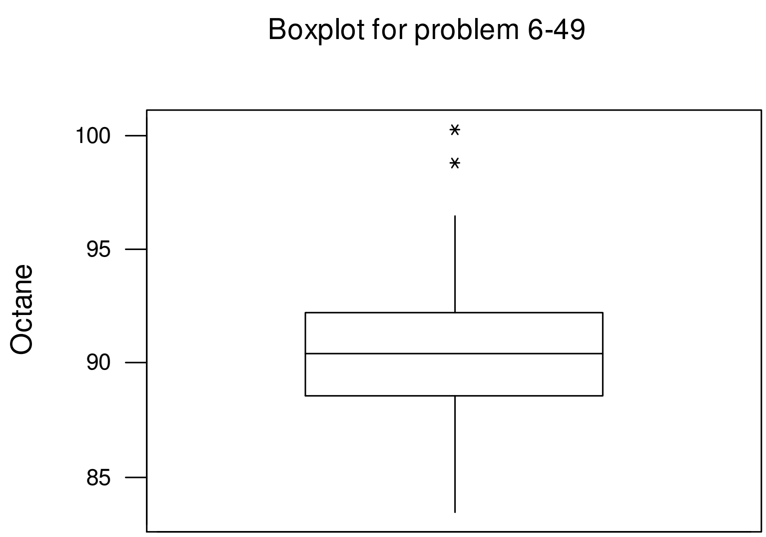


The data appear to be slightly skewed.

4-48 The golf course yardage data appear to be skewed. Also, there is an outlying data point above 7500 yards.

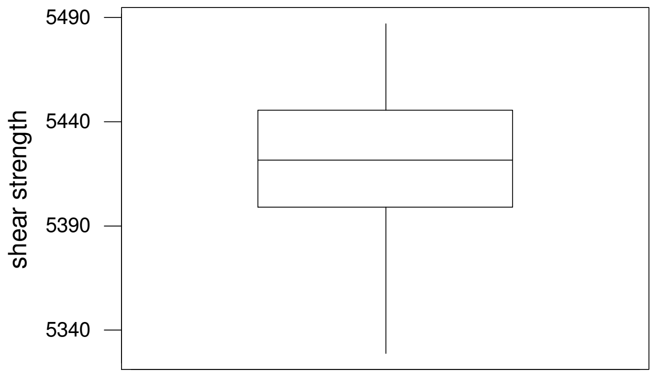


6-49

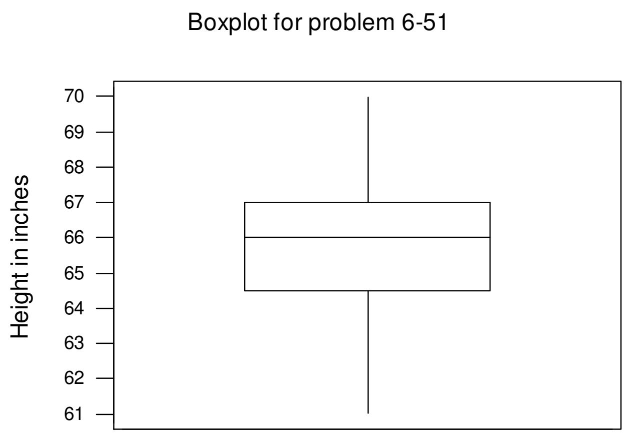


This plot conveys the same basic information as the stem and leaf plot but in a different format. The outliers that were separated from the main portion of the stem and leaf plot are shown here separated from the whiskers.

6-50 The box plot shows that the data are symmetrical about the mean. It also shows that there are no outliers in the data. These are the same interpretations seen in the stem-leaf-diagram.

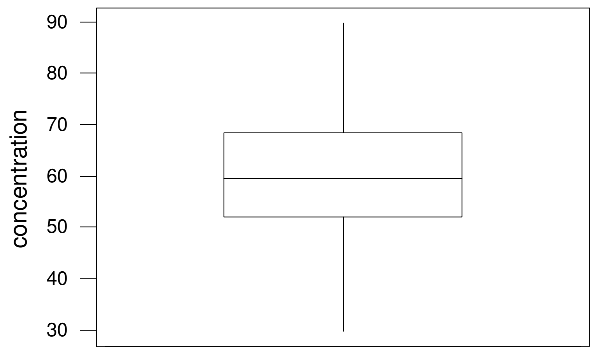


6-51

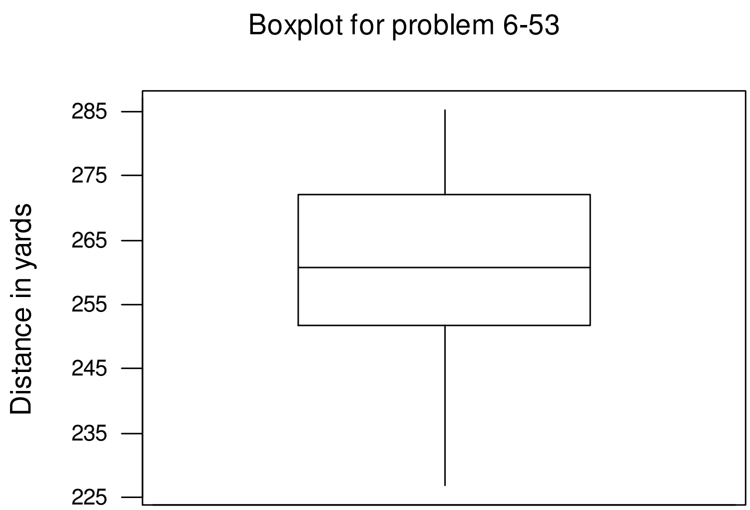


This plot, as the stem and leaf one, indicates that the data fall mostly in one region and that the measurements toward the ends of the range are more rare.

6-52 The box plot and the stem-leaf-diagram show that the data are very symmetrical about the mean. It also shows that there are no outliers in the data.

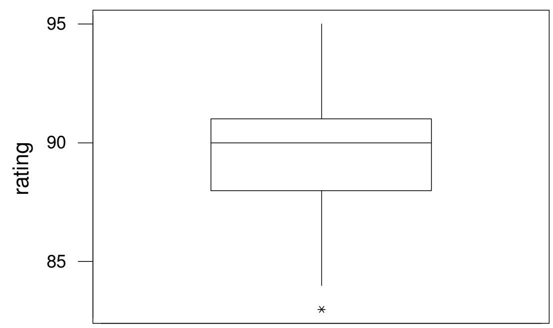


6-53

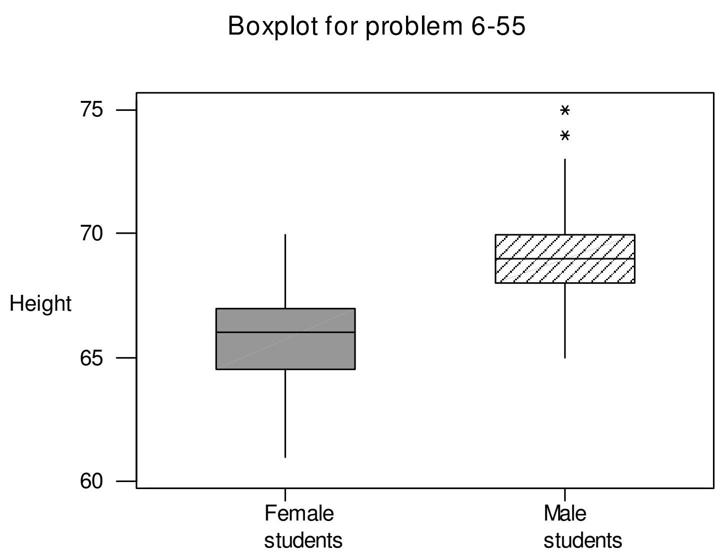


The plot indicates that most balls will fall somewhere in the 250-275 range. In general, the population is grouped more toward the high end of the region. This same type of information could have been obtained from the stem and leaf graph of problem 6-25.

6-54 The box plot shows that the data are not symmetrical about the mean. The data are skewed to the right and have a longer right tail (at the lower values). It also shows that there is an outlier in the data. These are the same interpretations seen in the stem-leaf-diagram.

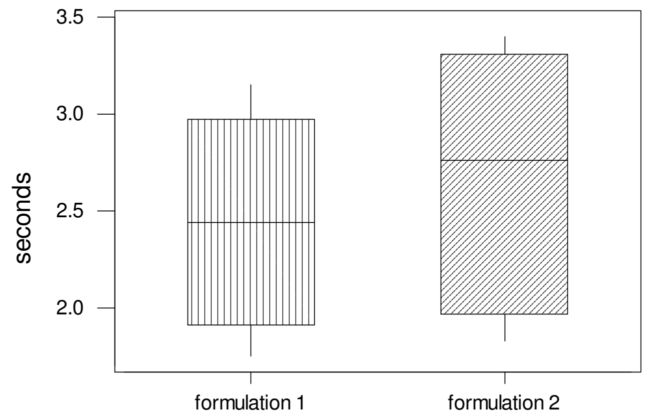


6-55



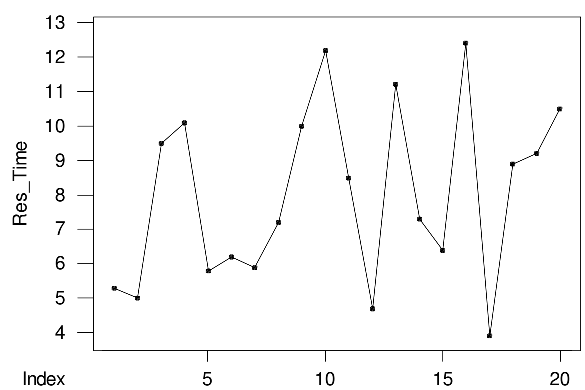
We can see that the two distributions seem to be centered at different values.

6-56 The box plot shows that there is a difference between the two formulations. Formulation 2 has a higher mean cold start ignition time and a larger variability in the values of the start times. The first formulation has a lower mean cold start ignition time and is more consistent. Care should be taken, though since these box plots for formula 1 and formula 2 are made using 8 and 10 data points respectively. More data should be collected on each formulation to get a better determination.



Section 6-6

6-57. Time Series Plot



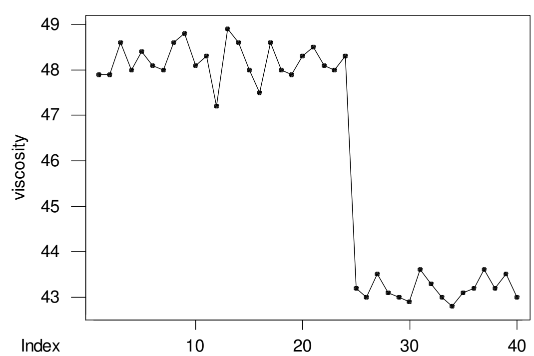
Computer response time appears random. No trends or patterns are obvious.

6-58 a.) Stem-leaf-plot of viscosity N = 40
Leaf Unit = 0.10

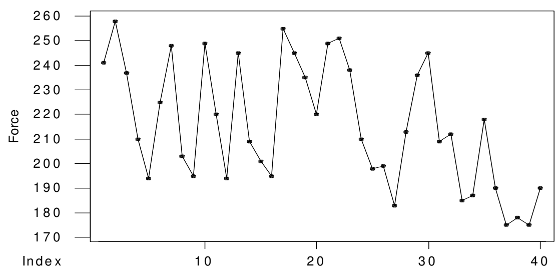
```
  2    42 89
 12    43 0000112223
 16    43 5566
 16    44
 16    44
 16    45
 16    45
 16    46
 16    46
 17    47 2
 (4)   47 5999
 19    48 000001113334
  7    48 5666689
```

The stem-leaf-plot shows that there are two “different” sets of data. One set of data is centered about 43 and the second set is centered about 48. The time series plot shows that the data starts out at the higher level and then drops down to the lower viscosity level at point 24. Each plot gives us a different set of information.

b.) If the specifications on the product viscosity are 48.0±2, then there is a problem with the process performance after data point 24. An investigation needs to take place to find out why the location of the process has dropped from around 48.0 to 43.0. The most recent product is not within specification limits.



6-59. a)



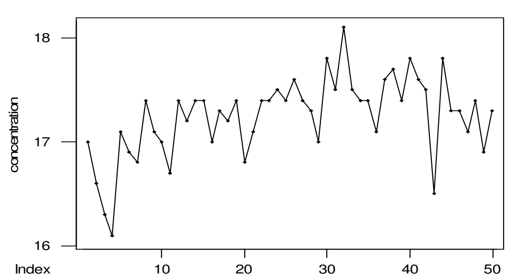
b) Stem-and-leaf display for Problem 2-23.Force: unit = 1   1|2 represents 12

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

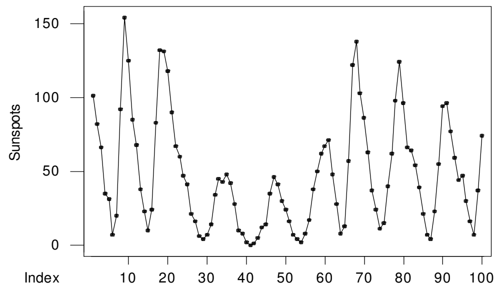
In the time series plot there appears to be a downward trend beginning after time 30. The stem and leaf plot does not reveal this.

6-60

```
1     18 | 1
4     17 | 888
8     17 | 6667
25    17 | 44444444444445555
(7)   17 | 2233333
18    17 | 000011111
9     16 | 8899
5     16 | 567
2     16 | 3
1     16 | 1
```



6-61 a)

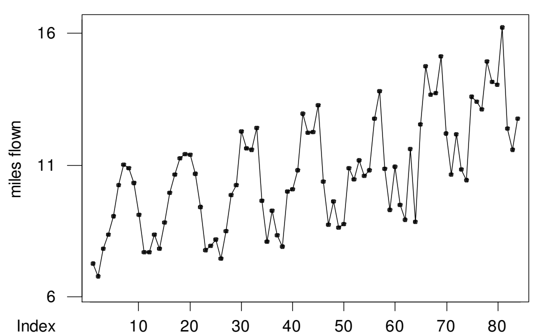


b) Stem-and-leaf display for Problem 2-25.Sunspots: unit = 1 1|2 represents 12

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

      HI|154
```

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

6-62 a.) Time Series Plot



Each year the miles flown peaks during the summer hours. The number of miles flown increased over the years 1964 to 1970.

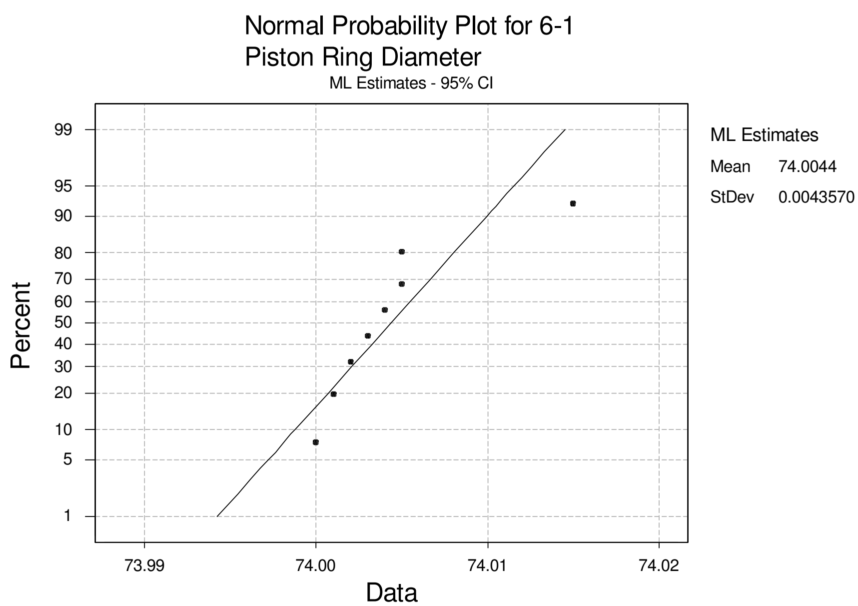
b.) `Stem-and-leaf of miles fl  N  = 84`

```
    Leaf Unit = 0.10

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

When grouped together, the yearly cycles in the data are not seen. The data in the stem-leaf-diagram appear to be nearly normally distributed.

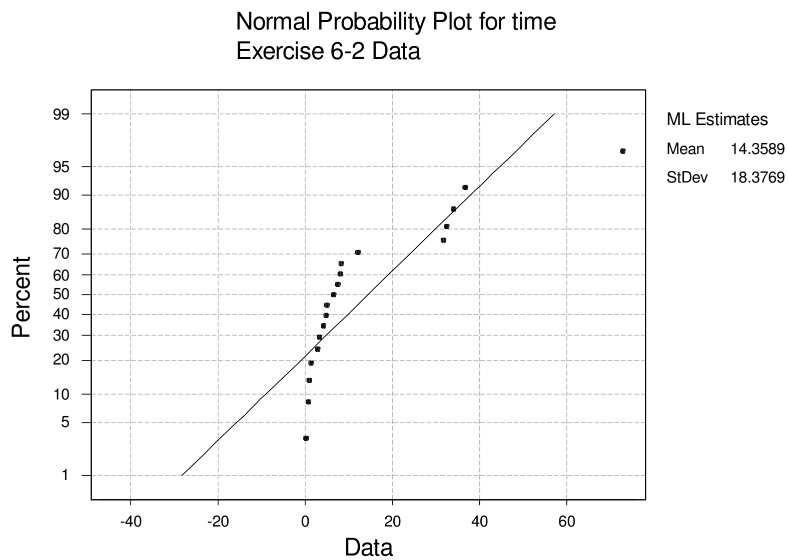
Section 6-7

6-63



The pattern of the data indicates that the sample may not come from a normally distributed population or that the largest observation is an outlier. Note the slight bending downward of the sample data at both ends of the graph.

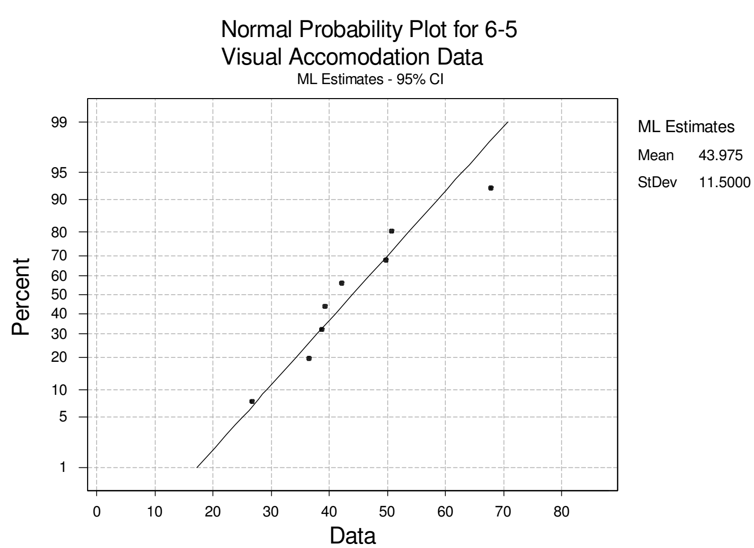
6-64



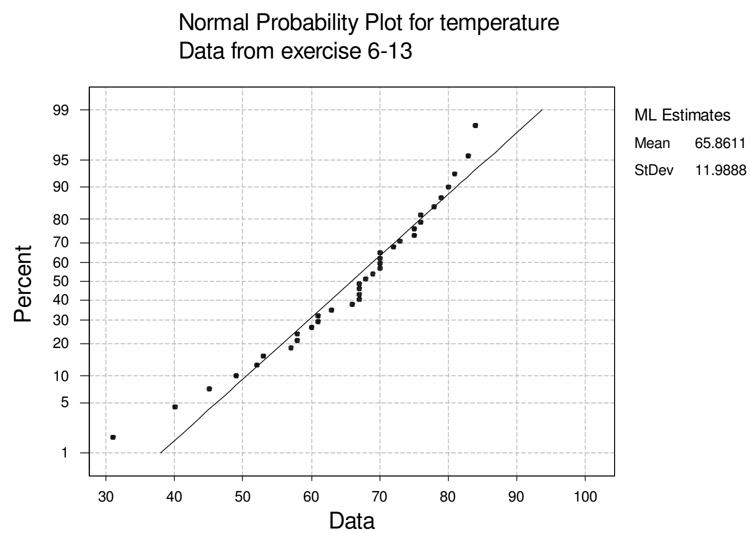
It appears that the data do not come from a normal distribution. Very few of the data points fall on the line.

6-65

There is no evidence to doubt that data are normally distributed

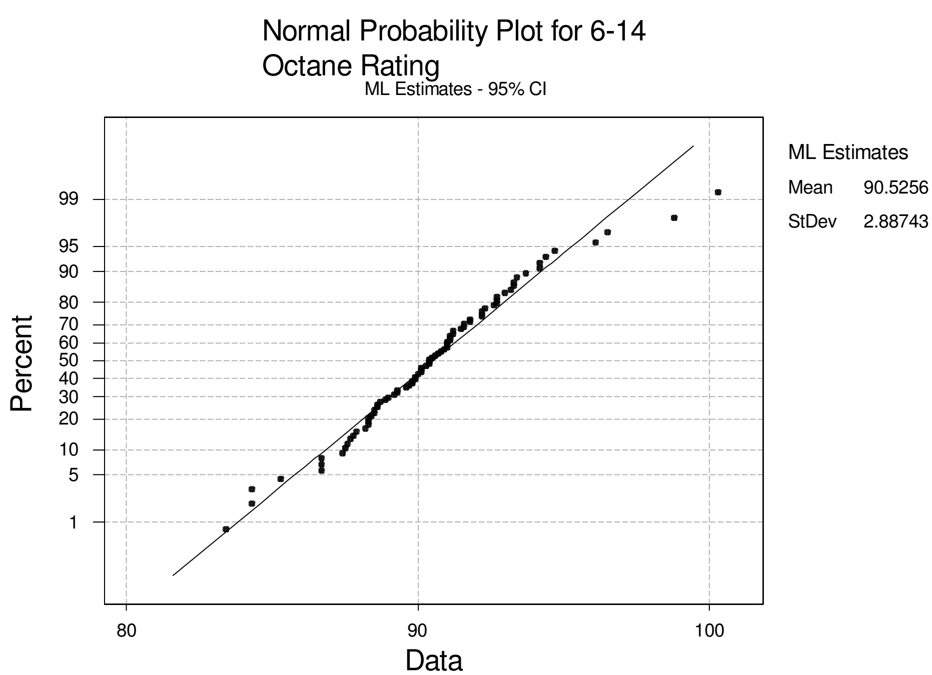


6-66



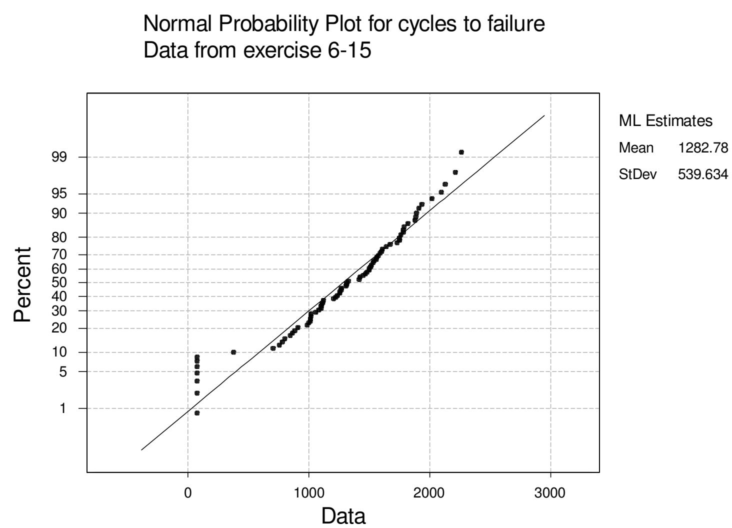
The data appear to be normally distributed. Although, there are some departures from the line at the ends of the distribution.

6-67



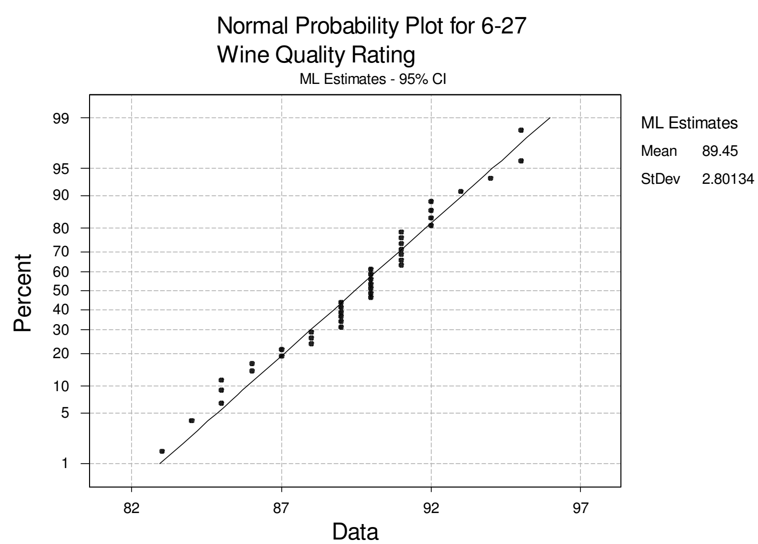
There are a few points outside the confidence limits, indicating that the sample is not perfectly normal. These deviations seem to be fairly small though.

6-68



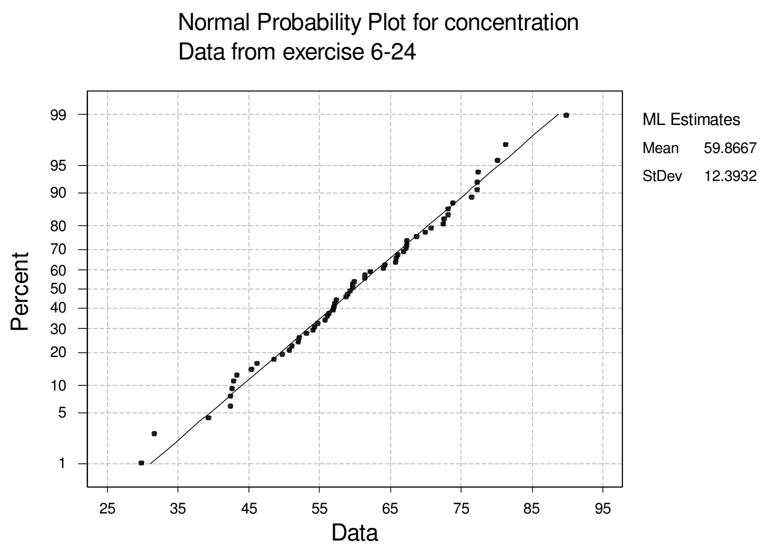
The data appear to be normally distributed. Although, there are some departures from the line at the ends of the distribution.

6-69



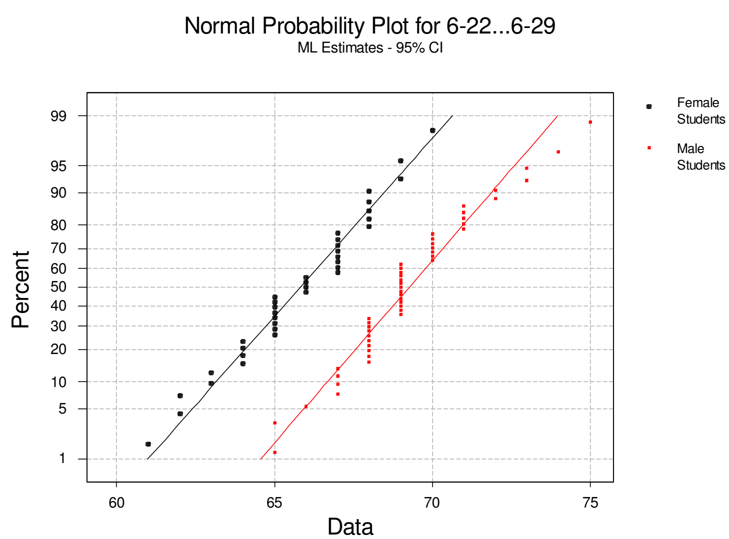
The data seem to be normally distributed. Notice that there are clusters of observations because of the discrete nature of the ratings.

6-70



The data appear to be normally distributed. Nearly all of the data points fall very close to, or on the line.

6-71



Both populations seem to be normally distributed, moreover, the lines seem to be roughly parallel indicating that the populations may have the same variance and differ only in the value of their mean.

6-72 Yes, it is possible to obtain an estimate of the mean from the 50th percentile value of the normal probability plot. The fiftieth percentile point is the point at which the sample mean should equal the population mean and 50% of the data would be above the value and 50% below. An estimate of the standard deviation would be to subtract the 50th percentile from the 64th percentile These values are based on the values from the z-table that could be used to estimate the standard deviation.

Supplemental Exercises

6-73. a) Sample Mean = 65.083
The sample mean value is close enough to the target value to accept the solution as conforming. There is a slight difference due to inherent variability.

b) $s^2 = 1.86869$ $s = 1.367$
A major source of variability would include measurement to measurement error.
A low variance is desirable since it may indicate consistency from measurement to measurement.

6-74 a) $\sum_{i=1}^{6} x_i^2 = 10{,}433$ $\left(\sum_{i=1}^{6} x_i\right)^2 = 62{,}001$ $n = 6$

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

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

b) $\sum_{i=1}^{6} x_i^2 = 353$ $\left(\sum_{i=1}^{6} x_i\right)^2 = 1521$ $n = 6$

$$s^2 = \frac{\sum_{i=1}^{6} x_i^2 - \frac{\left(\sum_{i=1}^{6} x_i\right)^2}{n}}{n-1} = \frac{353 - \frac{1{,}521}{6}}{6-1} = 19.9\Omega^2$$

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

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

c) $\sum_{i=1}^{6} x_i^2 = 1043300$ $\left(\sum_{i=1}^{6} x_i\right)^2 = 6200100$ $n = 6$

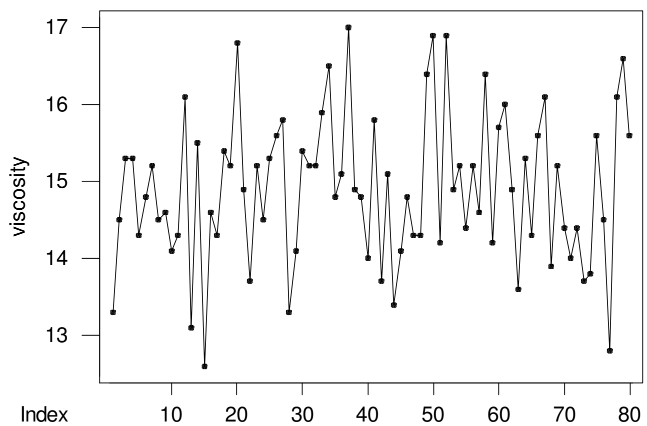
$$s^2 = \frac{\sum_{i=1}^{6} x_i^2 - \frac{\left(\sum_{i=1}^{6} x_i\right)^2}{n}}{n-1} = \frac{1043300 - \frac{6200100}{6}}{6-1} = 1990\Omega^2$$

$$s = \sqrt{1990\Omega^2} = 44.61\Omega$$

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

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.

6-76 a.) It appears that the data may shift up and then down over the 80 points.



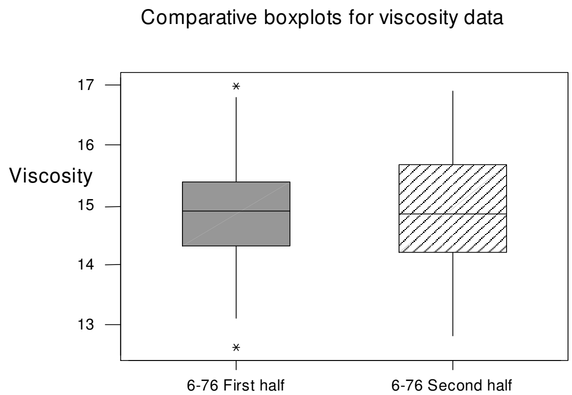
b.)It appears that the mean of the second set of 40 data points may be slightly higher than the first set of 40.

c.) Descriptive Statistics: viscosity 1, viscosity 2

Variable	N	Mean	Median	TrMean	StDev	SE Mean
Viscosity1	40	14.875	14.900	14.875	0.948	0.150
Viscosity2	40	14.923	14.850	14.914	1.023	0.162

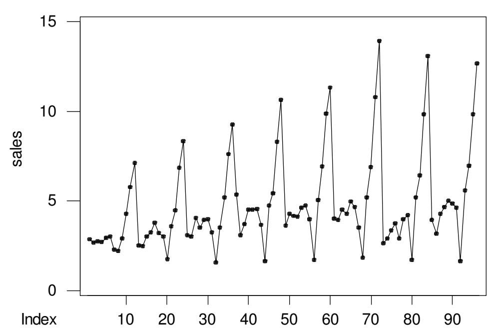
There is a slight difference in the mean levels and the standard deviations.

6-77



Both sets of data appear to have the same mean although the first half seem to be concentrated a little more tightly. Two data points appear as outliers in the first half of the data.

6-78



There appears to be a cyclic variation in the data with the high value of the cycle generally increasing. The high values are during the winter holiday months.

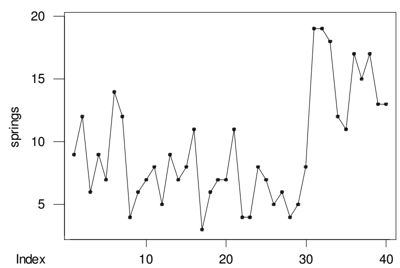
b) We might draw another cycle, with the peak similar to the last year's data (1969) at about 12.7 thousand bottles.

6-79 a) `Stem-and-leaf display for Problem 2-35: unit = 1    1|2  represents 12`

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

b) Sample Average = 9.325
Sample Standard Deviation = 4.4858

c)

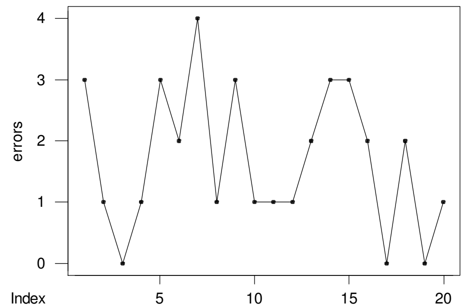


The time series plot indicates there was an increase in the average number of nonconforming springs made during the 40 days. In particular, the increase occurs during the last 10 days.

6-80 a.)
```
Stem-and-leaf of errors    N  = 20
Leaf Unit = 0.10

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

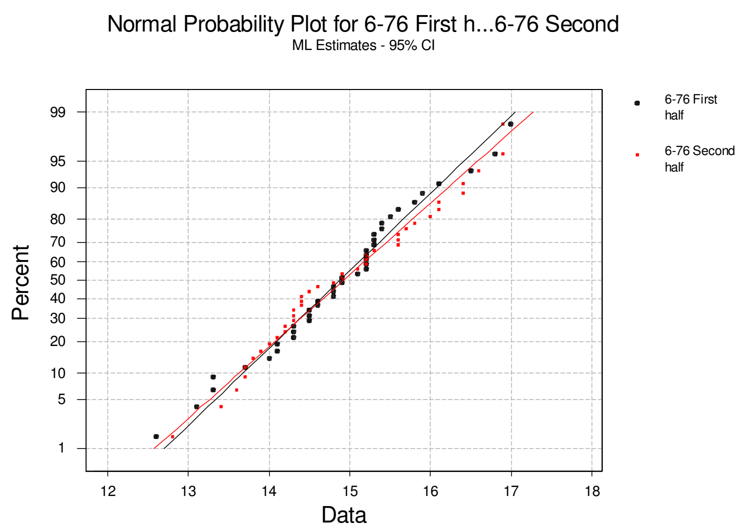
b.) Sample Average = 1.700
Sample Standard Deviation = 1.174

c.)



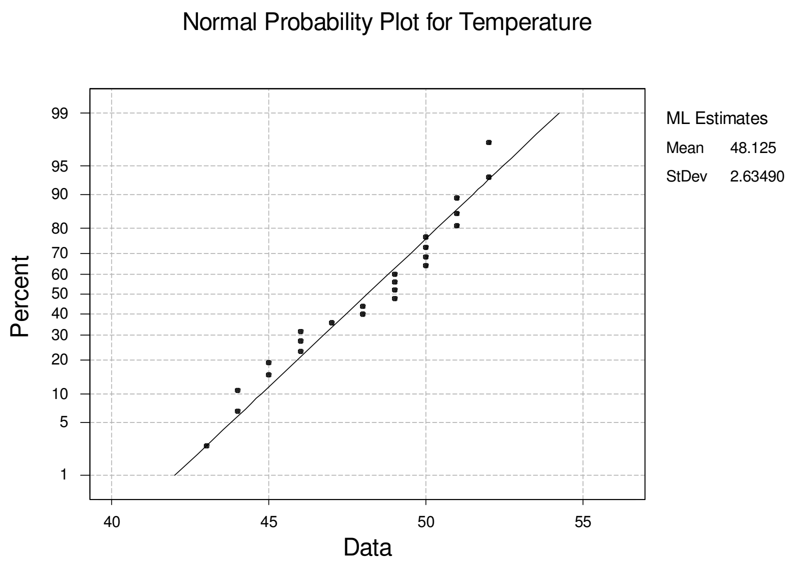
The time series plot indicates a slight decrease in the number of errors for strings 16 - 20.

6-81



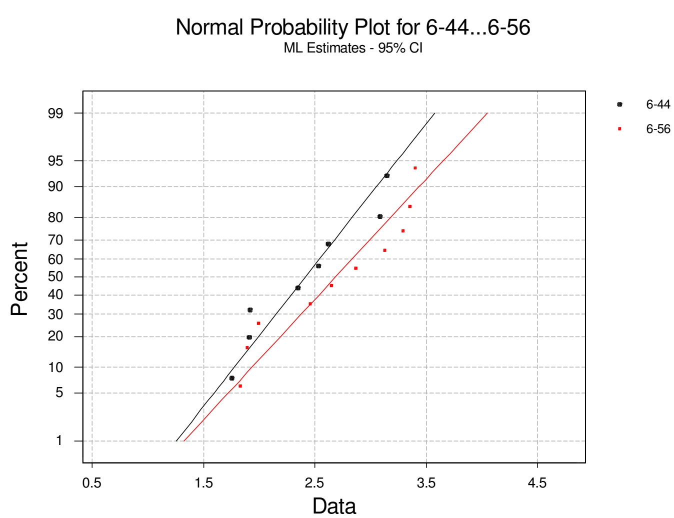
Both sets of data appear to be normally distributed and with roughly the same mean value. The difference in slopes for the two lines indicates that a change in variance might have occurred. This could have been the result of a change in processing conditions, the quality of the raw material or some other factor.

6-82



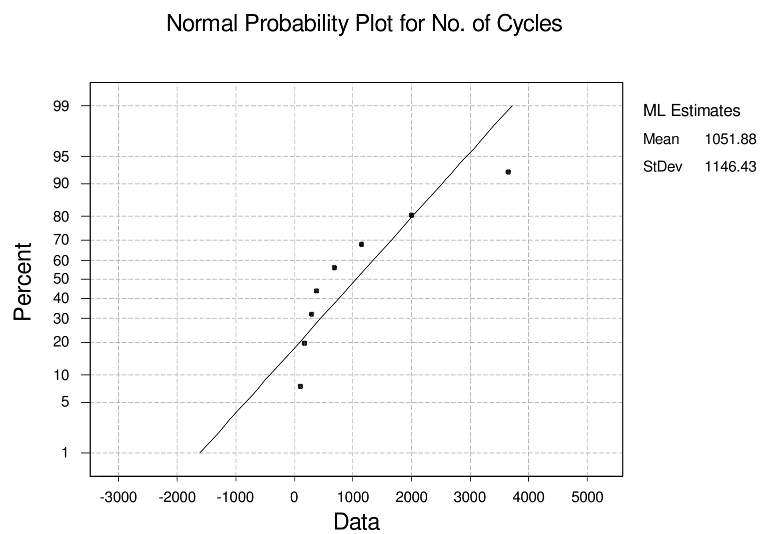
There appears to be no evidence that the data are not normally distributed. There are some repeat points in the data that cause some points to fall off the line.

6-83



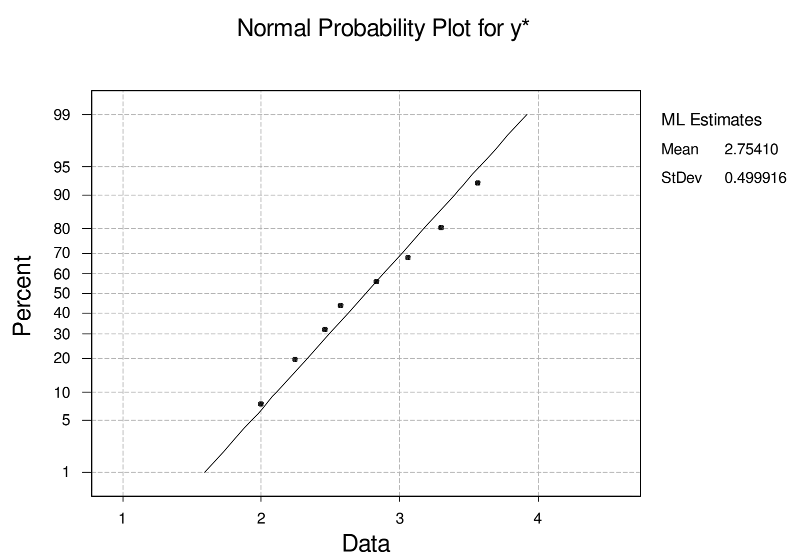
Although we do not have sufficient data points to really see a pattern, there seem to be no significant deviations from normality for either sample. The large difference in slopes indicates that the variances of the populations are very different.

8-84 a.)



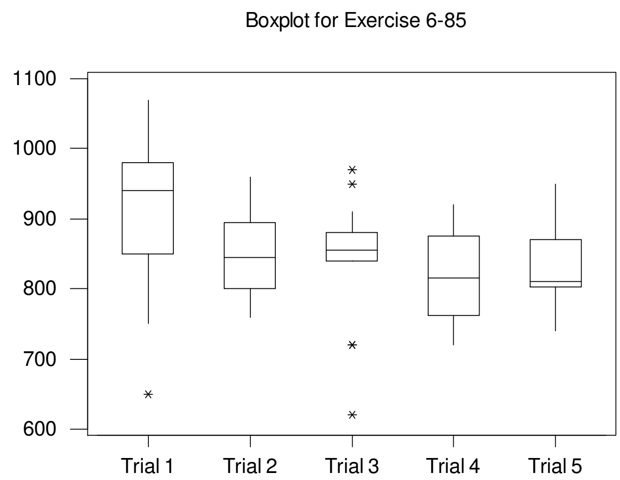
The data do not appear to be normally distributed. There is a curve in the line.

b.)



After the transformation y*=log(y), the normal probability plot shows no evidence that the data are not normally distributed.

6-85

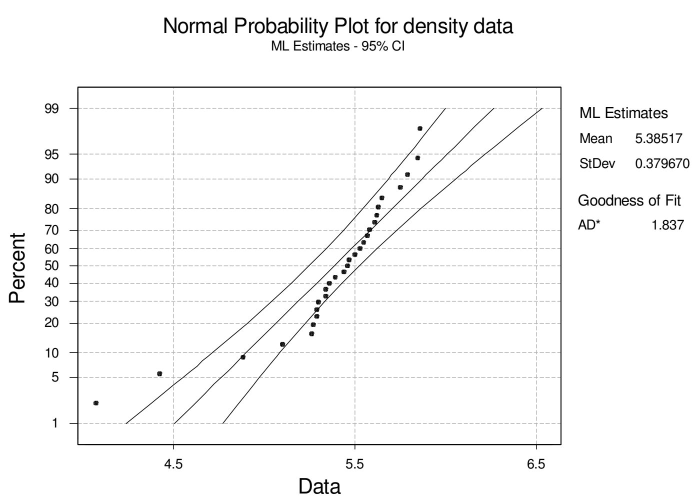


- There is a difference in the variability of the measurements in the trials. Trial 1 has the most variability in the measurements. Trial 3 has a small amount of variability in the main group of measurements, but there are four outliers. Trial 5 appears to have the least variability without any outliers.
- All of the trials except Trial 1 appear to be centered around 850. Trial 1 has a higher mean value
- All five trials appear to have measurements that are greater than the "true" value of 734.5.
- The difference in the measurements in Trial 1 may indicate a "start-up" effect in the data. There could be some bias in the measurements that is centering the data above the "true" value.

6-86 a.) Descriptive Statistics

Variable	N	Mean	Median	TrMean	StDev	SE Mean
density	29	5.4541	5.4600	5.4611	0.4072	0.0756

b.) There does appear to be a low outlier in the data.



c.)Due to the very low data point at 4.07, the mean may be lower than it should be. Therefore, the median would be a better estimate of the density of the earth. The median is not affected by outliers.

Mind Expanding Exercises

6-87 $\sum_{i=1}^{9} x_i^2 = 62572 \qquad \left(\sum_{i=1}^{9} x_i\right)^2 = 559504 \qquad n = 9$

$$s^2 = \frac{\sum_{i=1}^{9} x_i^2 - \frac{\left(\sum_{i=1}^{9} x_i\right)^2}{n}}{n-1} = \frac{62572 - \frac{559504}{9}}{9-1} = 50.61$$

$s = \sqrt{50.61} = 7.11$

Subtract 30 and multiply by 10

$\sum_{i=1}^{9} x_i^2 = 2579200 \qquad \left(\sum_{i=1}^{9} x_i\right)^2 = 22848400 \qquad n = 9$

$$s^2 = \frac{\sum_{i=1}^{9} x_i^2 - \frac{\left(\sum_{i=1}^{9} x_i\right)^2}{n}}{n-1} = \frac{2579200 - \frac{22848400}{9}}{9-1} = 5061.1$$

$s = \sqrt{5061.1} = 71.14$

Yes, the rescaling is by a factor of 10. Therefore, s^2 and s would be rescaled by multiplying s^2 by 10^2 (resulting in 5061.1) and s by 10 (71.14). Subtracting 30 from each value has no affect on the variance or standard deviation. This is because $V(aX + b) = a^2 V(X)$.

6-88 $\sum_{i=1}^{n} (x_i - a)^2 = \sum_{i=1}^{n} (x_i - \bar{x}) + n(\bar{x} - a)^2$; The sum written in this form shows that the quantity is minimized when a = $\bar{x}$.

6-89 Of the two quantities $\sum_{i=1}^{n} (x_i - \bar{x})^2$ and $\sum_{i=1}^{n} (x_i - \mu)^2$, the quantity $\sum_{i=1}^{n} (x_i - \bar{x})^2$ will be smaller given that $\bar{x} \neq \mu$. This is because $\bar{x}$ is based on the values of the x_i's. The value of μ may be quite different for this sample.

6-90 $y_i = a + bx_i$

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \quad \text{and} \quad \bar{y} = \frac{\sum_{i=1}^{n}(a + bx_i)}{n} = \frac{na + b\sum_{i=1}^{n} x_i}{n} = a + b\bar{x}$$

$$s_x{}^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1} \quad \text{and} \quad s_x = \sqrt{s_x{}^2}$$

$$s_y{}^2 = \frac{\sum_{i=1}^{n}(a + bx_i - a + b\bar{x})^2}{n-1} = \frac{\sum_{i=1}^{n}(bx_i - b\bar{x})^2}{n-1} = \frac{b^2\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1} = b^2 s_x{}^2$$

Therefore, $s_y = bs_x$

6-91 $\bar{x} = 835.00$ °F $\quad s_x = 10.5$ °F

The results in °C:

$\bar{y} = -32 + 5/9\bar{x} = -32 + 5/9(835.00) = 431.89$ °C

$s_y{}^2 = b^2 s_x{}^2 = (5/9)^2(10.5)^2 = 34.028$ °C

6-92 Using the results found in Exercise 6-90 with a = $-\frac{\bar{x}}{s}$ and b = 1/*s*, the mean and standard deviation of the z_i are $\bar{z} = 0$ and $s_Z = 1$.

6-93. Yes, in this case, since no upper bound on the last electronic component is available, use a measure of central location that is not dependent on this value. That measure is the median.

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

6-94 a) $\bar{x}_{n+1} = \dfrac{\sum_{i=1}^{n+1} x_i}{n+1} = \dfrac{\sum_{i=1}^{n} x_i + x_{n+1}}{n+1}$

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

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

b) $ns_{n+1}^2 = \sum_{i=1}^{n} x_1^2 + x_{n+1}^2 - \dfrac{\left(\sum_{i=1}^{n} x_i + x_{n+1}\right)^2}{n+1}$

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

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

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

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

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

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

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

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

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

c) $\overline{x}_n = 65.811$ inches $\quad x_{n+1} = 64$

$s_n^{\ 2} = 4.435 \quad n = 37 \quad s_n = 2.106$

$$\overline{x}_{n+1} = \frac{37(65.81) + 64}{37 + 1} = 65.76$$

$$s_{n+1} = \sqrt{\frac{(37-1)4.435 + \frac{37}{37+1}(64 - 65.811)^2}{37}}$$

$$= 2.098$$

6-95. The trimmed mean is pulled toward the median by eliminating outliers.
a) 10% Trimmed Mean = 89.29
b) 20% Trimmed Mean = 89.19
Difference is very small
c) No, the differences are very small, due to a very large data set with no significant outliers.

6-96. If nT/100 is not an integer, calculate the two surrounding integer values and interpolate between the two. For example, if nT/100 = 2/3, one could calculate the mean after trimming 2 and 3 observations from each end and then interpolate between these two means.

CHAPTER 7

Section 7-2

7-1. $\text{E}\left(\overline{X}_1\right) = E\left(\frac{\sum_{i=1}^{2n} X_i}{2n}\right) = \frac{1}{2n} E\left(\sum_{i=1}^{2n} X_i\right) = \frac{1}{2n}(2n\mu) = \mu$

$\text{E}\left(\overline{X}_2\right) = E\left(\frac{\sum_{i=1}^{n} X_i}{n}\right) = \frac{1}{n} E\left(\sum_{i=1}^{n} X_i\right) = \frac{1}{n}(n\mu) = \mu$, $\overline{X}_1$ and $\overline{X}_2$ are unbiased estimators of μ.

The variances are $\text{V}\left(\overline{X}_1\right) = \frac{\sigma^2}{2\text{n}}$ and $\text{V}\left(\overline{X}_2\right) = \frac{\sigma^2}{\text{n}}$; compare the MSE (variance in this case),

$$\frac{MSE(\hat{\Theta}_1)}{MSE(\hat{\Theta}_2)} = \frac{\sigma^2/2n}{\sigma^2/n} = \frac{n}{2n} = \frac{1}{2}$$

Since both estimators are unbiased, examination of the variances would conclude that $\overline{\text{X}}_1$ is the "better" estimator with the smaller variance.

7-2. $\text{E}\left(\hat{\Theta}_1\right) = \frac{1}{7}\left[E(X_1) + E(X_2) + \cdots + E(X_7)\right] = \frac{1}{7}(7E(X)) = \frac{1}{7}(7\mu) = \mu$

$\text{E}\left(\hat{\Theta}_2\right) = \frac{1}{2}\left[E(2X_1) + E(X_6) + E(X_7)\right] = \frac{1}{2}[2\mu - \mu + \mu] = \mu$

a) Both $\hat{\Theta}_1$ and $\hat{\Theta}_2$ are unbiased estimates of μ since the expected values of these statistics are equivalent to the true mean, μ.

b) $\text{V}\left(\hat{\Theta}_1\right) = V\left[\frac{X_1 + X_2 + ... + X_7}{7}\right] = \frac{1}{7^2}\left(V(X_1) + V(X_2) + \cdots + V(X_7)\right) = \frac{1}{49}(7\sigma^2) = \frac{1}{7}\sigma^2$

$$V(\hat{\Theta}_1) = \frac{\sigma^2}{7}$$

$$\text{V}\left(\hat{\Theta}_2\right) = V\left[\frac{2X_1 - X_6 + X_4}{2}\right] = \frac{1}{2^2}\left(V(2X_1) + V(X_6) + V(X_4)\right) = \frac{1}{4}(4V(X_1) + V(X_6) + V(X_4))$$

$$= \frac{1}{4}\left(4\sigma^2 + \sigma^2 + \sigma^2\right)$$

$$= \frac{1}{4}(6\sigma^2)$$

$$V(\hat{\Theta}_2) = \frac{3\sigma^2}{2}$$

Since both estimators are unbiased, the variances can be compared to decide which is the better estimator. The variance of $\hat{\Theta}_1$ is smaller than that of $\hat{\Theta}_2$, $\hat{\Theta}_1$ is the better estimator.

7-3. Since both $\hat{\Theta}_1$ and $\hat{\Theta}_2$ are unbiased, the variances of the estimators can be examined to determine which is the "better" estimator. The variance of $\hat{\theta}_2$ is smaller than that of $\hat{\theta}_1$ thus $\hat{\theta}_2$ may be the better estimator.

$$\text{Relative Efficiency} = \frac{MSE(\hat{\Theta}_1)}{MSE(\hat{\Theta}_2)} = \frac{V(\bar{\Theta}_1)}{V(\hat{\Theta}_2)} = \frac{10}{4} = 2.5$$

7-4. Since both estimators are unbiased:

Relative Efficiency = $\dfrac{MSE(\hat{\Theta}_1)}{MSE(\hat{\Theta}_2)} = \dfrac{V(\hat{\Theta}_1)}{V(\hat{\Theta}_2)} = \dfrac{\sigma^2/7}{3\sigma^2/2} = \dfrac{2}{21}$

7-5. $\dfrac{MSE(\hat{\Theta}_1)}{MSE(\hat{\Theta}_2)} = \dfrac{V(\hat{\Theta}_1)}{V(\hat{\Theta}_2)} = \dfrac{10}{4} = 2.5$

7-6. $E(\hat{\Theta}_1) = \theta \qquad E(\hat{\Theta}_2) = \theta/2$

$Bias = E(\hat{\Theta}_2) - \theta$

$= \dfrac{\theta}{2} - \theta = -\dfrac{\theta}{2}$

$V(\hat{\Theta}_1) = 10 \qquad V(\hat{\Theta}_2) = 4$

For unbiasedness, use $\hat{\Theta}_1$ since it is the only unbiased estimator.

As for minimum variance and efficiency we have:

Relative Efficiency = $\dfrac{(V(\hat{\Theta}_1) + Bias^2)_1}{(V(\hat{\Theta}_2) + Bias^2)_2}$ where, Bias for θ_1 is 0.

Thus,

Relative Efficiency = $\dfrac{(10+0)}{\left(4 + \left(\dfrac{-\theta}{2}\right)^2\right)} = \dfrac{40}{\left(16+\theta^2\right)}$

If the relative efficiency is less than or equal to 1, $\hat{\Theta}_1$ is the better estimator.

Use $\hat{\Theta}_1$, when $\dfrac{40}{(16+\theta^2)} \le 1$

$$40 \le (16+\theta^2)$$
$$24 \le \theta^2$$
$$\theta \le -4.899 \text{ or } \theta \ge 4.899$$

If $-4.899 < \theta < 4.899$ then use $\hat{\Theta}_2$.

For unbiasedness, use $\hat{\Theta}_1$. For efficiency, use $\hat{\Theta}_1$ when $\theta \le -4.899$ or $\theta \ge 4.899$ and use $\hat{\Theta}_2$ when $-4.899 < \theta < 4.899$.

7-7.

$E(\hat{\Theta}_1) = \theta$	No bias	$V(\hat{\Theta}_1) = 12 = MSE(\hat{\Theta}_1)$
$E(\hat{\Theta}_2) = \theta$	No bias	$V(\hat{\Theta}_2) = 10 = MSE(\hat{\Theta}_2)$
$E(\hat{\Theta}_3) \ne \theta$	Bias	$MSE(\hat{\Theta}_3) = 6$ [note that this includes (bias^2)]

To compare the three estimators, calculate the relative efficiencies:

$\dfrac{MSE(\hat{\Theta}_1)}{MSE(\hat{\Theta}_2)} = \dfrac{12}{10} = 1.2$, since rel. eff. > 1 use $\hat{\Theta}_2$ as the estimator for θ

$\dfrac{MSE(\hat{\Theta}_1)}{MSE(\hat{\Theta}_3)} = \dfrac{12}{6} = 2$, since rel. eff. > 1 use $\hat{\Theta}_3$ as the estimator for θ

$\dfrac{MSE(\hat{\Theta}_2)}{MSE(\hat{\Theta}_3)} = \dfrac{10}{6} = 1.8$, 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.

7-8. $n_1 = 20, n_2 = 10, n_3 = 8$

Show that S^2 is unbiased:

$$E(S^2) = E\left(\frac{20S_1^2 + 10S_2^2 + 8S_3^2}{38}\right)$$
$$= \frac{1}{38}\left(E(20S_1^2) + E(10S_2^2) + E(8S_3^2)\right)$$
$$= \frac{1}{38}\left(20\sigma_1^2 + 10\sigma_2^2 + 8\sigma_3^2\right)$$
$$= \frac{1}{38}\left(38\sigma^2\right)$$
$$= \sigma^2$$

$\therefore$ S^2 is an unbiased estimator of σ^2.

7-9. Show that $\dfrac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}{n}$ is a biased estimator of σ^2:

a)

$$E\left(\frac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}{n}\right)$$
$$= \frac{1}{n}E\left(\sum_{i=1}^{n}\left(X_i - n\overline{X}\right)^2\right)$$
$$= \frac{1}{n}\left(\sum_{i=1}^{n}E\left(X_i^2\right) - nE\left(\overline{X}^2\right)\right)$$
$$= \frac{1}{n}\left(\sum_{i=1}^{n}\left(\mu^2 + \sigma^2\right) - n\left(\mu^2 + \frac{\sigma^2}{n}\right)\right)$$
$$= \frac{1}{n}\left(n\mu^2 + n\sigma^2 - n\mu^2 - \sigma^2\right)$$
$$= \frac{1}{n}((n-1)\sigma^2)$$
$$= \sigma^2 - \frac{\sigma^2}{n}$$

$\therefore$ $\dfrac{\sum\left(X_i - \overline{X}\right)^2}{n}$ is a biased estimator of σ^2.

b) Bias $= E\left[\dfrac{\sum\left(X_i^2 - n\overline{X}\right)^2}{n}\right] - \sigma^2 = \sigma^2 - \dfrac{\sigma^2}{n} - \sigma^2 = -\dfrac{\sigma^2}{n}$

c) Bias decreases as n increases.

7-10 a) Show that $\overline{X}^2$ is a biased estimator of μ. Using $E(X^2) = V(X) + [E(X)]^2$

$$E(\overline{X}^2) = \frac{1}{n^2} E\left(\sum_{i=1}^{n} X_i\right)^2$$

$$= \frac{1}{n^2}\left(V\left(\sum_{i=1}^{n} X_i\right) + \left[E\left(\sum_{i=1}^{n} X_i\right)\right]^2\right)$$

$$= \frac{1}{n^2}\left(n\sigma^2 + \left(\sum_{i=1}^{n} \mu\right)^2\right)$$

$$= \frac{1}{n^2}\left(n\sigma^2 + (n\mu)^2\right)$$

$$= \frac{1}{n^2}\left(n\sigma^2 + n^2\mu^2\right)$$

$$E(\overline{X}^2) = \frac{\sigma^2}{n} + \mu^2$$

$\therefore$ $\overline{X}^2$ is a biased estimator of $\mu.^2$

b) Bias = $E(\overline{X}^2) - \mu^2 = \frac{\sigma^2}{n} + \mu^2 - \mu^2 = \frac{\sigma^2}{n}$

c) Bias decreases as n increases.

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 75.615 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 75.2 pounds.

c.) Our estimate of the population variance is the sample variance or 2.738 square pounds. Similarly, our estimate of the population standard deviation is the sample standard deviation or 1.655 pounds.

d.) The standard error of the mean pull force, estimated from the data provided is 0.325 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

7-12. **Descriptive Statistics**

Variable	N	Mean	Median	TrMean	StDev	SE Mean
Oxide Thickness	24	423.33	424.00	423.36	9.08	1.85

a) The mean oxide thickness, as estimated by Minitab from the sample, is 423.33 Angstroms.
b) Standard deviation for the population can be estimated by the sample standard deviation, or 9.08 Angstroms.
c) The standard error of the mean is 1.85 Angstroms.
d) Our estimate for the median is 424 Angstroms.
e) Seven of the measurements exceed 430 Angstroms, so our estimate of the proportion requested is 7/24 = 0.2917

7.13 a)

$$E(X) = \int_{-1}^{1} x\frac{1}{2}(1+\theta x)dx = \int_{-1}^{1}\frac{1}{2}xdx + \int_{-1}^{1}\frac{\theta}{2}x^2dx$$
$$= 0 + \frac{\theta}{3} = \frac{\theta}{3}$$

b) Let $\overline{X}$ be the sample average of the observations in the random sample. We know that $E(\overline{X}) = \mu$, the mean of the distribution. However, the mean of the distribution is $\theta/3$, so $\hat{\theta} = 3\overline{X}$ is an unbiased estimator of θ.

7.14 a.) $E(\hat{p}) = E(X/n) = \frac{1}{n}E(X) = \frac{1}{n}np = p$

b.) We know that the variance of $\hat{p}$ is $\frac{p\cdot(1-p)}{n}$ so its standard error must be $\sqrt{\frac{p\cdot(1-p)}{n}}$. To estimate this parameter we would substitute our estimate of p into it.

7.15 a.) $E(\overline{X}_1 - \overline{X}_2) = E(\overline{X}_1) - E(\overline{X}_2) = \mu_1 - \mu_2$

b.) $s.e. = \sqrt{V(\overline{X}_1 - \overline{X}_2)} = \sqrt{V(\overline{X}_1) + V(\overline{X}_2) + 2COV(\overline{X}_1, \overline{X}_2)} = \sqrt{\frac{\sigma_1^{\ 2}}{n_1} + \frac{\sigma_2^{\ 2}}{n_2}}$

This standard error could be estimated by using the estimates for the standard deviations of populations 1 and 2.

7-16

$$E(S_p^{\,2}) = E\left(\frac{(n_1-1)\cdot S_1^{\,2} + (n_2-1)\cdot S_2^{\,2}}{n_1+n_2-2}\right) = \frac{1}{n_1+n_2-2}\left[(n_1-1)E(S_1^{\,2}) + (n_2-1)\cdot E(S_2^{\,2})\right] =$$

$$= \frac{1}{n_1+n_2-2}\left[(n_1-1)\cdot \sigma_1^{\,2} + (n_2-1)\cdot \sigma_2^{\,2})\right] = \frac{n_1+n_2-2}{n_1+n_2-2}\sigma^2 = \sigma^2$$

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:

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

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

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

7-18 a.) $E(\frac{X_1}{n_1} - \frac{X_2}{n_2}) = \frac{1}{n_1}E(X_1) - \frac{1}{n_2}E(X_2) = \frac{1}{n_1}n_1p_1 - \frac{1}{n_2}n_2p_2 = p_1 - p_2 = E(p_1 - p_2)$

b.) $\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$

c.) An estimate of the standard error could be obtained substituting $\frac{X_1}{n_1}$ for p_1 and $\frac{X_2}{n_2}$ for p_2 in the equation shown in (b).

d.) Our estimate of the difference in proportions is 0.01

e.) The estimated standard error is 0.0413

Section 7-3

7-19. $f(x)=\dfrac{e^{-\lambda}\lambda^{x}}{x!}$ $\quad L(\lambda)=\prod_{i=1}^{n}\dfrac{e^{-\lambda}\lambda^{x_i}}{x_i!}=\dfrac{e^{-n\lambda}\lambda^{\sum_{i=1}^{n}x_i}}{\prod_{i=1}^{n}x_i!}$

$$\ln L(\lambda)=-n\lambda\ln e+\sum_{i=1}^{n}x_i\ln\lambda-\sum_{i=1}^{n}\ln x_i!$$

$$\frac{d\ln L(\lambda)}{d\lambda}=-n+\frac{1}{\lambda}\sum_{i=1}^{n}x_i\equiv 0$$

$$-n+\frac{\sum_{i=1}^{n}x_i}{\lambda}=0$$

$$\sum_{i=1}^{n}x_i=n\lambda$$

$$\hat{\lambda}=\frac{\sum_{i=1}^{n}x_i}{n}$$

7-20. $f(x)=\lambda e^{-\lambda(x-\theta)}$ for $x\ge\theta$ $\quad L(\lambda)=\prod_{i=1}^{n}\lambda e^{-\lambda(x-\theta)}=\lambda^{n}e^{-\lambda\sum_{i=1}^{n}(x-\theta)}=\lambda^{n}e^{-\lambda\left(\sum_{i=1}^{n}x-n\theta\right)}$

$$\ln L(\lambda,\theta)=-n\ln\lambda-\lambda\sum_{i=1}^{n}x_i-\lambda n\theta$$

$$\frac{d\ln L(\lambda,\theta)}{d\lambda}=-\frac{n}{\lambda}-\sum_{i=1}^{n}x_i-n\theta\equiv 0$$

$$-\frac{n}{\lambda}=\sum_{i=1}^{n}x_i-n\theta$$

$$\hat{\lambda}=n\Big/\left(\sum_{i=1}^{n}x_i-n\theta\right)$$

$$\hat{\lambda}=\frac{1}{\bar{x}-\theta}$$

Let $\lambda=1$ then $L(\theta)=e^{-\sum_{i=1}^{n}(x_{i-\theta})}$ for $x_{i\ge}0$

$L(\theta)=e^{-\sum_{i=1}^{n}x_{i-n\theta}}=e^{n\theta-n\bar{x}}$

$LnL(\theta)=n\theta-n\bar{x}$

θ cannot be estimated using ML equations since

$\dfrac{dLnL(\theta)}{d(\theta)}=0$. Therefore, θ is estimated using $Min(X_1,X_2,\ldots,X_n)$.

Ln(θ) is maximized at $x_{\min}$ and $\hat{\theta}=x_{\min}$

b.) Example: Consider traffic flow and let the time that has elapsed between one car passing a fixed point and the instant that the next car begins to pass that point be considered time headway. This headway can be modeled by the shifted exponential distribution.

Example in Reliability: Consider a process where failures are of interest. Say that a sample or population is put into operation at x = 0, but no failures will occur until θ period of operation. Failures will occur after the time θ.

7-21. $f(x) = p(1-p)^{x-1}$

$$L(p) = \prod_{i=1}^{n} p(1-p)^{x_i - 1}$$

$$= p^n (1-p)^{\sum_{i=1}^{n} x_i - n}$$

$$\ln L(p) = n \ln p + \left(\sum_{i=1}^{n} x_i - n \right) \ln(1-p)$$

$$\frac{\partial \ln L(p)}{\partial p} = \frac{n}{p} - \frac{\sum_{i=1}^{n} x_i - n}{1-p} \equiv 0$$

$$0 = \frac{(1-p)n - p\left(\sum_{i=1}^{n} x_i - n \right)}{p(1-p)}$$

$$0 = \frac{n - np - p\sum_{i=1}^{n} x_i + pn}{p(1-p)}$$

$$0 = n - p\sum_{i=1}^{n} x_i$$

$$\hat{p} = \frac{n}{\sum_{i=1}^{n} x_i}$$

7-22. $f(x) = (\theta + 1)x^{\theta}$

$$L(\theta) = \prod_{i=1}^{n} (\theta+1)x_i^{\theta} = (\theta+1)x_1^{\theta} \times (\theta+1)x_2^{\theta} \times \ldots$$

$$= (\theta+1)^n \prod_{i=1}^{n} x_i^{\theta}$$

$$\ln L(\theta) = n\ln(\theta+1) + \theta \ln x_1 + \theta \ln x_2 + \ldots$$

$$= n\ln(\theta+1) + \theta \sum_{i=1}^{n} \ln x_i$$

$$\frac{\partial \ln L(\theta)}{\partial \theta} = \frac{n}{\theta+1} + \sum_{i=1}^{n} \ln x_i = 0$$

$$\frac{n}{\theta+1} = -\sum_{i=1}^{n} \ln x_i$$

$$\hat{\theta} = \frac{n}{-\sum_{i=1}^{n} \ln x_i} - 1$$

7-23. a)

$$L(\beta,\delta)=\prod_{i=1}^{n}\frac{\beta}{\delta}\left(\frac{x_i}{\delta}\right)^{\beta-1}e^{-\left(\frac{x_i}{\delta}\right)^{\beta}}$$

$$=e^{-\sum_{i=1}^{n}\left(\frac{x_i}{\delta}\right)^{\beta}}\prod_{i=1}^{n}\frac{\beta}{\delta}\left(\frac{x_i}{\delta}\right)^{\beta-1}$$

$$\ln L(\beta,\delta)=\sum_{i=1}^{n}\ln\left[\frac{\beta}{\delta}\left(\frac{x_i}{\delta}\right)^{\beta-1}\right]-\sum\left(\frac{x_i}{\delta}\right)^{\beta}$$

$$=n\ln(\tfrac{\beta}{\delta})+(\beta-1)\sum\ln\left(\tfrac{x_i}{\delta}\right)-\sum\left(\tfrac{x_i}{\delta}\right)^{\beta}$$

b)

$$\frac{\partial\ln L(\beta,\delta)}{\partial\beta}=\frac{n}{\beta}+\sum\ln\left(\tfrac{x_i}{\delta}\right)-\sum\ln\left(\tfrac{x_i}{\delta}\right)\left(\tfrac{x_i}{\delta}\right)^{\beta}$$

$$\frac{\partial\ln L(\beta,\delta)}{\partial\delta}=-\frac{n}{\delta}-(\beta-1)\frac{n}{\delta}+\beta\frac{\sum x_i^{\beta}}{\delta^{\beta+1}}$$

Upon setting $\frac{\partial\ln L(\beta,\delta)}{\partial\delta}$ equal to zero, we obtain

$$\delta^{\beta}n=\sum x_i^{\beta}\quad and\quad \delta=\left[\frac{\sum x_i^{\beta}}{n}\right]^{1/\beta}$$

Upon setting $\frac{\partial\ln L(\beta,\delta)}{\partial\beta}$ equal to zero and substituting for δ, we obtain

$$\frac{n}{\beta}+\sum\ln x_i-n\ln\delta=\frac{1}{\delta^{\beta}}\sum x_i^{\beta}(\ln x_i-\ln\delta)$$

$$\frac{n}{\beta}+\sum\ln x_i-\frac{n}{\beta}\ln\left(\frac{\sum x_i^{\beta}}{n}\right)=\frac{n}{\sum x_i^{\beta}}\sum x_i^{\beta}\ln x_i-\frac{n}{\sum x_i^{\beta}}\sum x_i^{\beta}\frac{1}{\beta}\ln\left(\frac{\sum x_i^{\beta}}{n}\right)$$

and $\frac{1}{\beta}=\left[\frac{\sum x_i^{\beta}\ln x_i}{\sum x_i^{\beta}}+\frac{\sum\ln x_i}{n}\right]$

c) Numerical iteration is required.

7-24 $$\hat{\theta}=\frac{1}{n}\sum_{i=1}^{n}(\theta+1)x_i^{\theta}=\frac{(\theta+1)}{n}\sum_{i=1}^{n}x_i^{\theta}$$

7-25 $E(X)=\frac{a-0}{2}=\frac{1}{n}\sum_{i=1}^{n}X_i=\overline{X}$, therefore: $\hat{a}=2\overline{X}$

The expected value of this estimate is the true parameter so it must be unbiased. This estimate is reasonable in one sense because it is unbiased. However, there are obvious problems. Consider the sample $x_1=1$, $x_2=2$ and $x_3=10$. Now $\overline{x}=4.37$ and $\hat{a}=2\overline{x}=8.667$. This is an unreasonable estimate of a, because clearly $a\geq 10$.

7-26. a) $\hat{a}$ cannot be unbiased since it will always be less than a.

b) bias = $\frac{na}{n+1} - \frac{a(n+1)}{n+1} = -\frac{a}{n+1} \underset{n\to\infty}{\to} 0$.

c) $2\overline{X}$

d) $P(Y \le y) = P(X_1, \ldots, X_n \le y) = (P(X_1 \le y))^n = \left(\frac{y}{a}\right)^n$. Thus, f(y) is as given. Thus,

bias=E(Y)-a = $\frac{an}{n+1} - a = -\frac{a}{n+1}$.

7-27 For any $n>1$ $n(n+2) > 3n$ so the variance of $\hat{a}_2$ is less than that of $\hat{a}_1$. It is in this sense that the second estimator is better than the first.

7-28

$$L(\theta) = \prod_{i=1}^{n} \frac{x_i e^{-x_i/\theta}}{\theta^2} \qquad \ln L(\theta) = \sum \ln(x_i) - \sum \frac{x_i}{\theta} - 2n \ln \theta$$

$$\frac{\partial \ln L(\theta)}{\partial \theta} = \frac{1}{\theta^2} \sum x_i - \frac{2n}{\theta}$$

setting the last equation equal to zero and solving for theta, we find:

$$\hat{\theta} = \frac{\sum_{i=1}^{n} x_i}{2n}$$

7-29 a) $$E(X^2) = 2\theta = \frac{1}{n} \sum_{i=1}^{n} X_i^2 \text{ so}$$

$$\hat{\Theta} = \frac{1}{2n} \sum_{i=1}^{n} X_i^2$$

b)

$$L(\theta) = \prod_{i=1}^{n} \frac{x_i e^{-x_i^2/2\theta}}{\theta} \qquad \ln L(\theta) = \sum \ln(x_i) - \sum \frac{x_i^2}{2\theta} - n \ln \theta$$

$$\frac{\partial \ln L(\theta)}{\partial \theta} = \frac{1}{2\theta^2} \sum x_i^2 - \frac{n}{\theta}$$

setting the last equation equal to zero, we obtain the maximum likelihood estimate

$$\hat{\Theta} = \frac{1}{2n} \sum_{i=1}^{n} X_i^2$$

which is the same result we obtained in part (a) c)

$$\int_0^a f(x)dx = 0.5 = 1 - e^{-a^2/2\theta}$$

$$a = \sqrt{-2\theta \ln(0.5)}$$

We can estimate the median (*a*) by substituting our estimate for theta into the equation for *a*.

7-30 a) $\int_{-1}^{1} c(1+\theta x)dx = 1 = (cx + c\theta \frac{x^2}{2})\Big|_{-1}^{1} = 2c$

so the constant *c* should equal 0.5

b)

$$E(X) = \frac{1}{n}\sum_{i=1}^{n} X_i = \frac{\theta}{3}$$

$$\theta = 3 \cdot \frac{1}{n}\sum_{i=1}^{n} X_i$$

c)

$$E(\hat{\theta}) = E\left(3 \cdot \frac{1}{n}\sum_{i=1}^{n} X_i\right) = E(3\overline{X}) = 3E(\overline{X}) = 3\frac{\theta}{3} = \theta$$

d)

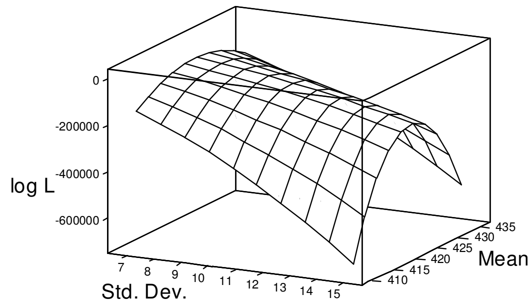
$$L(\theta) = \prod_{i=1}^{n} \frac{1}{2}(1+\theta X_i) \qquad \ln L(\theta) = n\ln(\frac{1}{2}) + \sum_{i=1}^{n} \ln(1+\theta X_i)$$

$$\frac{\partial \ln L(\theta)}{\partial \theta} = \sum_{i=1}^{n} \frac{X_i}{(1+\theta X_i)}$$

by inspection, the value of θ that maximizes the likelihood is max (X_i)

7-31 a)Using the results from Example 7-12 we obtain that the estimate of the mean is 423.33 and the estimate of the variance is 82.4464

b)



The function seems to have a ridge and its curvature is not too pronounced. The maximum value for std deviation is at 9.08, although it is difficult to see on the graph.

7-32 When n is increased to 40, the graph will look the same although the curvature will be more pronounced. As n increases it will be easier to determine the where the maximum value for the standard deviation is on the graph.

Section 7-5

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

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

$$= 0.9772 - 0.1586 = 0.8186$$

7-34. $X_i \sim N(100, 10^2)$ $\quad n = 25$

$$\mu_{\overline{X}} = 100 \quad \sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}} = \frac{10}{\sqrt{25}} = 2$$

$$P[(100 - 1.8(2)) \le \overline{X} \le (100 + 2)] = P(96.4 \le \overline{X} \le 102)$$

$$= P\left(\frac{96.4-100}{2} \le \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \le \frac{102-100}{2} \right)$$

$$= P(-1.8 \le Z \le 1) = P(Z \le 1) - P(Z \le -1.8)$$

$$= 0.8413 - 0.0359 = 0.8054$$

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

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

$$= P(Z \geq 0.175) = 1 - P(Z \leq 0.175)$$

$$= 1 - 0.56945 = 0.43055$$

7-36.

n = 6	n = 49
$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{3.5}{\sqrt{6}}$	$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{3.5}{\sqrt{49}}$
$= 1.429$	$= 0.5$
$\sigma_{\bar{X}}$ is reduced by 0.929 psi	

7-37. Assuming a normal distribution,

$\mu_{\bar{X}} = 2500 \quad \sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{50}{\sqrt{5}} = 22.361$

$$P(2499 \leq \bar{X} \leq 2510) = P\left(\frac{2499-2500}{22.361} \leq \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \leq \frac{2510-2500}{22.361}\right)$$

$$= P(-0.045 \leq Z \leq 0.45) = P(Z \leq 0.45) - P(Z \leq -0.045)$$

$$= 0.6736 - 0.4821 = 0.1915$$

7-38. $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{50}{\sqrt{5}} = 22.361\,psi$ = standard error of $\bar{X}$

7-39. $\sigma^2 = 25$

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

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

$$= \left(\frac{5}{1.5}\right)^2$$

$$= 11.11 \sim 12$$

7-40. Let $Y = \overline{X} - 6$

$$\mu_X = \frac{a+b}{2} = \frac{(0+1)}{2} = \tfrac{1}{2}$$

$$\mu_{\overline{X}} = \mu_X$$

$$\sigma_X^2 = \frac{(b-a)^2}{12} = \tfrac{1}{12}$$

$$\sigma_{\overline{X}}^2 = \frac{\sigma^2}{n} = \frac{\frac{1}{12}}{12} = \frac{1}{144}$$

$$\sigma_{\overline{X}} = \tfrac{1}{12}$$

$$\mu_Y = \tfrac{1}{2} - 6 = -5\tfrac{1}{2}$$

$$\sigma_Y^2 = \tfrac{1}{144}$$

$Y = \overline{X} - 6 \sim N(-5\frac{1}{2}, \frac{1}{144})$, approximately, using the central limit theorem.

7-41. $n = 36$

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

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

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

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

Using the central limit theorem:

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

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

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

$$= 1 - 0.7688 = 0.2312$$

7-42.

$\mu_X = 8.2$ minutes	n = 49
$\sigma_X = 1.5$ minutes	$\sigma_{\overline{X}} = \frac{\sigma_X}{\sqrt{n}} = \frac{1.5}{\sqrt{49}} = 0.2143$ $\mu_{\overline{X}} = \mu_X = 8.2$ mins

Using the central limit theorem, $\overline{X}$ is approximately normally distributed.

a) $P(\overline{X} < 10) = P(Z < \frac{10 - 8.2}{0.2143}) = P(Z < 8.4) = 1$

b) $P(5 < \overline{X} < 10) = P(\frac{5-8.2}{0.2143} < Z < \frac{10-8.2}{0.2143})$

$$= P(Z < 8.4) - P(Z < -14.932) = 1 - 0 = 1$$

c) $P(\overline{X} < 6) = P(Z < \frac{6 - 8.2}{0.2143}) = P(Z < -10.27) = 0$

7-43.

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

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

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

$= 1 - 0.4115 = 0.5885$

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

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

$= 0.5445 - 0.3687 = 0.1759$

7-44. If $\mu_B = \mu_A$, then $\overline{X}_B - \overline{X}_A$ is approximately normal with mean 0 and variance $\frac{\sigma_B^2}{25} + \frac{\sigma_A^2}{25} = 20.48$.

Then, $P(\overline{X}_B - \overline{X}_A > 3.5) = P(Z > \frac{3.5-0}{\sqrt{20.48}}) = P(Z > 0.773) = 0.2196$

The probability that $\overline{X}_B$ exceeds $\overline{X}_A$ by 3.5 or more is not that unusual when μ_B *and* μ_A are equal. Therefore, there is not strong evidence that μ_B is greater than μ_A.

7-45. Assume approximate normal distributions.

$(\overline{X}_{high} - \overline{X}_{low}) \sim N(60-55, \frac{4^2}{16} + \frac{4^2}{16})$

$\sim N(5,2)$

$P(\overline{X}_{high} - \overline{X}_{low} \ge 2) = P(Z \ge \frac{2-5}{\sqrt{2}}) = 1 - P(Z \le -2.12) = 1 - 0.0170 = 0.983$

Supplemental Exercises

7-46. $$f(x_1, x_2, x_3, x_4, x_5) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\sum_{i=1}^{5} \frac{(x_i-\mu)^2}{2\sigma^2}}$$

7-47. $$f(x_1, x_2, ..., x_n) = \prod_{i=1}^{n} \lambda e^{-\lambda x_i} \quad for \quad x_1 > 0, x_2 > 0, ..., x_n > 0$$

7-48. $f(x_1, x_2, x_3, x_4) = 1 \quad for \quad 0 \le x_1 \le 1, 0 \le x_2 \le 1, 0 \le x_3 \le 1, 0 \le x_4 \le 1$

7-49.

$$\overline{X}_1 - \overline{X}_2 \sim N(100-105, \frac{1.5^2}{25} + \frac{2^2}{30})$$

$$\sim N(-5, 0.2233)$$

7-50. $\sigma_{\overline{X}_1-\overline{X}_2} = \sqrt{0.2233} = 0.4726$

7-51. $X \sim N(50,144)$

$$P(47 \le \overline{X} \le 53) = P\left(\frac{47-50}{12/\sqrt{36}} \le Z \le \frac{53-50}{12/\sqrt{36}}\right)$$

$$= P(-1.5 \le Z \le 1.5)$$

$$= P(Z \le 1.5) - P(Z \le -1.5)$$

$$= 0.9332 - 0.0668 = 0.8664$$

7-52. No, because Central Limit Theorem states that with large samples ($n \ge 30$), $\overline{X}$ is approximately normally distributed.

7-53. Assume $\overline{X}$ is approximately normally distributed.

$$P(\overline{X} > 4985) = 1 - P(\overline{X} \le 4985) = 1 - P(Z \le \tfrac{4985-5500}{100/\sqrt{9}})$$

$$= 1 - P(Z \le -15.45) = 1 - 0 = 1$$

7-54. $$t = \frac{\overline{X}-\mu}{s/\sqrt{n}} = \frac{52-50}{\sqrt{2/16}} = 5.6569$$

$t_{.05,15} = 1.753$. Since 5.33 >> $t_{.05,15}$, the results are <u>very unusual</u>.

7-55. $P(\overline{X} \le 37) = P(Z \le -5.36) = 0$

7-56. Binomial with p equal to the proportion of defective chips and n = 100.

7-57. $E(a\overline{X}_1 + (1-a)\overline{X}_2 = a\mu + (1-a)\mu = \mu$

$$V(\overline{X}) = V[a\overline{X}_1 + (1-a)\overline{X}_2]$$
$$= a^2V(\overline{X}_1) + (1-a)^2V(\overline{X}_2)$$
$$= a^2(\tfrac{\sigma^2}{n_1}) + (1-2a+a^2)(\tfrac{\sigma^2}{n_2})$$
$$= \frac{a^2\sigma^2}{n_1} + \frac{\sigma^2}{n_2} - \frac{2a\sigma^2}{n_2} + \frac{a^2\sigma^2}{n_2}$$
$$= (n_2a^2 + n_1 - 2n_1a + n_1a^2)(\frac{\sigma^2}{n_1n_2})$$

$$\frac{\partial V(\overline{X})}{\partial a} = (\frac{\sigma^2}{n_1n_2})(2n_2a - 2n_1 + 2n_1a) \equiv 0$$
$$0 = 2n_2a - 2n_1 + 2n_1a$$
$$2a(n_2+n_1) = 2n_1$$
$$a(n_2+n_1) = n_1$$
$$a = \frac{n_1}{n_2+n_1}$$

7-58

$$L(\theta) = \left(\frac{1}{2\theta^3}\right)^n e^{\sum_{i=1}^{n} \frac{-x_i}{\theta}} \prod_{i=1}^{n} x_i^2$$

$$\ln L(\theta) = n \ln\left(\frac{1}{2\theta^3}\right) + 2\sum_{i=1}^{n} \ln x_i - \sum_{i=1}^{n} \frac{x_i}{\theta}$$

$$\frac{\partial \ln L(\theta)}{\partial \theta} = \frac{-3n}{\theta} + \sum_{i=1}^{n} \frac{x_i}{\theta^2}$$

Making the last equation equal to zero and solving for theta, we obtain:

$$\hat{\Theta} = \frac{\sum_{i=1}^{n} x_i}{3n}$$

as the maximum likelihood estimate.

7-59

$$L(\theta) = \theta^n \prod_{i=1}^{n} x_i^{\theta - 1}$$

$$\ln L(\theta) = n \ln \theta + (\theta - 1) \sum_{i=1}^{n} \ln(x_i)$$

$$\frac{\partial \ln L(\theta)}{\partial \theta} = \frac{n}{\theta} + \sum_{i=1}^{n} \ln(x_i)$$

making the last equation equal to zero and solving for theta, we obtain the maximum likelihood estimate.

$$\hat{\Theta} = \frac{-n}{\sum_{i=1}^{n} \ln(x_i)}$$

7-60

$$L(\theta) = \frac{1}{\theta^n} \prod_{i=1}^{n} x_i^{\frac{1-\theta}{\theta}}$$

$$\ln L(\theta) = -n \ln \theta + \frac{1-\theta}{\theta} \sum_{i=1}^{n} \ln(x_i)$$

$$\frac{\partial \ln L(\theta)}{\partial \theta} = -\frac{n}{\theta} - \frac{1}{\theta^2} \sum_{i=1}^{n} \ln(x_i)$$

making the last equation equal to zero and solving for the parameter of interest, we obtain the maximum likelihood estimate.

$$\hat{\Theta} = -\frac{1}{n} \sum_{i=1}^{n} \ln(x_i)$$

$$E(\hat{\theta}) = E\left[-\frac{1}{n} \sum_{i=1}^{n} \ln(x_i)\right] = \frac{1}{n} E\left[-\sum_{i=1}^{n} \ln(x_i)\right] = -\frac{1}{n} \sum_{i=1}^{n} E[\ln(x_i)]$$

$$= \frac{1}{n} \sum_{i=1}^{n} \theta = \frac{n\theta}{n} = \theta$$

$$E(\ln(X_i)) = \int_0^1 (\ln x) x^{\frac{1-\theta}{\theta}} dx$$ let $u = \ln x$ and $dv = x^{\frac{1-\theta}{\theta}} dx$

then, $E(\ln(X)) = -\theta \int_0^1 x^{\frac{1-\theta}{\theta}} dx = -\theta$

Mind-Expanding Exercises

7-61. $P(X_1 = 0, X_2 = 0) = \frac{M(M-1)}{N(N-1)}$

$P(X_1 = 0, X_2 = 1) = \frac{M(N-M)}{N(N-1)}$

$P(X_1 = 1, X_2 = 0) = \frac{(N-M)M}{N(N-1)}$

$P(X_1 = 1, X_2 = 1) = \frac{(N-M)(N-M-1)}{N(N-1)}$

$P(X_1 = 0) = M/N$

$P(X_1 = 1) = \frac{N-M}{N}$

$$P(X_2 = 0) = P(X_2 = 0 \mid X_1 = 0)P(X_1 = 0) + P(X_2 = 0 \mid X_1 = 1)P(X_1 = 1)$$
$$= \frac{M-1}{N-1} \times \frac{M}{N} + \frac{M}{N-1} \times \frac{N-M}{N} = \frac{M}{N}$$

$$P(X_2 = 1) = P(X_2 = 1 \mid X_1 = 0)P(X_1 = 0) + P(X_2 = 1 \mid X_1 = 1)P(X_1 = 1)$$
$$= \frac{N-M}{N-1} \times \frac{M}{N} + \frac{N-M-1}{N-1} \times \frac{N-M}{N} = \frac{N-M}{N}$$

Because $P(X_2 = 0 \mid X_1 = 0) = \frac{M-1}{N-1}$ is not equal to $P(X_2 = 0) = \frac{M}{N}$, X_1 *and* X_2 are not independent.

7-62 a)

$$c_n = \frac{\Gamma[(n-1)/2]}{\Gamma(n/2)\sqrt{2/(n-1)}}$$

b.) When $n = 10$, $c_n = 1.0281$. When $n = 25$, $c_n = 1.0105$. So S is a pretty good estimator for the standard deviation even when relatively small sample sizes are used.

7-63 (a) $Z_i = Y_i - X_i$; so Z_i is N(0, $2\sigma^2$). Let $\sigma^{*2} = 2\sigma^2$. The likelihood function is

$$L(\sigma^{*2}) = \prod_{i=1}^{n} \frac{1}{\sigma^* \sqrt{2\pi}} e^{-\frac{1}{2\sigma^{*2}}(z_i^2)}$$
$$= \frac{1}{(\sigma^{*2}\, 2\pi)^{n/2}} e^{-\frac{1}{2\sigma^{*2}}\sum_{i=1}^{n} z_i^2}$$

The log-likelihood is $\ln[L(\sigma^{*2})] = -\frac{n}{2}(2\pi\sigma^{*2}) - \frac{1}{2\sigma^{*2}}\sum_{i=1}^{n} z_i^2$

Finding the maximum likelihood estimator:

$$\frac{d\ln[L(\sigma^{*2})]}{d\sigma^*} = -\frac{n}{2\sigma^{*2}} + \frac{1}{2\sigma^{*4}}\sum_{i=1}^{n} z_i^2 = 0$$
$$n\sigma^{*2} = \sum_{i=1}^{n} z_i^2$$
$$\hat{\sigma}^{*2} = \frac{1}{n}\sum_{i=1}^{n} z_i^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - x_i)^2$$

But $\sigma^{*2} = 2\sigma^2$, so the MLE is

$$2\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - x_i)^2$$
$$\hat{\sigma}^2 = \frac{1}{2n}\sum_{i=1}^{n}(y_i - x_i)^2$$

b)

$$E(\hat{\sigma}^2) = E\left[\frac{1}{2n}\sum_{i=1}^{n}(Y_i - X_i)^2\right]$$
$$= \frac{1}{2n}\sum_{i=1}^{n} E(Y_i - X_i)^2$$
$$= \frac{1}{2n}\sum_{i=1}^{n} E(Y_i^2 - 2Y_iX_i + X_i^2)$$
$$= \frac{1}{2n}\sum_{i=1}^{n}[E(Y_i^2) - E(2Y_iX_i) + E(X_i^2)]$$
$$= \frac{1}{2n}\sum_{i=1}^{n}[\sigma^2 - 0 + \sigma^2]$$
$$= \frac{2n\sigma^2}{2n}$$
$$= \sigma^2$$

So the estimator is unbiased.

7-64. $P\left(|\overline{X}-\mu| \geq \frac{c\sigma}{\sqrt{n}}\right) \leq \frac{1}{c^2}$ from Chebyshev's inequality.

Then, $P\left(|\overline{X}-\mu| < \frac{c\sigma}{\sqrt{n}}\right) \geq 1-\frac{1}{c^2}$. Given an ε, n and c can be chosen sufficiently large that the last probability is near 1 and $\frac{c\,\sigma}{\sqrt{n}}$ is equal to ε.

7-65 $P\left(X_{(n)} \leq t\right) = P\left(X_i \leq t \;\; for \;\; i = 1,...,n\right) = [F(t)]^n$

$P\left(X_{(1)} > t\right) = P\left(X_i > t \;\; for \;\; i = 1,...,n\right) = [1-F(t)]^n$

Then, $P\left(X_{(1)} \leq t\right) = 1-[1-F(t)]^n$

$f_{X_{(1)}}(t) = \frac{\partial}{\partial t}F_{X_{(1)}}(t) = n[1-F(t)]^{n-1}f(t)$

$f_{X_{(n)}}(t) = \frac{\partial}{\partial t}F_{X_{(n)}}(t) = n[F(t)]^{n-1}f(t)$

7-66 $P\left(X_{(1)} = 0\right) = F_{X_{(1)}}(0) = 1-[1-F(0)]^n = 1-p^n$ because F(0) = 1 - p.

$P\left(X_{(n)} = 1\right) = 1-F_{X_{(n)}}(0) = 1-[F(0)]^n = 1-(1-p)^n$

7-67. $P(X \leq t) = F(t) = \Phi\left[\frac{t-\mu}{\sigma}\right]$. From Exercise 7-65,

$f_{X_{(1)}}(t) = n\left\{1-\Phi\left[\frac{t-\mu}{\sigma}\right]\right\}^{n-1}\frac{1}{\sqrt{2\pi}\sigma}\;e^{-\frac{(t-\mu)^2}{2\sigma^2}}$

$f_{X_{(n)}}(t) = n\left\{\Phi\left[\frac{t-\mu}{\sigma}\right]\right\}^{n-1}\frac{1}{\sqrt{2\pi}\sigma}\;e^{-\frac{(t-\mu)^2}{2\sigma^2}}$

7-68. $P(X \leq t) = 1-e^{-\lambda t}$. From Exercise 7-65,

$F_{X_{(1)}}(t) = 1-e^{-n\lambda t}$	$f_{X_{(1)}}(t) = n\lambda e^{-n\lambda t}$
$F_{X_{(n)}}(t) = [1-e^{-\lambda t}]^n$	$f_{X_{(n)}}(t) = n[1-e^{-\lambda t}]^{n-1}\lambda e^{-\lambda t}$

7-69. $P\left(F\left(X_{(n)}\right) \le t\right) = P\left(X_{(n)} \le F^{-1}(t)\right) = t^n$ from Exercise 7-65 for $0 \le t \le 1$.

If $Y = F(X_{(n)})$, then $f_Y(y) = ny^{n-1}, 0 \le y \le 1$. Then, $E(Y) = \int_0^1 ny^n dy = \frac{n}{n+1}$

$P\left(F\left(X_{(1)}\right) > t\right) = P\left(X_{(1)} < F^{-1}(t)\right) = 1 - (1-t)^n$ from Exercise 7-65 for $0 \le t \le 1$.

If $Y = F(X_{(1)})$, then $f_Y(y) = n(1-t)^{n-1}, 0 \le y \le 1$.

Then, $E(Y) = \int_0^1 yn(1-y)^{n-1} dy = \frac{1}{n+1}$ where integration by parts is used. Therefore,

$E[F(X_{(n)})] = \frac{n}{n+1}$ *and* $E[F(X_{(1)})] = \frac{1}{n+1}$

7-70.
$$\begin{aligned} E(V) &= k\sum_{i=1}^{n-1} [E(X_{i+1}^2) + E(X_i^2) - 2E(X_i X_{i+1})] \\ &= k\sum_{i=1}^{n-1} (\sigma^2 + \mu^2 + \sigma^2 + \mu^2 - 2\mu^2) \\ &= k(n-1)2\sigma^2 \end{aligned}$$
Therefore, $k = \frac{1}{2(n-1)}$

7-71 a.)The traditional estimate of the standard deviation, S, is 3.26. The mean of the sample is 13.43 so the values of $\left|X_i - \overline{X}\right|$ corresponding to the given observations are 3.43, 1.43, 4.43, 0.57, 4.57, 1.57 and 2.57. The median of these new quantities is 2.57 so the new estimate of the standard deviation is 3.81; slightly larger than the value obtained with the traditional estimator.
b.) Making the first observation in the original sample equal to 50 produces the following results. The traditional estimator, S, is equal to 13.91. The new estimator remains unchanged.

7-72 a.)
$$\begin{aligned} T_r = X_1 &+ \\ &X_1 + X_2 - X_1 + \\ &X_1 + X_2 - X_1 + X_3 - X_2 + \\ &\ldots + \\ &X_1 + X_2 - X_1 + X_3 - X_2 + \ldots + X_r - X_{r-1} + \\ &(n-r)(X_1 + X_2 - X_1 + X_3 - X_2 + \ldots + X_r - X_{r-1}) \end{aligned}$$

Because X_1 is the minimum lifetime of n items, $E(X_1) = \frac{1}{n\lambda}$.

Then, $X_2 - X_1$ is the minimum lifetime of (n-1) items from the memoryless property of the exponential and $E(X_2 - X_1) = \frac{1}{(n-1)\lambda}$.

Similarly, $E(X_k - X_{k-1}) = \frac{1}{(n-k+1)\lambda}$. Then,

$E(T_r) = \frac{n}{n\lambda} + \frac{n-1}{(n-1)\lambda} + \ldots + \frac{n-r+1}{(n-r+1)\lambda} = \frac{r}{\lambda}$ and $E\left(\frac{T_r}{r}\right) = \frac{1}{\lambda} = \mu$

b.) $V(T_r / r) = 1/(\lambda^2 r)$ is related to the variance of the Erlang distribution $V(X) = r/\lambda^2$. They are related by the value $(1/r^2)$. The censored variance is $(1/r^2)$ times the uncensored variance.

Section 7.3.3 on CD

S7-1 From Example S7-2 the posterior distribution for μ is normal with mean $\dfrac{(\sigma^2/n)\mu_0 + \sigma_0^2\bar{x}}{\sigma_0^2 + \sigma^2/n}$ and

variance $\dfrac{\sigma_0^2/(\sigma^2/n)}{\sigma_0^2 + \sigma^2/n}$.

The Bayes estimator for μ goes to the MLE as n increases. This follows since σ^2/n goes to 0, and the estimator approaches $\dfrac{\sigma_0^2\bar{x}}{\sigma_0^2}$ (the σ_0^2 's cancel). Thus, in the limit $\hat{\mu} = \bar{x}$.

S7-2 a) Because $f(x|\mu) = \dfrac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ and $f(\mu) = \dfrac{1}{b-a}$ for $a \le \mu \le b$, the joint distribution is $f(x,\mu) = \dfrac{1}{(b-a)\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ for $-\infty < x < \infty$ and $a \le \mu \le b$. Then,

$f(x) = \dfrac{1}{b-a}\displaystyle\int_a^b \dfrac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} d\mu$ and this integral is recognized as a normal probability. Therefore,

$f(x) = \dfrac{1}{b-a}\left[\Phi\left(\frac{b-x}{\sigma}\right) - \Phi\left(\frac{a-x}{\sigma}\right)\right]$ where Φ(x) is the standard normal cumulative distribution function.

Then, $f(\mu|x) = \dfrac{f(x,\mu)}{f(x)} = \dfrac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma\left[\Phi\left(\frac{b-x}{\sigma}\right) - \Phi\left(\frac{a-x}{\sigma}\right)\right]}$

b) The Bayes estimator is $\tilde{\mu} = \displaystyle\int_a^b \dfrac{\mu e^{-\frac{(x-\mu)^2}{2\sigma^2}} d\mu}{\sqrt{2\pi}\sigma\left[\Phi\left(\frac{b-x}{\sigma}\right) - \Phi\left(\frac{a-x}{\sigma}\right)\right]}$.

Let v = (x - μ). Then, dv = - dμ and

$$\tilde{\mu} = \int_{x-b}^{x-a} \frac{(x-v)e^{-\frac{v^2}{2\sigma^2}} dv}{\sqrt{2\pi}\sigma\left[\Phi\left(\frac{b-x}{\sigma}\right) - \Phi\left(\frac{a-x}{\sigma}\right)\right]}$$

$$= \frac{x\left[\Phi\left(\frac{x-a}{\sigma}\right) - \Phi\left(\frac{x-b}{\sigma}\right)\right]}{\left[\Phi\left(\frac{b-x}{\sigma}\right) - \Phi\left(\frac{a-x}{\sigma}\right)\right]} - \int_{x-b}^{x-a} \frac{ve^{-\frac{v^2}{2\sigma^2}} dv}{\sqrt{2\pi}\sigma\left[\Phi\left(\frac{b-x}{\sigma}\right) - \Phi\left(\frac{a-x}{\sigma}\right)\right]}$$

Let w = $\dfrac{v^2}{2\sigma^2}$. Then, $dw = [\frac{2v}{2\sigma^2}]dv = [\frac{v}{\sigma^2}]dv$ and

$$\tilde{\mu} = x - \int_{\frac{(x-b)^2}{2\sigma^2}}^{\frac{(x-a)^2}{2\sigma^2}} \frac{\sigma e^{-w} dw}{\sqrt{2\pi}\left[\Phi\left(\frac{b-x}{\sigma}\right) - \Phi\left(\frac{a-x}{\sigma}\right)\right]}$$

$$= x + \frac{\sigma}{\sqrt{2\pi}}\left[\frac{e^{-\frac{(x-a)^2}{2\sigma^2}} - e^{-\frac{(x-b)^2}{2\sigma^2}}}{\Phi\left(\frac{b-x}{\sigma}\right) - \Phi\left(\frac{a-x}{\sigma}\right)}\right]$$

S7-3. a) $f(x) = \frac{e^{-\lambda}\lambda^x}{x!}$ for x = 0, 1, 2, and $f(\lambda) = \left(\frac{m+1}{\lambda_0}\right)^{m+1} \frac{\lambda^m e^{-(m+1)\frac{\lambda}{\lambda_0}}}{\Gamma(m+1)}$ for λ > 0. Then,

$$f(x,\lambda) = \frac{(m+1)^{m+1}\lambda^{m+x} e^{-\lambda-(m+1)\frac{\lambda}{\lambda_0}}}{\lambda_0^{m+1}\Gamma(m+1)x!}.$$

This last density is recognized to be a gamma density as a function of λ. Therefore, the posterior distribution of λ is a gamma distribution with parameters m + x + 1 and 1 + $\frac{m+1}{\lambda_0}$.

b) The mean of the posterior distribution can be obtained from the results for the gamma distribution to be

$$\frac{m+x+1}{\left[1+\frac{m+1}{\lambda_0}\right]} = \lambda_0\left(\frac{m+x+1}{m+\lambda_0+1}\right)$$

S7-4 a) From Example S7-2, the Bayes estimate is $\tilde{\mu} = \frac{\frac{9}{25}(4)+1(4.85)}{\frac{9}{25}+1} = 4.625$

b.) $\hat{\mu} = \bar{x} = 4.85$ The Bayes estimate appears to underestimate the mean.

S7-5. a) From Example S7-2, $\tilde{\mu} = \frac{(0.01)(5.03)+(\frac{1}{25})(5.05)}{0.01+\frac{1}{25}} = 5.046$

b.) $\hat{\mu} = \bar{x} = 5.05$ The Bayes estimate is very close to the MLE of the mean.

S7-6. a) $f(x \mid \lambda) = \lambda e^{-\lambda x},\ x \ge 0$ and $f(\lambda) = 0.01e^{-0.01\lambda}$. Then,

$f(x_1, x_2, \lambda) = \lambda^2 e^{-\lambda(x_1+x_2)} 0.01e^{-0.01\lambda} = 0.01\lambda^2 e^{-\lambda(x_1+x_2+0.01)}$. As a function of λ, this is recognized as a gamma density with parameters 3 and $x_1 + x_2 + 0.01$. Therefore, the posterior mean for λ is

$$\tilde{\lambda} = \frac{3}{x_1+x_2+0.01} = \frac{3}{2\bar{x}+0.01} = 0.00133.$$

b) Using the Bayes estimate for λ, P(X<1000)= $\int_0^{1000} 0.00133e^{-.00133x}dx$ =0.736.

CHAPTER 8

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 Φ(2.14) = P(Z<2.14) = 0.9838 and the confidence level is 100(1-.032354) = 96.76%.

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 Φ(2.49) = P(Z<2.49) = 0.9936 and the confidence level is is 100(1-.012774) = 98.72%.

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 Φ(1.85) = P(Z<1.85) = 0.9678 and the confidence level is 93.56%.

8-2 a.) A z_α = 2.33 would give result in a 98% two-sided confidence interval.
b.) A z_α = 1.29 would give result in a 80% two-sided confidence interval.
c.) A z_α = 1.15 would give result in a 75% two-sided confidence interval.

8-3 a.) A z_α = 1.29 would give result in a 90% one-sided confidence interval.
b.) A z_α = 1.65 would give result in a 95% one-sided confidence interval.
c.) A z_α = 2.33 would give result in a 99% one-sided confidence interval.

8-4 a.) 95% CI for μ, $n = 10, \quad \sigma = 20 \quad \bar{x} = 1000, \quad z = 1.96$

$$\bar{x} - z\sigma/\sqrt{n} \le \mu \le \bar{x} + z\sigma/\sqrt{n}$$

$$1000 - 1.96(20/\sqrt{10}) \le \mu \le 1000 + 1.96(20/\sqrt{10})$$

$$987.6 \le \mu \le 1012.4$$

b.) .95% CI for μ, $n = 25, \quad \sigma = 20 \quad \bar{x} = 1000, \quad z = 1.96$

$$\bar{x} - z\sigma/\sqrt{n} \le \mu \le \bar{x} + z\sigma/\sqrt{n}$$

$$1000 - 1.96(20/\sqrt{25}) \le \mu \le 1000 + 1.96(20/\sqrt{25})$$

$$992.2 \le \mu \le 1007.8$$

c.) 99% CI for μ, $n = 10, \quad \sigma = 20 \quad \bar{x} = 1000, \quad z = 2.58$

$$\bar{x} - z\sigma/\sqrt{n} \le \mu \le \bar{x} + z\sigma/\sqrt{n}$$

$$1000 - 2.58(20/\sqrt{10}) \le \mu \le 1000 + 2.58(20/\sqrt{10})$$

$$983.7 \le \mu \le 1016.3$$

d.) 99% CI for μ, $n = 25, \quad \sigma = 20 \quad \bar{x} = 1000, \quad z = 2.58$

$$\bar{x} - z\sigma/\sqrt{n} \le \mu \le \bar{x} + z\sigma/\sqrt{n}$$

$$1000 - 2.58(20/\sqrt{25}) \le \mu \le 1000 + 2.58(20/\sqrt{25})$$

$$989.7 \le \mu \le 1010.3$$

8-5 Find n for the length of the 95% CI to be 40. $Z_{a/2} = 1.96$

$$1/2 \text{ length} = (1.96)(20)/\sqrt{n} = 20$$

$$39.2 = 20\sqrt{n}$$

$$n = \left(\frac{39.2}{20}\right)^2 = 3.84$$

Therefore, $n = 4$.

8-6 Interval (1): $3124.9 \le \mu \le 3215.7$ and Interval (2):: $3110.5 \le \mu \le 3230.1$

Interval (1): half-length =90.8/2=45.4 Interval (2): half-length =119.6/2=59.8

a.) $\bar{x}_1 = 3124.9 + 45.4 = 3170.3$

$\bar{x}_2 = 3110.5 + 59.8 = 3170.3$ The sample means are the same.

b.) Interval (1): $3124.9 \le \mu \le 3215.7$ was calculated with 95% Confidence because it has a smaller half-length, and therefore a smaller confidence interval. The 99% confidence level will make the interval larger.

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.

8-8 95% Two-sided CI on the breaking strength of yarn: where $\bar{x} = 98$, $\sigma = 2$, n=9 and $z_{0.025} = 1.96$

$$\bar{x} - z_{0.025}\sigma/\sqrt{n} \le \mu \le \bar{x} + z_{0.025}\sigma/\sqrt{n}$$

$$98 - 1.96(2)/\sqrt{9} \le \mu \le 98 + 1.96(2)/\sqrt{9}$$

$$96.7 \le \mu \le 99.3$$

8-9 95% Two-sided CI on the true mean yield: where $\bar{x} = 90.480$, $\sigma = 3$, n=5 and $z_{0.025} = 1.96$

$$\bar{x} - z_{0.025}\sigma/\sqrt{n} \le \mu \le \bar{x} + z_{0.025}\sigma/\sqrt{n}$$

$$90.480 - 1.96(3)/\sqrt{5} \le \mu \le 90.480 + 1.96(3)/\sqrt{5}$$

$$87.85 \le \mu \le 93.11$$

8-10 99% Two-sided CI on the diameter cable harness holes: where $\bar{x}$ =1.5045 , $\sigma = 0.01$, n=10 and $z_{0.005} = 2.58$

$$\bar{x} - z_{0.005}\sigma/\sqrt{n} \le \mu \le \bar{x} + z_{0.005}\sigma/\sqrt{n}$$

$$1.5045 - 2.58(0.01)/\sqrt{10} \le \mu \le 1.5045 + 2.58(0.01)/\sqrt{10}$$

$$1.4963 \le \mu \le 1.5127$$

8-11 a.) 99% Two-sided CI on the true mean piston ring diameter

For $\alpha = 0.01$, $z_{\alpha/2} = z_{0.005} = 2.58$, and $\bar{x} = 74.036$, $\sigma = 0.001$, n=15

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

$$74.036 - 2.58\left(\frac{0.001}{\sqrt{15}}\right) \le \mu \le 74.036 + 2.58\left(\frac{0.001}{\sqrt{15}}\right)$$

$$74.0353 \le \mu \le 74.0367$$

b.) 95% One-sided CI on the true mean piston ring diameter

For $\alpha = 0.05$, $z_{\alpha} = z_{0.05} = 1.65$ and $\bar{x} = 74.036$, $\sigma = 0.001$, n=15

$$\bar{x} - z_{0.05}\frac{\sigma}{\sqrt{n}} \le \mu$$

$$74.036 - 1.65\left(\frac{0.001}{\sqrt{15}}\right) \le \mu$$

$$74.0356 \le \mu$$

8-12 a.) 95% Two-sided CI on the true mean life of a 75-watt light bulb

For $\alpha = 0.05$, $z_{\alpha/2} = z_{0.025} = 1.96$, and $\bar{x} = 1014$, $\sigma = 25$, n=20

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

$$1014 - 1.96\left(\frac{25}{\sqrt{20}}\right) \le \mu \le 1014 + 1.96\left(\frac{25}{\sqrt{20}}\right)$$

$$1003 \le \mu \le 1025$$

b.) 95% One-sided CI on the true mean piston ring diameter

For $\alpha = 0.05$, $z_{\alpha} = z_{0.05} = 1.65$ and $\bar{x} = 1014$, $\sigma = 25$, n=20

$$\bar{x} - z_{0.05}\frac{\sigma}{\sqrt{n}} \le \mu$$

$$1014 - 1.65\left(\frac{25}{\sqrt{20}}\right) \le \mu$$

$$1005 \le \mu$$

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

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

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

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

$$3226.4 \le \mu \le 3273.6$$

8-14 95% Confident that the error of estimating the true mean life of a 75-watt light bulb is less than 5 hours.

For $\alpha = 0.05$, $z_{\alpha/2} = z_{0.025} = 1.96$, and $\sigma = 25$, $E=5$

$$n = \left(\frac{z_{a/2}\sigma}{E}\right)^2 = \left(\frac{1.96(25)}{5}\right)^2 = 96.04$$

Always round up to the next number, therefore n=97

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.

8-16 99% Confident that the error of estimating the true compressive strength is less than 15 psi.

For $\alpha = 0.01$, $z_{\alpha/2} = z_{0.005} = 2.58$, and $\sigma = 31.62$, $E=15$

$$n = \left(\frac{z_{a/2}\sigma}{E}\right)^2 = \left(\frac{2.58(31.62)}{15}\right)^2 = 29.6 \cong 30$$

Therefore, n=30

8-17 To decrease the length of the CI by one half, the sample size must be increased by 4 times (2^2).

$$z_{\alpha/2}\sigma/\sqrt{n} = 0.5l$$

Now, to decrease by half, divide both sides by 2.

$$(z_{\alpha/2}\sigma/\sqrt{n})/2 = (l/2)/2$$

$$(z_{\alpha/2}\sigma/2\sqrt{n}) = l/4$$

$$(z_{\alpha/2}\sigma/\sqrt{2^2 n}) = l/4$$

Therefore, the sample size must be increased by 2^2.

8-18 If n is doubled in Eq 8-7: $\bar{x} - z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right) \le \mu \le \bar{x} + z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$

$$\frac{z_{\alpha/2}\sigma}{\sqrt{2n}} = \frac{z_{\alpha/2}\sigma}{1.414\sqrt{n}} = \frac{z_{\alpha/2}\sigma}{1.414\sqrt{n}} = \frac{1}{1.414}\left(\frac{z_{\alpha/2}\sigma}{\sqrt{n}}\right)$$

The interval is reduced by 0.293 29.3%

If n is increased by a factor of 4 Eq 8-7:

$$\frac{z_{\alpha/2}\sigma}{\sqrt{4n}} = \frac{z_{\alpha/2}\sigma}{2\sqrt{n}} = \frac{z_{\alpha/2}\sigma}{2\sqrt{n}} = \frac{1}{2}\left(\frac{z_{\alpha/2}\sigma}{\sqrt{n}}\right)$$

The interval is reduced by 0.5 or ½.

Section 8-3

8-19 $t_{0.025,15} = 2.131$ $t_{0.05,10} = 1.812$ $t_{0.10,20} = 1.325$

$t_{0.005,25} = 2.787$ $t_{0.001,30} = 3.385$

8-20 a.) $t_{0.025,12} = 2.179$

b.) $t_{0.025,24} = 2.064$

c.) $t_{0.005,13} = 3.012$

d.) $t_{0.0005,15} = 4.073$

8-21 a.) $t_{0.05,14} = 1.761$

b.) $t_{0.01,19} = 2.539$

c.) $t_{0.001,24} = 3.467$

8-22 95% confidence interval on mean tire life

$$n = 16 \quad \bar{x} = 60{,}139.7 \quad s = 3645.94 \quad t_{0.025,15} = 2.131$$

$$\bar{x} - t_{0.025,15}\left(\frac{s}{\sqrt{n}}\right) \le \mu \le \bar{x} + t_{0.025,15}\left(\frac{s}{\sqrt{n}}\right)$$

$$60139.7 - 2.131\left(\frac{3645.94}{\sqrt{16}}\right) \le \mu \le 60139.7 + 2.131\left(\frac{3645.94}{\sqrt{16}}\right)$$

$$58197.33 \le \mu \le 62082.07$$

8-23 99% lower confidence bound on mean Izod impact strength

$$n = 20 \quad \bar{x} = 1.25 \quad s = 0.25 \quad t_{0.01,19} = 2.539$$

$$\bar{x} - t_{0.01,19}\left(\frac{s}{\sqrt{n}}\right) \le \mu$$

$$1.25 - 2.539\left(\frac{0.25}{\sqrt{20}}\right) \le \mu$$

$$1.108 \le \mu$$

8-24 99% confidence interval on mean current required
Assume that the data are normally distributed and that the variance is unknown.

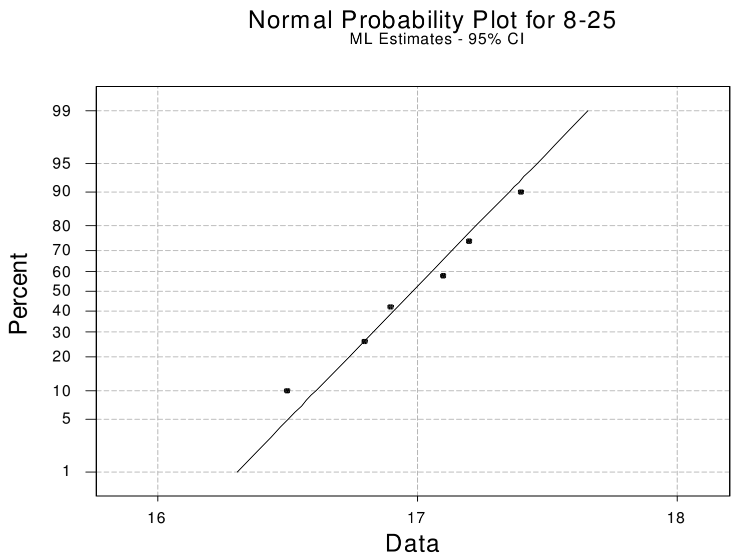
$$n = 10 \quad \bar{x} = 317.2 \quad s = 15.7 \quad t_{0.005,9} = 3.250$$

$$\bar{x} - t_{0.005,9}\left(\frac{s}{\sqrt{n}}\right) \le \mu \le \bar{x} + t_{0.005,9}\left(\frac{s}{\sqrt{n}}\right)$$

$$317.2 - 3.250\left(\frac{15.7}{\sqrt{10}}\right) \le \mu \le 317.2 + 3.250\left(\frac{15.7}{\sqrt{10}}\right)$$

$$301.06 \le \mu \le 333.34$$

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 $\alpha = 0.01$, $t_{\alpha/2,n-1} = t_{0.005,5} = 4.032$

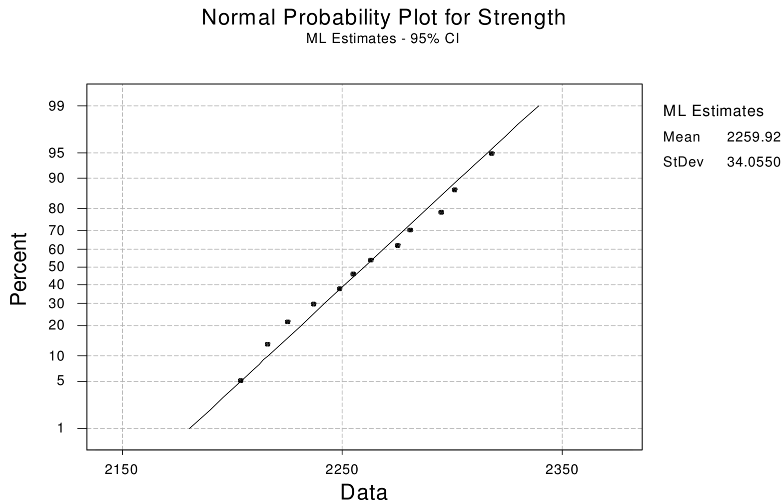
$$\bar{x} - t_{0.005,5}\left(\frac{s}{\sqrt{n}}\right) \le \mu \le \bar{x} + t_{0.005,5}\left(\frac{s}{\sqrt{n}}\right)$$

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

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

The 99% confidence for the mean polyunsaturated fat is (16.455, 17.505). We would look for the true mean to be somewhere in this region.

8-26 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 compressive strength is normally distributed.



b.) 95% two-sided confidence interval on mean comprehensive strength

$n = 12 \quad \bar{x} = 2259.9 \quad s = 35.6 \quad t_{0.025,11} = 2.201$

$$\bar{x} - t_{0.025,11}\left(\frac{s}{\sqrt{n}}\right) \le \mu \le \bar{x} + t_{0.025,11}\left(\frac{s}{\sqrt{n}}\right)$$

$$2259.9 - 2.201\left(\frac{35.6}{\sqrt{12}}\right) \le \mu \le 2259.9 + 2.201\left(\frac{35.6}{\sqrt{12}}\right)$$

$$2237.3 \le \mu \le 2282.5$$

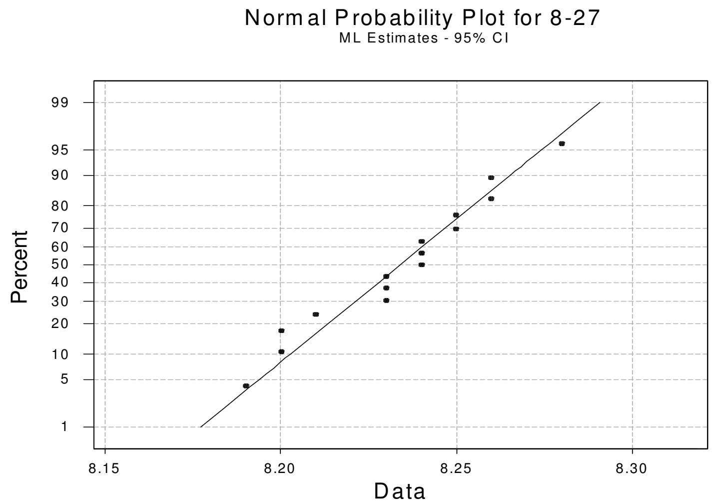
c.) 95% lower-confidence bound on mean strength

$$\bar{x} - t_{0.05,11}\left(\frac{s}{\sqrt{n}}\right) \le \mu$$

$$2259.9 - 1.796\left(\frac{35.6}{\sqrt{12}}\right) \le \mu$$

$$2241.4 \le \mu$$

8-27 a.) According to the normal probability plot there does not seem to be a severe deviation from normality for this data. This is evident by the fact that the data appears to fall along a straight line.



b.) 95% two-sided confidence interval on mean rod diameter
For $\alpha = 0.05$ and $n = 15$, $t_{\alpha/2,n-1} = t_{0.025,14} = 2.145$

$$\bar{x} - t_{0.025,14}\left(\frac{s}{\sqrt{n}}\right) \le \mu \le \bar{x} + t_{0.025,14}\left(\frac{s}{\sqrt{n}}\right)$$

$$8.23 - 2.145\left(\frac{0.025}{\sqrt{15}}\right) \le \mu \le 8.23 + 2.145\left(\frac{0.025}{\sqrt{15}}\right)$$

$$8.216 \le \mu \le 8.244$$

8-28 95% lower confidence bound on mean rod diameter $t_{0.05,14} = 1.761$

$$\bar{x} - t_{0.05,14}\left(\frac{s}{\sqrt{n}}\right) \le \mu$$

$$8.23 - 1.761\left(\frac{0.025}{\sqrt{15}}\right) \le \mu$$

$$8.219 \le \mu$$

The lower bound of the one sided confidence interval is lower than the lower bound of the two-sided confidence interval even though the level of significance is the same. This is because all of the alpha value is in the left tail (or in the lower bound).

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

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

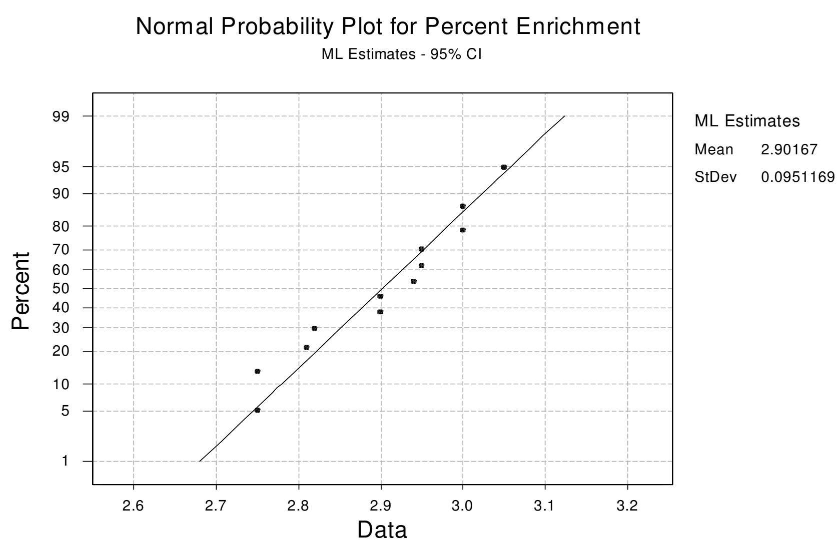
$$\overline{x} - t_{0.05,24}\left(\frac{s}{\sqrt{n}}\right) \le \mu$$

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

$$4.023 \le \mu$$

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

8-30 a.) The data appear to be normally distributed. Therefore, there is no evidence to support that the percentage of enrichment is not normally distributed.



b.) 99% two-sided confidence interval on mean percentage enrichment
For $\alpha = 0.01$ and $n = 12$, $t_{\alpha/2,n-1} = t_{0.005,12} = 3.106$, $\overline{x} = 2.9017$ $s = 0.0993$

$$\overline{x} - t_{0.005,11}\left(\frac{s}{\sqrt{n}}\right) \le \mu \le \overline{x} + t_{0.005,11}\left(\frac{s}{\sqrt{n}}\right)$$

$$2.902 - 3.106\left(\frac{0.0993}{\sqrt{12}}\right) \le \mu \le 2.902 + 3.106\left(\frac{0.0993}{\sqrt{12}}\right)$$

$$2.813 \le \mu \le 2.991$$

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

95% CI on the mean volume of syrup dispensed

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

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

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

$$1.094 \le \mu \le 1.106$$

8-32 90% CI on the mean frequency of a beam subjected to loads

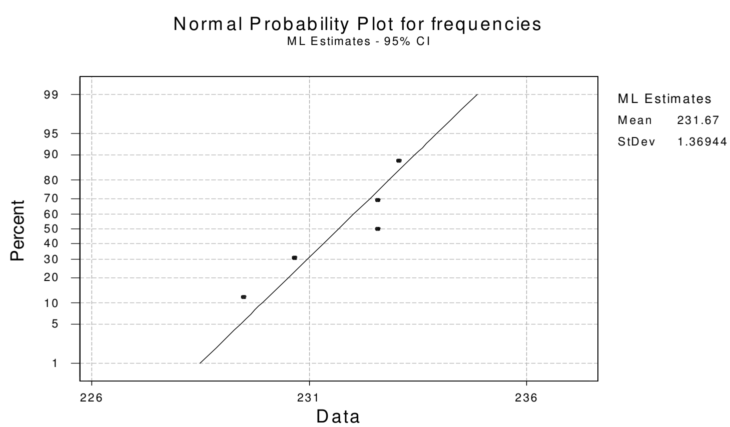
$\bar{x} = 231.67, \quad s = 1.53, \quad n = 5, \quad t_{\alpha/2,n-1} = t_{.05,4} = 2.132$

$$\bar{x} - t_{0.05,4}\left(\frac{s}{\sqrt{n}}\right) \le \mu \le \bar{x} + t_{0.05,4}\left(\frac{s}{\sqrt{n}}\right)$$

$$231.67 - 2.132\left(\frac{1.53}{\sqrt{5}}\right) \le \mu \le 231.67 - 2.132\left(\frac{1.53}{\sqrt{5}}\right)$$

$$230.2 \le \mu \le 233.1$$

By examining the normal probability plot, it appears that the data are normally distributed. There does not appear to be enough evidence to reject the hypothesis that the frequencies are normally distributed.



Section 8-4

8-33 $\chi^2_{0.05,10} = 18.31$ $\chi^2_{0.025,15} = 27.49$ $\chi^2_{0.01,12} = 26.22$

$\chi^2_{0.005,25} = 46.93$ $\chi^2_{0.95,20} = 10.85$ $\chi^2_{0.99,18} = 7.01$ $\chi^2_{0.995,16} = 5.14$

8-34 a.) $\chi^2_{0.05,24} = 36.42$

b.) $\chi^2_{0.99,9} = 2.09$

c.) $\chi^2_{0.95,19} = 10.12$ and $\chi^2_{0.05,19} = 30.14$

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

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

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

$$0.00003075 \le \sigma^2$$

8-36 95% two sided confidence interval for σ

$n = 10 \quad s = 4.8$

$\chi^2_{\alpha/2,n-1} = \chi^2_{0.025,9} = 19.02$ and $\chi^2_{1-\alpha/2,n-1} = \chi^2_{0.975,9} = 2.70$

$$\frac{9(4.8)^2}{19.02} \le \sigma^2 \le \frac{9(4.8)^2}{2.70}$$

$$10.90 \le \sigma^2 \le 76.80$$

$$3.30 < \sigma < 8.76$$

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

For $\alpha = 0.05$ and n = 16, $\chi^2_{\alpha,n-1} = \chi^2_{0.05,15} = 25.00$

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

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

8-38 99% two-sided confidence interval on σ^2 for Izod impact test data

$n = 20 \quad s = 0.25 \quad \chi^2_{0.005,19} = 38.58$ and $\chi^2_{0.995,19} = 6.84$

$$\frac{19(0.25)^2}{38.58} \le \sigma^2 \le \frac{19(0.25)^2}{6.84}$$

$$0.03078 \le \sigma^2 \le 0.1736$$

$$0.1754 < \sigma < 0.4167$$

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 $\chi^2_{1-\alpha/2,n-1} = \chi^2_{0.975,50} = 32.36$

$$\frac{50(0.37)^2}{71.42} \le \sigma^2 \le \frac{50(0.37)^2}{32.36}$$

$$0.096 \le \sigma^2 \le 0.2115$$

Taking the square root of the endpoints of this interval we obtain,

$$0.31 < \sigma < 0.46$$

8-40 99% two-sided confidence interval for σ for the hole diameter data. (Exercise 8-35)

For $\alpha = 0.01$ and n = 15, $\chi^2_{\alpha/2,n-1} = \chi^2_{0.005,14} = 31.32$ and $\chi^2_{1-\alpha/2,n-1} = \chi^2_{0.995,14} = 4.07$

$$\frac{14(0.008)^2}{31.32} \le \sigma^2 \le \frac{14(0.008)^2}{4.07}$$

$$0.00002861 \le \sigma^2 \le 0.0002201$$

$$0.005349 < \sigma < 0.01484$$

8-41 90% lower confidence bound on σ (the standard deviation the sugar content)

given n = 10, $s^2 = 23.04$

For $\alpha = 0.1$ and n = 10, $\chi^2_{\alpha,n-1} = \chi^2_{0.1,9} = 14.68$

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

$$14.13 \le \sigma^2$$

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

$$3.8 < \sigma$$

<u>Section 8-5</u>

8-42 95% Confidence Interval on the death rate from lung cancer.

$$\hat{p} = \frac{823}{1000} = 0.823 \quad n = 1000 \quad z_{\alpha/2} = 1.96$$

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

$$0.823 - 1.96\sqrt{\frac{0.823(0.177)}{1000}} \le p \le 0.823 + 1.96\sqrt{\frac{0.823(0.177)}{1000}}$$

$$0.7993 \le p \le 0.8467$$

8-43 E = 0.03, $\alpha = 0.05$, $z_{\alpha/2} = z_{0.025} = 1.96$ and $\hat{p} = 0.823$ as the initial estimate of p,

$$n = \left(\frac{z_{\alpha/2}}{E}\right)^2 \hat{p}(1-\hat{p}) = \left(\frac{1.96}{0.03}\right)^2 0.823(1-0.823) = 621.79,$$

$n \cong 622$.

8-44 a.) 95% Confidence Interval on the true proportion of helmets showing damage.

$$\hat{p} = \frac{18}{50} = 0.36 \qquad n = 50 \qquad z_{\alpha/2} = 1.96$$

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

$$0.36 - 1.96\sqrt{\frac{0.36(0.64)}{50}} \le p \le 0.36 + 1.96\sqrt{\frac{0.36(0.64)}{50}}$$

$$0.227 \le p \le 0.493$$

b.) $n = \left(\frac{z_{\alpha/2}}{E}\right)^2 p(1-p) = \left(\frac{1.96}{0.02}\right)^2 0.36(1-0.36) = 2212.76$

$n \cong 2213$

c.) $n = \left(\frac{z_{\alpha/2}}{E}\right)^2 p(1-p) = \left(\frac{1.96}{0.02}\right)^2 0.5(1-0.5) = 2401$

8-45 The worst case would be for p = 0.5, thus with E = 0.05 and $\alpha = 0.01$, $z_{\alpha/2} = z_{0.005} = 2.58$ we obtain a sample size of:

$$n = \left(\frac{z_{\alpha/2}}{E}\right)^2 p(1-p) = \left(\frac{2.58}{0.05}\right)^2 0.5(1-0.5) = 665.64, \quad n \cong 666$$

8-46 99% Confidence Interval on the fraction defective.

$$\hat{p} = \frac{10}{800} = 0.0125 \qquad n = 800 \qquad z_{\alpha} = 2.33$$

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

$$\infty \le p \le 0.0125 + 2.33\sqrt{\frac{0.0125(0.9875)}{800}}$$

$$\infty \le p \le 0.0217$$

8-47 E = 0.017, $\alpha = 0.01$, $z_{\alpha/2} = z_{0.005} = 2.58$

$$n = \left(\frac{z_{\alpha/2}}{E}\right)^2 p(1-p) = \left(\frac{2.58}{0.017}\right)^2 0.5(1-0.5) = 5758.13, \quad n \cong 5759$$

8-48 95% Confidence Interval on the fraction defective produced with this tool.

$$\hat{p} = \frac{13}{300} = 0.04333 \quad n = 300 \quad z_{\alpha/2} = 1.96$$

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

$$0.04333 - 1.96\sqrt{\frac{0.04333(0.95667)}{300}} \le p \le 0.04333 + 1.96\sqrt{\frac{0.04333(0.95667)}{300}}$$

$$0.02029 \le p \le 0.06637$$

Section 8-6

8-49 95% prediction interval on the life of the next tire
given $\bar{x} = 60139.7$ $s = 3645.94$ $n = 16$
for $\alpha=0.05$ $t_{\alpha/2,n-1} = t_{0.025,15} = 2.131$

$$\bar{x} - t_{0.025,15}s\sqrt{1+\frac{1}{n}} \le x_{n+1} \le \bar{x} + t_{0.025,15}s\sqrt{1+\frac{1}{n}}$$

$$60139.7 - 2.131(3645.94)\sqrt{1+\frac{1}{16}} \le x_{n+1} \le 60139.7 + 2.131(3645.94)\sqrt{1+\frac{1}{16}}$$

$$52131.1 \le x_{n+1} \le 68148.3$$

The prediction interval is considerably wider than the 95% confidence interval ($58,197.3 \le \mu \le 62,082.07$), which is to be expected since we are going beyond our data.

8-50 99% prediction interval on the Izod impact data

$$n = 20 \quad \bar{x} = 1.25 \quad s = 0.25 \quad t_{0.005,19} = 2.861$$

$$\bar{x} - t_{0.005,19}s\sqrt{1+\frac{1}{n}} \le x_{n+1} \le \bar{x} + t_{0.005,19}s\sqrt{1+\frac{1}{n}}$$

$$1.25 - 2.861(0.25)\sqrt{1+\frac{1}{20}} \le x_{n+1} \le 1.25 + 2.861(0.25)\sqrt{1+\frac{1}{20}}$$

$$0.517 \le x_{n+1} \le 1.983$$

The lower bound of the 99% prediction interval is considerably lower than the 99% confidence interval ($1.108 \le \mu \le \infty$), which is to be expected since we are going beyond our data.

8-51 Given $\bar{x} = 317.2$ $s = 15.7$ $n = 10$ for $\alpha=0.05$ $t_{\alpha/2,n-1} = t_{0.005,9} = 3.250$

$$\bar{x} - t_{0.005,9} s\sqrt{1+\frac{1}{n}} \le x_{n+1} \le \bar{x} + t_{0.005,9} s\sqrt{1+\frac{1}{n}}$$

$$317.2 - 3.250(15.7)\sqrt{1+\frac{1}{10}} \le x_{n+1} \le 317.2 - 3.250(15.7)\sqrt{1+\frac{1}{10}}$$

$$263.7 \le x_{n+1} \le 370.7$$

The length of the prediction interval is longer.

8-52 99% prediction interval on the polyunsaturated fat

$n = 6 \quad \bar{x} = 16.98 \quad s = 0.319 \quad t_{0.005,5} = 4.032$

$$\bar{x} - t_{0.005,5} s\sqrt{1+\frac{1}{n}} \le x_{n+1} \le \bar{x} + t_{0.005,5} s\sqrt{1+\frac{1}{n}}$$

$$16.98 - 4.032(0.319)\sqrt{1+\frac{1}{6}} \le x_{n+1} \le 16.98 + 4.032(0.319)\sqrt{1+\frac{1}{6}}$$

$$15.59 \le x_{n+1} \le 18.37$$

The length of the prediction interval is a lot longer than the width of the confidence interval $16.455 \le \mu \le 17.505$.

8-53 90% prediction interval on the next specimen of concrete tested

given $\bar{x} = 2260$ $s = 35.57$ $n = 12$ for $\alpha = 0.05$ and n = 12, $t_{\alpha/2,n-1} = t_{0.05,11} = 1.796$

$$\bar{x} - t_{0.05,11} s\sqrt{1+\frac{1}{n}} \le x_{n+1} \le \bar{x} + t_{0.05,11} s\sqrt{1+\frac{1}{n}}$$

$$2260 - 1.796(35.57)\sqrt{1+\frac{1}{12}} \le x_{n+1} \le 2260 + 1.796(35.57)\sqrt{1+\frac{1}{12}}$$

$$2193.5 \le x_{n+1} \le 2326.5$$

8-54 95% prediction interval on the next rod diameter tested

$n = 15 \quad \bar{x} = 8.23 \quad s = 0.025 \quad t_{0.025,14} = 2.145$

$$\bar{x} - t_{0.025,14} s\sqrt{1+\frac{1}{n}} \le x_{n+1} \le \bar{x} + t_{0.025,14} s\sqrt{1+\frac{1}{n}}$$

$$8.23 - 2.145(0.025)\sqrt{1+\frac{1}{15}} \le x_{n+1} \le 8.23 - 2.145(0.025)\sqrt{1+\frac{1}{15}}$$

$$8.17 \le x_{n+1} \le 8.29$$

95% two-sided confidence interval on mean rod diameter is $8.216 \le \mu \le 8.244$

8-55 90% prediction interval on wall thickness on the next bottle tested.

given $\bar{x} = 4.05$ s = 0.08 n = 25 for $t_{\alpha/2,n-1} = t_{0.05,24} = 1.711$

$$\bar{x} - t_{0.05,24}s\sqrt{1+\frac{1}{n}} \le x_{n+1} \le \bar{x} + t_{0.05,24}s\sqrt{1+\frac{1}{n}}$$

$$4.05 - 1.711(0.08)\sqrt{1+\frac{1}{25}} \le x_{n+1} \le 4.05 - 1.711(0.08)\sqrt{1+\frac{1}{25}}$$

$$3.91 \le x_{n+1} \le 4.19$$

8-56 To obtain a one sided prediction interval, use $t_{\alpha,n-1}$ instead of $t_{\alpha/2,n-1}$
Since we want a 95% one sided prediction interval, $t_{\alpha/2,n-1} = t_{0.05,24} = 1.711$
and $\bar{x} = 4.05$ s = 0.08 n = 25

$$\bar{x} - t_{0.05,24}s\sqrt{1+\frac{1}{n}} \le x_{n+1}$$

$$4.05 - 1.711(0.08)\sqrt{1+\frac{1}{25}} \le x_{n+1}$$

$$3.91 \le x_{n+1}$$

The prediction interval bound is a lot lower than the confidence interval bound of 4.023 mm

8-57 99% prediction interval for enrichment data given $\bar{x} = 2.9$ s = 0.099 n = 12 for $\alpha = 0.01$ and n = 12, $t_{\alpha/2,n-1} = t_{0.005,11} = 3.106$

$$\bar{x} - t_{0.005,12}s\sqrt{1+\frac{1}{n}} \le x_{n+1} \le \bar{x} + t_{0.005,12}s\sqrt{1+\frac{1}{n}}$$

$$2.9 - 3.106(0.099)\sqrt{1+\frac{1}{12}} \le x_{n+1} \le 2.9 - 3.106(0.099)\sqrt{1+\frac{1}{12}}$$

$$2.58 \le x_{n+1} \le 3.22$$

The prediction interval is much wider than the 99% CI on the population mean ($2.813 \le \mu \le 2.991$).

8-58 95% Prediction Interval on the volume of syrup of the next beverage dispensed
$\bar{x} = 1.10$ s = 0.015 n = 25 $t_{\alpha/2,n-1} = t_{0.025,24} = 2.064$

$$\bar{x} - t_{0.025,24}s\sqrt{1+\frac{1}{n}} \le x_{n+1} \le \bar{x} + t_{0.025,24}s\sqrt{1+\frac{1}{n}}$$

$$1.10 - 2.064(0.015)\sqrt{1+\frac{1}{25}} \le x_{n+1} \le 1.10 - 2.064(0.015)\sqrt{1+\frac{1}{25}}$$

$$1.068 \le x_{n+1} \le 1.13$$

The prediction interval is wider than the confidence interval: $1.093 \le \mu \le 1.106$

8-59 90% prediction interval the value of the natural frequency of the next beam of this type that will be tested. given $\bar{x}$ = 231.67, s=1.53 For α = 0.10 and n = 5, $t_{\alpha/2,n-1} = t_{0.05,4} = 2.132$

$$\bar{x} - t_{0.05,4} s\sqrt{1+\frac{1}{n}} \le x_{n+1} \le \bar{x} + t_{0.05,4} s\sqrt{1+\frac{1}{n}}$$

$$231.67 - 2.132(1.53)\sqrt{1+\frac{1}{5}} \le x_{n+1} \le 231.67 - 2.132(1.53)\sqrt{1+\frac{1}{5}}$$

$$228.1 \le x_{n+1} \le 235.2$$

The 90% prediction in interval is greater than the 90% CI.

Section 8-7

8-60 95% tolerance interval on the life of the tires that has a 95% CL
given $\bar{x}$ = 60139.7 s = 3645.94 n = 16 we find *k*=2.903

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

$$60139.7 - 2.903(3645.94),\ 60139.7 + 2.903(3645.94)$$

$$(49555.54,\ 70723.86)$$

95% confidence interval (58,197.3 ≤ μ ≤ 62,082.07) is shorter than the 95%tolerance interval.

8-61 99% tolerance interval on the Izod impact strength PVC pipe that has a 90% CL
given $\bar{x}$=1.25, s=0.25 and n=20 we find *k*=3.368

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

$$1.25 - 3.368(0.25),\ 1.25 + 3.368(0.25)$$

$$(0.408,\ 2.092)$$

The 99% tolerance interval is much wider than the 99% confidence interval on the population mean (1.090 ≤ μ ≤ 1.410).

8-62 99% tolerance interval on the brightness of television tubes that has a 95% CL
given $\bar{x}$ = 317.2 s = 15.7 n = 10 we find *k*=4.433

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

$$317.2 - 4.433(15.7),\ 317.2 + 4.433(15.7)$$

$$(247.60,\ 386.80)$$

The 99% tolerance interval is much wider than the 95% confidence interval on the population mean $301.06 \le \mu \le 333.34$.

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 ($16.46 \le \mu \le 17.51$).

8-64 90% tolerance interval on the comprehensive strength of concrete that has a 90% CL given $\bar{x}$ = 2260 s = 35.57 n = 12 we find *k*=2.404

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

$$2260 - 2.404(35.57), 2260 + 2.404(35.57)$$

$$(2174.5, 2345.5)$$

The 90% tolerance interval is much wider than the 95% confidence interval on the population mean $2237.3 \le \mu \le 2282.5$.

8-65 95% tolerance interval on the diameter of the rods in exercise 8-27 that has a 90% confidence level. $\bar{x}$ = 8.23 s = 0.0.25 n=15 and k=2.713

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

$$8.23 - 2.713(0.025), 8.23 + 2.713(0.025)$$

$$(8.16, 8.30)$$

The 95% tolerance interval is wider than the 95% confidence interval on the population mean ($8.216 \le \mu \le 8.244$).

8-66 90% tolerance interval on wall thickness measurements that have a 90% CL given $\bar{x}$ = 4.05 s = 0.08 n = 25 we find *k*=2.077

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

$$4.05 - 2.077(0.08), 4.05 + 2.077(0.08)$$

$$(3.88, 4.22)$$

The lower bound of the 90% tolerance interval is much lower than the lower bound on the 95% confidence interval on the population mean ($4.023 \le \mu \le \infty$)
.

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

$$\bar{x} - ks$$

$$4.05 - 1.702(0.08)$$

$$3.91$$

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.

8-68 99% tolerance interval on rod enrichment data that have a 95% CL given $\overline{x} = 2.9$ $s = 0.099$ $n = 12$ we find k=4.150

$$\bar{x} - ks,\ \bar{x} + ks$$
$$2.9 - 4.150(0.099),\ 2.9 + 4.150(0.099)$$
$$(2.49,\ 3.31)$$

The 99% tolerance interval is much wider than the 95% CI on the population mean ($2.84 \le \mu \le 2.96$).

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

$\overline{x} = 1.10$ $s = 0.015$ $n = 25$ 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

8-70 Where $\alpha_1 + \alpha_2 = \alpha$. Let $\alpha = 0.05$

Interval for $\alpha_1 = \alpha_2 = \alpha/2 = 0.025$

The confidence level for $\bar{x} - 1.96\sigma/\sqrt{n} \le \mu \le \bar{x} + 1.96\sigma/\sqrt{n}$ is determined by the by the value of z_0 which is 1.96. From Table II, we find $\Phi(1.96) = P(Z<1.96) = 0.975$ and the confidence level is 95%.

Interval for $\alpha_1 = 0.01, \quad \alpha_2 = 0.04$

The confidence interval is $\bar{x} - 2.33\sigma/\sqrt{n} \le \mu \le \bar{x} + 1.75\sigma/\sqrt{n}$, the confidence level is the same since $\alpha = 0.05$. The symmetric interval does not affect the level of significance.

8-71 $\mu = 50$ σ unknown

a) $n = 16$ $\overline{x} = 52$ $s = 1.5$

$$t_o = \frac{52 - 50}{8/\sqrt{16}} = 1$$

The P-value for $t_0 = 1$, degrees of freedom = 15, is between 0.1 and 0.25. Thus we would conclude that the results are not very unusual.

b) $n = 30$

$$t_o = \frac{52 - 50}{8/\sqrt{30}} = 1.37$$

The P-value for $t_0 = 1.37$, degrees of freedom = 29, is between 0.05 and 0.1. Thus we would conclude that the results are somewhat unusual.

c) $n = 100$ (with $n > 30$, the standard normal table can be used for this problem)

$$z_o = \frac{52 - 50}{8/\sqrt{100}} = 2.5$$

The P-value for $z_0 = 2.5$, is 0.00621. Thus we would conclude that the results are very unusual.

d) For constant values of $\overline{x}$ and s, increasing only the sample size, we see that the standard error of $\overline{X}$ decreases and consequently a sample mean value of 52 when the true mean is 50 is more unusual for the larger sample sizes.

8-72 $\mu = 50, \quad \sigma^2 = 5$

a.) For $n = 16$ find $P(s^2 \geq 7.44)$ or $P(s^2 \leq 2.56)$

$$P(s^2 \geq 7.44) = P\left(\chi^2_{15} \geq \frac{15(7.44)}{5^2}\right) = 0.05 \leq P\left(\chi^2_{15} \geq 22.32\right) \leq 0.10$$

Using Minitab $P(s^2 \geq 7.44)$ =0.0997

$$P(s^2 \leq 2.56) = P\left(\chi^2_{15} \leq \frac{15(2.56)}{5}\right) = 0.05 \leq P\left(\chi^2_{15} \leq 7.68\right) \leq 0.10$$

Using Minitab $P(s^2 \leq 2.56)$ =0.064

b) For $n = 30$ find $P(s^2 \geq 7.44)$ or $P(s^2 \leq 2.56)$

$$P(s^2 \geq 7.44) = P\left(\chi^2_{29} \geq \frac{29(7.44)}{5}\right) = 0.025 \leq P\left(\chi^2_{29} \geq 43.15\right) \leq 0.05$$

Using Minitab $P(s^2 \geq 7.44)$ =0.044

$$P(s^2 \leq 2.56) = P\left(\chi^2_{29} \leq \frac{29(2.56)}{5}\right) = 0.01 \leq P\left(\chi^2_{29} \leq 14.85\right) \leq 0.025$$

Using Minitab $P(s \leq 2.56)$ =0.014.

c) For $n = 71$ $P(s^2 \geq 7.44)$ or $P(s^2 \leq 2.56)$

$$P(s^2 \geq 7.44) = P\left(\chi^2_{15} \geq \frac{70(7.44)}{5}\right) = 0.005 \leq P\left(\chi^2_{70} \geq 104.16\right) \leq 0.01$$

Using Minitab $P(s^2 \geq 7.44)$ =0.0051

$$P(s^2 \leq 2.56) = P\left(\chi^2_{70} \leq \frac{70(2.56)}{5}\right) = P\left(\chi^2_{70} \leq 35.84\right) \leq 0.005$$

Using Minitab $P(s^2 \leq 2.56)$ <0.001

d) The probabilities get smaller as n increases. As n increases, the sample variance should approach the population variance; therefore, the likelihood of obtaining a sample variance much larger than the population variance will decrease.

e) The probabilities get smaller as n increases. As n increases, the sample variance should approach the population variance; therefore, the likelihood of obtaining a sample variance much smaller than the population variance will decrease.

8-73 a) The data appear to follow a normal distribution based on the normal probability plot since the data fall along a straight line.

b) It is important to check for normality of the distribution underlying the sample data since the confidence intervals to be constructed should have the assumption of normality for the results to be reliable (especially since the sample size is less than 30 and the central limit theorem does not apply).

c) No, with 95% confidence, we can not infer that the true mean could be 14.05 since this value is not contained within the given 95% confidence interval.

d) As with part b, to construct a confidence interval on the variance, the normality assumption must hold for the results to be reliable.

e) Yes, it is reasonable to infer that the variance could be 0.35 since the 95% confidence interval on the variance contains this value.

f) i) & ii) No, doctors and children would represent two completely different populations not represented by the population of Canadian Olympic hockey players. Since doctors nor children were the target of this study or part of the sample taken, the results should not be extended to these groups.

8-74 a.) The probability plot shows that the data appear to be normally distributed. Therefore, there is no evidence conclude that the comprehensive strength data are normally distributed.

b.) 99% lower confidence bound on the mean $\vec{x} = 25.12, \quad s = 8.42, \quad n = 9 \quad t_{0.01,8} = 2.896$

$$\bar{x} - t_{0.01,8}\left(\frac{s}{\sqrt{n}}\right) \le \mu$$

$$25.12 - 2.896\left(\frac{8.42}{\sqrt{9}}\right) \le \mu$$

$$16.99 \le \mu$$

The lower bound on the 99% confidence interval shows that the mean comprehensive strength will most likely be greater than 16.99 Megapascals.

c.) 98% lower confidence bound on the mean $\vec{x} = 25.12, \quad s = 8.42, \quad n = 9 \quad t_{0.01,8} = 2.896$

$$\bar{x} - t_{0.01,8}\left(\frac{s}{\sqrt{n}}\right) \le \mu \le \bar{x} - t_{0.01,8}\left(\frac{s}{\sqrt{n}}\right)$$

$$25.12 - 2.896\left(\frac{8.42}{\sqrt{9}}\right) \le \mu \le 25.12 - 2.896\left(\frac{8.42}{\sqrt{9}}\right)$$

$$16.99 \le \mu \le 33.25$$

The bounds on the 98% two-sided confidence interval shows that the mean comprehensive strength will most likely be greater than 16.99 Megapascals and less than 33.25 Megapascals. The lower bound of the 99% one sided CI is the same as the lower bound of the 98% two-sided CI (this is because of the value of α)

d.) 99% one-sided upper bound on the confidence interval on σ^2 comprehensive strength

$$s = 8.42, \quad s^2 = 70.90 \quad \chi^2_{0.99,8} = 1.65$$

$$\sigma^2 \le \frac{8(8.42)^2}{1.65}$$

$$\sigma^2 \le 343.74$$

The upper bound on the 99% confidence interval on the variance shows that the variance of the comprehensive strength will most likely be less than 343.74 Megapascals^2.

e.) 98% one-sided upper bound on the confidence interval on σ^2 comprehensive strength

$$s = 8.42, \quad s^2 = 70.90 \quad \chi^2_{0.01,9} = 20.09 \quad \chi^2_{0.99,8} = 1.65$$

$$\frac{8(8.42)^2}{20.09} \le \sigma^2 \le \frac{8(8.42)^2}{1.65}$$

$$28.23 \le \sigma^2 \le 343.74$$

The bounds on the 98% two-sided confidence-interval on the variance shows that the variance of the comprehensive strength will most likely be less than 343.74 Megapascals^2 and greater than 28.23 Megapascals^2. The upper bound of the 99% one-sided CI is the same as the upper bound of the 98% two-sided CI (this is because of the value of α)

f.) 98% lower confidence bound on the mean $\vec{x} = 23, \quad \text{s} = 6.07, \quad n = 9 \quad t_{0.01,8} = 2.896$

$$\bar{x} - t_{0.01,8}\left(\frac{s}{\sqrt{n}}\right) \le \mu \le \bar{x} - t_{0.01,8}\left(\frac{s}{\sqrt{n}}\right)$$

$$23 - 2.896\left(\frac{6.07}{\sqrt{9}}\right) \le \mu \le 23 - 2.896\left(\frac{6.07}{\sqrt{9}}\right)$$

$$17.14 \le \mu \le 28.86$$

98% one-sided upper bound on the confidence interval on σ^2 comprehensive strength

$$s = 6.07, \quad s^2 = 36.9 \quad \chi^2_{0.01,9} = 20.09 \quad \chi^2_{0.99,8} = 1.65$$

$$\frac{8(6.07)^2}{20.09} \le \sigma^2 \le \frac{8(6.07)^2}{1.65}$$

$$14.67 \le \sigma^2 \le 178.64$$

Fixing the mistake decreased the values of the sample mean and the sample standard deviation. Since the sample standard deviation was decreased. The width of the confidence intervals were also decreased.

g.) 98% lower confidence bound on the mean $\vec{x} = 25, \quad \text{s} = 8.41, \quad n = 9 \quad t_{0.01,8} = 2.896$

$$\bar{x} - t_{0.01,8}\left(\frac{s}{\sqrt{n}}\right) \le \mu \le \bar{x} - t_{0.01,8}\left(\frac{s}{\sqrt{n}}\right)$$

$$25 - 2.896\left(\frac{8.41}{\sqrt{9}}\right) \le \mu \le 25 - 2.896\left(\frac{8.41}{\sqrt{9}}\right)$$

$$16.88 \le \mu \le 33.12$$

98% one-sided upper bound on the confidence interval on σ^2 comprehensive strength

$$s = 8.41, \quad s^2 = 70.73 \quad \chi^2_{0.01,9} = 20.09 \quad \chi^2_{0.99,8} = 1.65$$

$$\frac{8(8.41)^2}{20.09} \le \sigma^2 \le \frac{8(8.41)^2}{1.65}$$

$$28.16 \le \sigma^2 \le 342.94$$

Fixing the mistake did not have an affect on the sample mean or the sample standard deviation. They are very close to the original values. The width of the confidence intervals are also very similar.

h.) When a mistaken value is near the sample mean, the mistake will not affect the sample mean, standard deviation or confidence intervals greatly. However, when the mistake is not near the sample mean, the value can greatly affect the sample mean, standard deviation and confidence intervals. The farther from the mean, the greater the effect.

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.

8-76 $\bar{x} = 15.33 \quad s = 0.62 \quad n = 20 \quad k = 2.564$

a.) 95% Tolerance Interval of hemoglobin values with 90% confidence

$$\bar{x} - ks, \ \bar{x} + ks$$
$$15.33 - 2.564(0.62), 15.33 + 2.564(0.62)$$
$$(13.74, \quad 16.92)$$

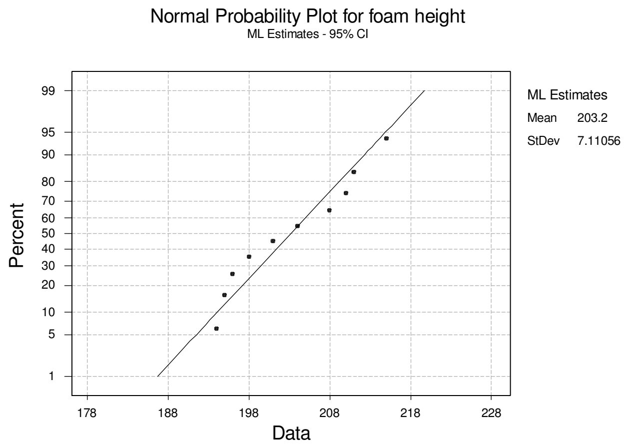
b.) 99% Tolerance Interval of hemoglobin values with 90% confidence $k = 3.368$

$$\bar{x} - ks, \ \bar{x} + ks$$
$$15.33 - 3.368(0.62), 15.33 + 3.368(0.62)$$
$$(13.24, 17.42)$$

8-77 95% prediction interval for the next sample of concrete that will be tested.
given $\bar{x} = 25.12$ $s = 8.42$ $n = 9$ for $\alpha = 0.05$ and $n = 9$, $t_{\alpha/2,n-1} = t_{0.025,8} = 2.306$

$$\bar{x} - t_{0.025,8}\, s\sqrt{1+\frac{1}{n}} \le x_{n+1} \le \bar{x} + t_{0.025,8}\, s\sqrt{1+\frac{1}{n}}$$
$$25.12 - 2.306(8.42)\sqrt{1+\frac{1}{9}} \le x_{n+1} \le 25.12 + 2.306(8.42)\sqrt{1+\frac{1}{9}}$$
$$4.65 \le x_{n+1} \le 45.59$$

8-78 a.) There is no evidence to reject the assumption that the data are normally distributed.



b.) 95% confidence interval on the mean $\bar{x} = 203.20, \quad s = 7.5, \quad n = 10 \quad t_{0.025,9} = 2.262$

$$\bar{x} - t_{0.025,9}\left(\frac{s}{\sqrt{n}}\right) \le \mu \le \bar{x} - t_{0.025,9}\left(\frac{s}{\sqrt{n}}\right)$$

$$203.2 - 2.262\left(\frac{7.50}{\sqrt{10}}\right) \le \mu \le 203.2 + 2.262\left(\frac{7.50}{\sqrt{10}}\right)$$

$$197.84 \le \mu \le 208.56$$

c.) 95% prediction interval on a future sample

$$\bar{x} - t_{0.025,9}s\sqrt{1+\frac{1}{n}} \le \mu \le \bar{x} - t_{0.025,9}s\sqrt{1+\frac{1}{n}}$$

$$203.2 - 2.262(7.50)\sqrt{1+\frac{1}{10}} \le \mu \le 203.2 + 2.262(7.50)\sqrt{1+\frac{1}{10}}$$

$$185.41 \le \mu \le 220.99$$

d.) 95% tolerance interval on foam height with 99% confidence $k = 4.265$

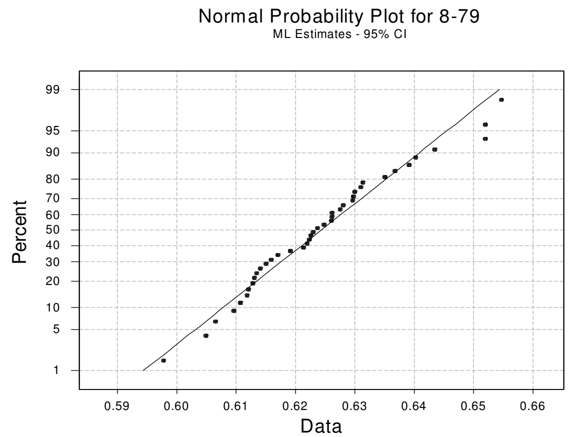
$$\bar{x} - ks, \bar{x} + ks$$

$$203.2 - 4.265(7.5), 203.2 + 4.265(7.5)$$

$$(171.21, 235.19)$$

e.) The 95% CI on the population mean has the smallest interval. This type of interval tells us that 95% of such intervals would contain the population mean. The 95% prediction interval, tell us where, most likely, the next data point will fall. This interval is quite a bit larger than the CI on the mean. The tolerance interval is the largest interval of all. It tells us the limits that will include 95% of the data with 99% confidence.

8-79 a) Normal probability plot for the coefficient of restitution



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

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

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

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

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

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

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

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

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

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

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

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

$$(0.582, 0.666)$$

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. And the tolerance interval captures 99% of the values of the normal distribution with a 95% level of confidence.

8-80 95% Confidence Interval on the death rate from lung cancer.

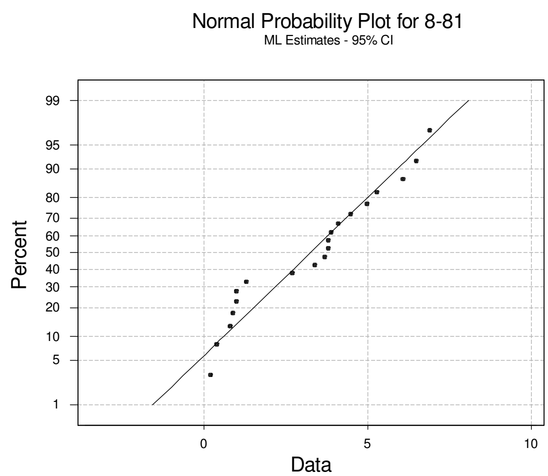
$$\hat{p} = \frac{8}{40} = 0.2 \qquad n = 40 \qquad z_\alpha = 1.65$$

$$\hat{p} - z_\alpha \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \le p$$

$$0.2 - 1.65\sqrt{\frac{0.2(0.8)}{40}} \le p$$

$$0.0956 \le p$$

8-81 a.)The normal probability shows that the data are mostly follow the straight line, however, there are some points that deviate from the line near the middle. It is probably safe to assume that the data are normal.



b.) 95% CI on the mean dissolved oxygen concentration

$\bar{x} = 3.265$, s = 2.127, n = 20 $t_{a/2,\ n-1} = t_{0.025,\ 19} = 2.093$

$$\bar{x} - t_{0.025,19}\frac{s}{\sqrt{n}} \le \mu \le \bar{x} + t_{0.025,19}\frac{s}{\sqrt{n}}$$

$$3.265 - 2.093\frac{2.127}{\sqrt{20}} \le \mu \le 3.265 + 2.093\frac{2.127}{\sqrt{20}}$$

$$2.270 \le \mu \le 4.260$$

c.) 95% prediction interval on the oxygen concentration for the next stream in the system that will be tested..

$$\bar{x} - t_{0.025,19}\, s\sqrt{1+\frac{1}{n}} \le x_{n+1} \le \bar{x} + t_{0.025,19}\, s\sqrt{1+\frac{1}{n}}$$

$$3.265 - 2.093(2.127)\sqrt{1+\frac{1}{20}} \le x_{n+1} \le 3.265 + 2.093(2.127)\sqrt{1+\frac{1}{20}}$$

$$-1.297 \le x_{n+1} \le 7.827$$

d.) 95% tolerance interval on the values of the dissolved oxygen concentration with a 99% level of confidence

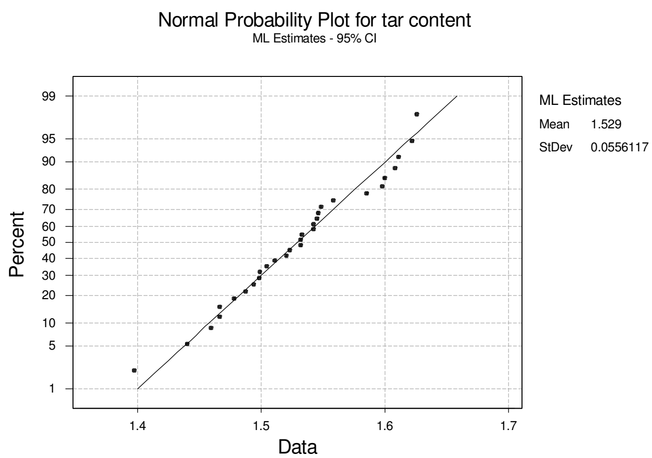
$$(\bar{x} - ks,\ \bar{x} + ks)$$

$$(3.265 - 3.168(2.127),\ 3.265 + 3.168(2.127))$$

$$(-3.473,\ 10.003)$$

e.)The confidence interval in part (b) describes the confidence interval on the population mean and we may interpret this to mean that there is a 95% probability that the interval may cover the population mean. The prediction interval tells us that within that within a 95% probability that the next stream will have an oxygen concentration between –1.297 and 7.827mg/L. And the tolerance interval captures 95% of the values of the normal distribution with a 99% confidence level.

8-82 a.) There is no evidence to support that the data are not normally distributed. The data points appear to fall along the normal probability line.



b.) 99% CI on the mean tar content

$\bar{x} = 1.529$, s = 0.0566, n = 30 $t_{a/2,\ n-1} = t_{0.005,\ 29} = 2.756$

$$\bar{x} - t_{0.005,29}\frac{s}{\sqrt{n}} \le \mu \le \bar{x} + t_{0.005,29}\frac{s}{\sqrt{n}}$$

$$1.529 - 2.756\frac{0.0566}{\sqrt{30}} \le \mu \le 1.529 + 2.756\frac{0.0566}{\sqrt{30}}$$

$$1.501 \le \mu \le 1.557$$

e.) 99% prediction interval on the tar content for the next sample that will be tested..

$$\bar{x} - t_{0.005,19}s\sqrt{1+\frac{1}{n}} \le x_{n+1} \le \bar{x} + t_{0.005,19}s\sqrt{1+\frac{1}{n}}$$

$$1.529 - 2.756(0.0566)\sqrt{1+\frac{1}{30}} \le x_{n+1} \le 1.529 + 2.756(0.0566)\sqrt{1+\frac{1}{30}}$$

$$1.370 \le x_{n+1} \le 1.688$$

f.) 99% tolerance interval on the values of the tar content with a 95% level of confidence

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

$$(1.529 - 3.350(0.0566), 1.529 + 3.350(0.0566))$$

$$(1.339, 1.719)$$

e.)The confidence interval in part (b) describes the confidence interval on the population mean and we may interpret this to mean that 95% of such intervals will cover the population mean. The prediction interval tells us that within that within a 95% probability that the sample will have a tar content between 1.370 and 1.688. And the tolerance interval captures 95% of the values of the normal distribution with a 99% confidence level.

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

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

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

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

$$0.0021 \le p \le 0.0113$$

b) No, there is not evidence to support the claim that the fraction of defective units produced is one percent or less. This is because the upper limit of the control limit is greater than 0.01.

8-84 a.) 99% Confidence Interval on the population proportion

n=1600 x=8 $\hat{p} = 0.005$ $z_{\alpha/2}=z_{0.005}=2.58$

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

$$0.005 - 2.58\sqrt{\frac{0.005(1-0.005)}{1600}} \le p \le 0.005 + 2.58\sqrt{\frac{0.005(1-0.005)}{1600}}$$

$$0.0004505 \le p \le 0.009549$$

b.) E = 0.008, α = 0.01, $z_{\alpha/2} = z_{0.005} = 2.58$

$$n = \left(\frac{z_{\alpha/2}}{E}\right)^2 p(1-p) = \left(\frac{2.58}{0.008}\right)^2 0.005(1-0.005) = 517.43, \quad n \cong 518$$

c.) E = 0.008, α = 0.01, $z_{\alpha/2} = z_{0.005} = 2.58$

$$n = \left(\frac{z_{\alpha/2}}{E}\right)^2 p(1-p) = \left(\frac{2.58}{0.008}\right)^2 0.5(1-0.5) = 26001.56, \quad n \cong 26002$$

d.)Knowing an estimate of the population proportion reduces the required sample size by a significant amount. A sample size of 518 is much more reasonable than a sample size of over 26,000.

8-85 $\hat{p} = \frac{117}{484} = 0.242$

a) 90% confidence interval; $z_{\alpha/2} = 1.645$

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

$$0.210 \le p \le 0.274$$

With 90% confidence, we believe the true proportion of new engineering graduates who were planning to continue studying for an advanced degree lies between 0.210 and 0.274.

b) 95% confidence interval; $z_{\alpha/2} = 1.96$

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

$$0.204 \le p \le 0.280$$

With 95% confidence, we believe the true proportion of new engineering graduates who were planning to continue studying for an advanced degree lies between 0.204 and 0.280.

c) Comparison of parts a and b:

The 95% confidence interval is larger than the 90% confidence interval. Higher confidence always yields larger intervals, all other values held constant.

d) Yes, since both intervals contain the value 0.25, thus there in not enough evidence to determine that the true proportion is not actually 0.25.

Mind Expanding Exercises

8-86 a.) $P(\chi^2_{1-\frac{\alpha}{2},2r} < 2\lambda T_r < \chi^2_{\frac{\alpha}{2},2r}) = 1-\alpha$

$$= P\left(\frac{\chi^2_{1-\frac{\alpha}{2},2r}}{2T_r} < \lambda < \frac{\chi^2_{\frac{\alpha}{2},2r}}{2T_r}\right)$$

Then a confidence interval for $\mu = \frac{1}{\lambda}$ is $\left(\frac{2T_r}{\chi^2_{\frac{\alpha}{2},2r}}, \frac{2T_r}{\chi^2_{1-\frac{\alpha}{2},2r}}\right)$

b) n = 20 , r = 10 , and the observed value of T_r is 199 + 10(29) = 489.

A 95% confidence interval for $\frac{1}{\lambda}$ is $\left(\frac{2(489)}{34.17}, \frac{2(489)}{9.59}\right) = (28.62, 101.98)$

8-87 $\alpha_1 = \int_{z_{\alpha_1}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 1 - \int_{-\infty}^{z_{\alpha_1}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$

Therefore, $1-\alpha_1 = \Phi(z_{\alpha_1})$.

To minimize L we need to minimize $\Phi^{-1}(1-\alpha_1)+\Phi(1-\alpha_2)$ subject to $\alpha_1+\alpha_2=\alpha$. Therefore, we need to minimize $\Phi^{-1}(1-\alpha_1)+\Phi(1-\alpha+\alpha_1)$.

$$\frac{\partial}{\partial\alpha_1}\Phi^{-1}(1-\alpha_1) = -\sqrt{2\pi}\ e^{\frac{z_{\alpha_1}^2}{2}}$$

$$\frac{\partial}{\partial\alpha_1}\Phi^{-1}(1-\alpha+\alpha_1) = \sqrt{2\pi}\ e^{\frac{z_{\alpha-\alpha_1}^2}{2}}$$

Upon setting the sum of the two derivatives equal to zero, we obtain $e^{\frac{z_{\alpha-\alpha_1}^2}{2}} = e^{\frac{z_{\alpha_1}^2}{2}}$. This is solved by $z_{\alpha_1} = z_{\alpha-\alpha_1}$. Consequently, $\alpha_1 = \alpha-\alpha_1$, $2\alpha_1=\alpha$ and $\alpha_1=\alpha_2=\frac{\alpha}{2}$.

8.88 a.) n=1/2+(1.9/.1)(9.4877/4)
n=46

b.) (10-.5)/(9.4877/4)=(1+p)/(1-p)
p=0.6004 between 10.19 and 10.41.

8-89 a)

$$P(X_i \le \tilde{\mu}) = 1/2$$
$$P(allX_i \le \tilde{\mu}) = (1/2)^n$$
$$P(allX_i \ge \tilde{\mu}) = (1/2)^n$$

$$P(A\cup B) = P(A)+P(B)-P(A\cap B)$$
$$= \left(\frac{1}{2}\right)^n + \left(\frac{1}{2}\right)^n = 2\left(\frac{1}{2}\right)^n = \left(\frac{1}{2}\right)^{n-1}$$
$$1-P(A\cup B) = P(\min(X_i)<\tilde{\mu}<\max(X_i)) = 1-\left(\frac{1}{2}\right)^n$$

b.) $P(\min(X_i)<\tilde{\mu}<\max(X_i)) = 1-\alpha$

The confidence interval is min(X_i), max(X_i)

8-90 We would expect that 950 of the confidence intervals would include the value of μ. This is due to the definition of a confidence interval.

Let X bet the number of intervals that contain the true mean (μ). We can use the large sample approximation to determine the probability that P(930<X<970).

Let $p = \frac{950}{1000} = 0.950 \quad p_1 = \frac{930}{1000} = 0.930$ and $p_2 = \frac{970}{1000} = 0.970$

The variance is estimated by $\frac{p(1-p)}{n} = \frac{0.950(0.050)}{1000}$

$$P(0.930 < p < 0.970) = P\left(Z < \frac{(0.970-0.950)}{\sqrt{\frac{0.950(0.050)}{1000}}}\right) - P\left(Z < \frac{(0.930-0.950)}{\sqrt{\frac{0.950(0.050)}{1000}}}\right)$$

$$= P\left(Z < \frac{0.02}{0.006892}\right) - P\left(Z < \frac{-0.02}{0.006892}\right) = P(Z < 2.90) - P(Z < -2.90) = 0.9963$$

CHAPTER 9

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 null 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.

9-2 a) α = P(reject H_0 when H_0 is true)

= P($\bar{X} \le 11.5$ when $\mu = 12$) = $P\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \le \frac{11.5-12}{0.5/\sqrt{4}}\right)$ = P(Z ≤ −2)

= 0.02275.

The probability of rejecting the null hypothesis when it is true is 0.02275.

b) β = P(accept H_0 when $\mu = 11.25$) = $P(\bar{X} > 11.5 \text{ when } \mu = 11.25)$ = $P\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} > \frac{11.5-11.25}{0.5/\sqrt{4}}\right)$

P(Z > 1.0) = 1 − P(Z ≤ 1.0) = 1 − 0.84134 = 0.15866

The probability of accepting the null hypothesis when it is false is 0.15866.

9-3 a) α = P($\bar{X} \le 11.5 \mid \mu = 12$) = $P\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \le \frac{11.5-12}{0.5/\sqrt{16}}\right)$ = P(Z ≤ −4) = 0.

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

b) β = 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 ≤ 2)

= 1− 0.97725 = 0.02275.

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

9-4 Find the boundary of the critical region if $\alpha = 0.01$:

$$0.01 = P\left(Z \le \frac{c-12}{0.5/\sqrt{4}}\right)$$

What Z value will give a probability of 0.01? Using Table 2 in the appendix, Z value is −2.33.

Thus, $\frac{c-12}{0.5/\sqrt{4}} = -2.33$, c = 11.4175

9-5. $P\left(Z \le \frac{c-12}{0.5/\sqrt{4}}\right) = 0.05$

What Z value will give a probability of 0.05? Using Table 2 in the appendix, Z value is −1.65.

Thus, $\frac{c-12}{0.5/\sqrt{4}} = -1.65$, c = 11.5875

9-6 a) $\alpha = P(\overline{X} \le 98.5) + P(\overline{X} > 101.5)$

$$= P\left(\frac{\overline{X}-100}{2/\sqrt{9}} \le \frac{98.5-100}{2/\sqrt{9}}\right) + P\left(\frac{\overline{X}-100}{2/\sqrt{9}} > \frac{101.5-100}{2/\sqrt{9}}\right)$$

= P(Z ≤ −2.25) + P(Z > 2.25)
= (P(Z ≤- 2.25)) + (1 − P(Z ≤ 2.25))
= 0.01222 + 1 − 0.98778
= 0.01222 + 0.01222 = 0.02444

b) $\beta = P(98.5 \le \overline{X} \le 101.5 \text{ when } \mu = 103)$

$$= P\left(\frac{98.5-103}{2/\sqrt{9}} \le \frac{\overline{X}-103}{2/\sqrt{9}} \le \frac{101.5-103}{2/\sqrt{9}}\right)$$

= P(−6.75 ≤ Z ≤ −2.25)
= P(Z ≤ −2.25) − P(Z ≤ −6.75)
= 0.01222 − 0 = 0.01222

c) $\beta = P(98.5 \le \overline{X} \le 101.5 \text{ when } \mu = 105)$

$$= P\left(\frac{98.5-105}{2/\sqrt{9}} \le \frac{\overline{X}-105}{2/\sqrt{9}} \le \frac{101.5-105}{2/\sqrt{9}}\right)$$

= P(−9.75≤ Z ≤ −5.25)
= P(Z ≤ −5.25) − P(Z ≤ −9.75)
= 0 − 0
= 0.

The probability of accepting the null hypothesis when it is actually false is smaller in part c since the true mean, μ = 105, is further from the acceptance region. A larger difference exists.

9-7 Use *n* = 5, everything else held constant (from the values in exercise 9-6):

a) $P(\overline{X} \le 98.5) + P(\overline{X} > 101.5)$

$$= P\left(\frac{\overline{X}-100}{2/\sqrt{5}} \le \frac{98.5-100}{2/\sqrt{5}}\right) + P\left(\frac{\overline{X}-100}{2/\sqrt{5}} > \frac{101.5-100}{2/\sqrt{5}}\right)$$

= P(Z ≤ −1.68) + P(Z > 1.68)
= P(Z ≤ −1.68) + (1 − P(Z ≤ 1.68))
= 0.04648 + (1 − 0.95352)
= 0.09296

b) $\beta = P(98.5 \le \overline{X} \le 101.5 \text{ when } \mu = 103)$

$$= P\left(\frac{98.5-103}{2/\sqrt{5}} \le \frac{\overline{X}-103}{2/\sqrt{5}} \le \frac{101.5-103}{2/\sqrt{5}}\right)$$

= P(−5.03 ≤ Z ≤ −1.68)
= P(Z ≤ −1.68) − P(Z ≤ −5.03)
= 0.04648 − 0
= 0.04648

c) $\beta = P(98.5 \le \bar{x} \le 101.5 \text{ when } \mu = 105)$

$$= P\left(\frac{98.5-105}{2/\sqrt{5}} \le \frac{\bar{X}-105}{2/\sqrt{5}} \le \frac{101.5-105}{2/\sqrt{5}}\right)$$

$= P(-7.27 \le Z \le -3.91)$
$= P(Z \le -3.91) - P(Z \le -7.27)$
$= 0.00005 - 0$
$= 0.00005$

It is smaller, because it is not likely to accept the product when the true mean is as high as 105.

9-8 a) $\alpha = P(\bar{X} > 185 \text{ when } \mu = 175)$

$$= P\left(\frac{\bar{X}-175}{20/\sqrt{10}} > \frac{185-175}{20/\sqrt{10}}\right)$$

$= P(Z > 1.58)$
$= 1 - P(Z \le 1.58)$
$= 1 - 0.94295$
$= 0.05705$

b) $\beta = P(\bar{X} \le 185 \text{ when } \mu = 195)$

$$= P\left(\frac{\bar{X}-195}{20/\sqrt{10}} \le \frac{185-195}{20/\sqrt{10}}\right)$$

$= P(Z \le -1.58)$
$= 0.05705.$

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

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

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

$= P(Z > 2.37) = 1 - P(Z \le 2.37)$
$= 1 - 0.99111$
$= 0.00889.$

The probability that a value of at least 190 mm would be observed (if the true mean height is 175 mm) is only 0.00889. Thus, the sample value of $\bar{x} = 190$ mm would be an unusual result.

9-10 Using n = 16:

a) $\alpha = P(\bar{X} > 185 \text{ when } \mu = 175)$

$$= P\left(\frac{\bar{X}-175}{20/\sqrt{16}} > \frac{185-175}{20/\sqrt{16}}\right)$$

$= P(Z > 2)$
$= 1 - P(Z \le 2)$
$= 1 - 0.97725$
$= 0.02275$

b) β = P($\bar{X}$ ≤ 185 when μ = 195)

$$= P\left(\frac{\bar{X}-195}{20/\sqrt{16}} \le \frac{185-195}{20/\sqrt{16}}\right)$$

= P(Z ≤ −2)
= 0.02275.

9-11 a) $P(\bar{X} > c \,|\, \mu = 175) = 0.0571$

$$P\left(Z > \frac{c-175}{20/\sqrt{16}}\right) = P(Z \ge 1.58)$$

Thus, $1.58 = \dfrac{c-175}{20/\sqrt{16}}$, and $c = 182.9$

b) If the true mean foam height is 195 mm, then

β = P($\bar{X}$ ≤ 182.9 when μ = 195)

$$= P\left(Z \le \frac{182.9-195}{20/\sqrt{16}}\right)$$

= P(Z ≤ −2.42)
= 0.00776

c) For the same level of α, with the increased sample size, β is reduced.

9-12 a) α = P($\bar{X}$ ≤ 4.85 when μ = 5) + P($\bar{X}$ > 5.15 when μ = 5)

$$= P\left(\frac{\bar{X}-5}{0.25/\sqrt{8}} \le \frac{4.85-5}{0.25/\sqrt{8}}\right) + P\left(\frac{\bar{X}-5}{0.25/\sqrt{8}} > \frac{5.15-5}{0.25/\sqrt{8}}\right)$$

= P(Z ≤ −1.7) + P(Z > 1.7)
= P(Z ≤ −1.7) + (1 − P(Z ≤ 1.7)
= 0.04457 + (1 − 0.95543)
= 0.08914.

b) Power = 1 − β

β = P(4.85 ≤ $\bar{X}$ ≤ 5.15 when μ = 5.1)

$$= P\left(\frac{4.85-5.1}{0.25/\sqrt{8}} \le \frac{\bar{X}-5.1}{0.25/\sqrt{8}} \le \frac{5.15-5.1}{0.25/\sqrt{8}}\right)$$

= P(−2.83 ≤ Z ≤ 0.566)
= P(Z ≤ 0.566) − P(Z ≤ −2.83)
= 0.71566 − 0.00233
= 0.71333
1 − β = 0.2867.

9-13 Using n = 16:

a) $\alpha = P(\overline{X} \le 4.85 \mid \mu = 5) + P(\overline{X} > 5.15 \mid \mu = 5)$

$$= P\left(\frac{\overline{X}-5}{0.25/\sqrt{16}} \le \frac{4.85-5}{0.25/\sqrt{16}}\right) + P\left(\frac{\overline{X}-5}{0.25/\sqrt{16}} > \frac{5.15-5}{0.25/\sqrt{16}}\right)$$

$= P(Z \le -2.4) + P(Z > 2.4)$
$= P(Z \le -2.4) + (1 - P(Z \le 2.4))$
$= 2(1 - P(Z \le 2.4))$
$= 2(1 - 0.99180)$
$= 2(0.0082)$
$= 0.0164.$

b) $\beta = P(4.85 \le \overline{X} \le 5.15 \mid \mu = 5.1)$

$$= P\left(\frac{4.85-5.1}{0.25/\sqrt{16}} \le \frac{\overline{X}-5.1}{0.25/\sqrt{16}} \le \frac{5.15-5.1}{0.25/\sqrt{16}}\right)$$

$= P(-4 \le Z \le 0.8) = P(Z \le 0.8) - P(Z \le -4)$
$= 0.78814 - 0$
$= 0.78814$
$1 - \beta = 0.21186$

9-14 Find the boundary of the critical region if $\alpha = 0.05$:

$$0.025 = P\left(Z \le \frac{c-5}{0.25/\sqrt{8}}\right)$$

What Z value will give a probability of 0.025? Using Table 2 in the appendix, Z value is −1.96.

Thus, $\frac{c-5}{0.25/\sqrt{8}} = -1.96$, c = 4.83 and

$$\frac{c-5}{0.25/\sqrt{8}} = 1.96, \quad c=5.17$$

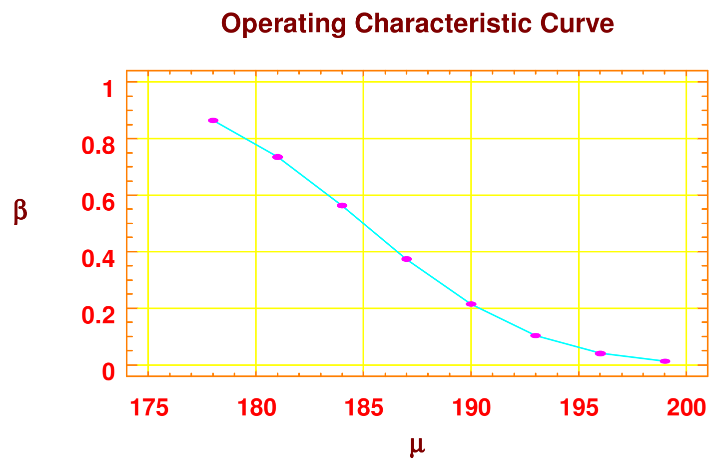
The acceptance region should be $(4.83 \le \overline{X} \le 5.17)$.

9-15 Operating characteristic curve:

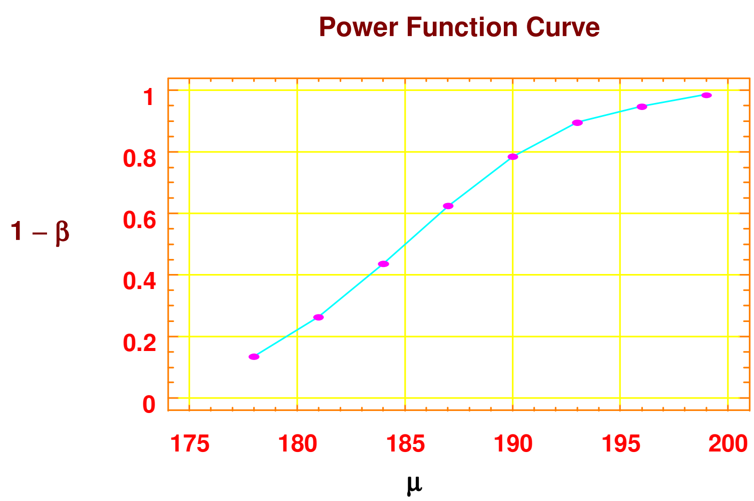
$\overline{x} = 185$

$$\beta = P\left(Z \le \frac{\overline{x}-\mu}{20/\sqrt{10}}\right) = P\left(Z \le \frac{185-\mu}{20/\sqrt{10}}\right)$$

μ	$P\left(Z \le \frac{185-\mu}{20/\sqrt{10}}\right) =$	β	$1-\beta$
178	$P(Z \le 1.11) =$	0.8665	0.1335
181	$P(Z \le 0.63) =$	0.7357	0.2643
184	$P(Z \le 0.16) =$	0.5636	0.4364
187	$P(Z \le -0.32) =$	0.3745	0.6255
190	$P(Z \le -0.79) =$	0.2148	0.7852
193	$P(Z \le -1.26) =$	0.1038	0.8962
196	$P(Z \le -1.74) =$	0.0409	0.9591
199	$P(Z \le -2.21) =$	0.0136	0.9864



9-16



9-17. The problem statement implies H_0: $p = 0.6$, H_1: $p > 0.6$ and defines an acceptance region as

$$\hat{p} \le \frac{400}{500} = 0.80 \text{ and rejection region as } \hat{p} > 0.80$$

a) $\alpha = P(\hat{p} > 0.80 \mid p = 0.60) = P\left(Z > \dfrac{0.80 - 0.60}{\sqrt{\dfrac{0.6(0.4)}{500}}} \right)$

$= P(Z > 9.13) = 1 - P(Z \le 9.13) \approx 0$

b) $\beta = P(\hat{p} \le 0.8 \text{ when } p = 0.75) = P(Z \le 2.58) = 0.99506.$

9-18 $X \sim \text{bin}(10, 0.3)$ Implicitly, H_0: $p = 0.3$ and H_1: $p < 0.3$

$n = 10$

Accept region: $\hat{p} > 0.1$

Reject region: $\hat{p} \le 0.1$

Use the normal approximation for parts a), b) and c):

a) When p =0.3 α = $P(\hat{p} < 0.1) = P\left(Z \le \frac{0.1 - 0.3}{\sqrt{\frac{0.3(0.7)}{10}}}\right)$

$$= P(Z \le -1.38)$$
$$= 0.08379$$

b) When p = 0.2 β = $P(\hat{p} > 0.1) = P\left(Z > \frac{0.1 - 0.2}{\sqrt{\frac{0.2(0.8)}{10}}}\right)$

$$= P(Z > -0.79)$$
$$= 1 - P(Z < -0.79)$$
$$= 0.78524$$

c) Power = 1 − β = 1 − 078524 = 0.21476

9-19 X ~ bin(15, 0.4) H_0: $p = 0.4$ and H_1: $p \neq 0.4$

p_1= 4/15 = 0.267 p_2 = 8/15 = 0.533

Accept Region: $0.267 \le \hat{p} \le 0.533$

Reject Region: $\hat{p} < 0.267$ or $\hat{p} > 0.533$

Use the normal approximation for parts a) and b)

a) When p = 0.4, $\alpha = P(\hat{p} < 0.267) + P(\hat{p} > 0.533)$

$$= P\left(Z < \frac{0.267 - 0.4}{\sqrt{\frac{0.4(0.6)}{15}}}\right) + P\left(Z > \frac{0.533 - 0.4}{\sqrt{\frac{0.4(0.6)}{15}}}\right)$$
$$= P(Z < -1.05) + P(Z > 1.05)$$
$$= P(Z < -1.05) + (1 - P(Z < 1.05))$$
$$= 0.14686 + 0.14686$$
$$= 0.29372$$

b) When p = 0.2,

$$\beta = P(0.267 \le \hat{p} \le 0.533) = P\left(\frac{0.267-0.2}{\sqrt{\frac{0.2(0.8)}{15}}} \le Z \le \frac{0.533-0.2}{\sqrt{\frac{0.2(0.8)}{15}}}\right)$$
$$= P(0.65 \le Z \le 3.22)$$
$$= P(Z \le 3.22) - P(Z \le 0.65)$$
$$= 0.99936 - 0.74215$$
$$= 0.25721$$

Section 9-2

9-20 a.) 1) The parameter of interest is the true mean water temperature, μ.
2) $H_0 : \mu = 100$
3) $H_1 : \mu > 100$
4) $\alpha = 0.05$
5) $z_0 = \frac{\overline{x} - \mu}{\sigma / \sqrt{n}}$
6) Reject H_0 if $z_0 > z_\alpha$ where $z_{0.05} = 1.65$
7) $\overline{x} = 98$, $\sigma = 2$

$$z_0 = \frac{98 - 100}{2/\sqrt{9}} = -3.0$$

8) Since -3.0 < 1.65 do not reject H_0 and conclude the water temperature is not significantly different greater than 100 at $\alpha = 0.05$.

b) P-value = $1 - \Phi(-3.0) = 1 - 0.00135 = 0.99865$

c) $\beta = \Phi\left(z_{0.05} + \frac{100-104}{2/\sqrt{9}}\right)$

$= \Phi(1.65 + -6)$
$= \Phi(-4.35)$
$\cong 0$

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) $z_0 = \frac{\overline{x} - \mu}{\sigma / \sqrt{n}}$
6) Reject H_0 if $z_0 < -z_{\alpha/2}$ where $-z_{0.025} = -1.96$ or $z_0 > z_{\alpha/2}$ where $z_{0.025} = 1.96$
7) $\overline{x} = 90.48$, $\sigma = 3$

$$z_0 = \frac{90.48 - 90}{3/\sqrt{5}} = 0.36$$

8) Since $-1.96 < 0.36 < 1.96$ do not reject H_0 and conclude the yield is not significantly different from 90% at $\alpha = 0.05$.

b) P-value = $2[1 - \Phi(0.36)] = 2[1 - 0.64058] = 0.71884$

c) $n = \frac{(z_{\alpha/2} + z_\beta)^2 \sigma^2}{\delta^2} = \frac{(z_{0.025} + z_{0.05})^2 3^2}{(85-90)^2} = \frac{(1.96 + 1.65)^2 9}{(-5)^2} = 4.69$

$n \cong 5$.

d) $\beta = \Phi\left(z_{0.025} + \frac{90-92}{3/\sqrt{5}}\right) - \Phi\left(-z_{0.025} + \frac{90-92}{3/\sqrt{5}}\right)$

$= \Phi(1.96 + -1.491) - \Phi(-1.96 + -1.491)$
$= \Phi(0.47) - \Phi(-3.45)$
$= \Phi(0.47) - (1 - \Phi(3.45))$
$= 0.68082 - (1 - 0.99972)$
$= 0.68054.$

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

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

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

$87.85 \le \mu \le 93.11$

With 95% confidence, we believe the true mean yield of the chemical process is between 87.85% and 93.11%. Since 90% is contained in the confidence interval, our decision in (a) agrees with the confidence interval.

9-22 a) 1) The parameter of interest is the true mean crankshaft wear, μ.
2) $H_0 : \mu = 3$
3) $H_1 : \mu \neq 3$
4) $\alpha = 0.05$
5) $z_0 = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$
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) $\bar{x} = 2.78$, $\sigma = 0.9$

$$z_0 = \frac{2.78 - 3}{0.9/\sqrt{15}} = -0.95$$

8) Since $-0.95 > -1.96$, do not reject the null hypothesis and conclude there is not sufficient evidence to support the claim the mean crankshaft wear is not equal to 3 at $\alpha = 0.05$.

b) $\beta = \Phi\left(z_{0.025} + \frac{3 - 3.25}{0.9/\sqrt{15}}\right) - \Phi\left(-z_{0.025} + \frac{3 - 3.25}{0.9/\sqrt{15}}\right)$

$= \Phi(1.96 + -1.08) - \Phi(-1.96 + -1.08)$
$= \Phi(0.88) - \Phi(-3.04)$
$= 0.81057 - (0.00118)$
$= 0.80939$

c.) $n = \frac{(z_{\alpha/2} + z_\beta)^2 \sigma^2}{\delta^2} = \frac{(z_{0.025} + z_{0.10})^2 \sigma^2}{(3.75-3)^2} = \frac{(1.96 + 1.29)^2 (0.9)^2}{(0.75)^2} = 15.21,$

$n \cong 16$

9-23. a) 1) The parameter of interest is the true mean melting point, μ.

2) $H_0 : \mu = 155$

3) $H_1 : \mu \neq 155$

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) $\overline{x} = 154.2$, $\sigma = 1.5$

$$z_0 = \frac{154.2 - 155}{1.5/\sqrt{10}} = -1.69$$

8) Since –1.69 > -2.58, do not reject the null hypothesis and conclude there is not sufficient evidence to support the claim the mean melting point is not equal to 155 °F at $\alpha = 0.01$.

b) P-value = 2*P(Z <- 1.69) =2* 0.045514 = 0.091028

c)
$$\beta = \Phi\left(z_{0.005} - \frac{\delta\sqrt{n}}{\sigma}\right) - \Phi\left(-z_{0.005} - \frac{\delta\sqrt{n}}{\sigma}\right)$$
$$= \Phi\left(2.58 - \frac{(155-150)\sqrt{10}}{1.5}\right) - \Phi\left(-2.58 - \frac{(155-150)\sqrt{10}}{1.5}\right)$$

$= \Phi(-7.96) - \Phi(-13.12) = 0 - 0 = 0$

d)
$$n = \frac{(z_{\alpha/2} + z_\beta)^2 \sigma^2}{\delta^2} = \frac{(z_{0.005} + z_{0.10})^2 \sigma^2}{(150-155)^2} = \frac{(2.58+1.29)^2 (1.5)^2}{(5)^2} = 1.35,$$

$n \cong 2$.

9-24 a.) 1) The parameter of interest is the true mean battery life in hours, μ.

2) $H_0 : \mu = 40$

3) $H_1 : \mu > 40$

4) $\alpha = 0.05$

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

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

7) $\overline{x} = 40.5$, $\sigma = 1.25$

$$z_0 = \frac{40.5 - 40}{1.25/\sqrt{10}} = 1.26$$

8) Since 1.26 < 1.65 do not reject H_0 and conclude the battery life is not significantly different greater than 40 at $\alpha = 0.05$.

b) P-value = $1 - \Phi(1.26) = 1 - 0.8962 = 0.1038$

c) $\beta = \Phi\left(z_{0.05} + \dfrac{40-42}{1.25/\sqrt{10}}\right)$

$= \Phi(1.65 + -5.06)$

$= \Phi(-3.41)$

$\cong 0.000325$

d.) $n = \dfrac{(z_\alpha + z_\beta)^2 \sigma^2}{\delta^2} = \dfrac{(z_{0.05} + z_{0.10})^2 \sigma^2}{(40-44)^2} = \dfrac{(1.65+1.29)^2 (1.25)^2}{(4)^2} = 0.844,$

$n \cong 1$

e.)95% Confidence Interval

$$\bar{x} + z_{0.05}\sigma / \sqrt{n} \le \mu$$

$$40.5 + 1.65(1.25)/\sqrt{10} \le \mu$$

$$39.85 \le \mu$$

The lower bound of the 90 % confidence interval must be greater than 40 to verify that the true mean exceeds 40 hours.

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

2) $H_0 : \mu = 3500$

3) $H_1 : \mu \ne 3500$

4) $\alpha = 0.01$

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

6) Reject H_0 if $z_0 < -z_{\alpha/2}$ where $-z_{0.005} = -2.58$ or $z_0 > z_{\alpha/2}$ where $z_{0.005} = 2.58$

7) $\bar{x} = 3250$, $\sigma = 60$

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

8) Since $-14.43 < -2.58$, reject the null hypothesis and conclude the true mean tensile 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) $z_{\alpha/2} = z_{0.005} = 2.58$

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

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

$3205.31 \le \mu \le 3294.69$

With 95% confidence, we believe the true mean tensile strength is between 3205.31 psi and 3294.69 psi. We can test the hypotheses that the true mean tensile strength is not equal to 3500 by noting that the value is not within the confidence interval.

9-26

$$n = \frac{(z_{\alpha/2} + z_\beta)^2\sigma^2}{\delta^2} = \frac{(z_{0.05} + z_{0.20})^2\sigma^2}{(3250 - 3500)^2} = \frac{(1.65 + .84)^2(60)^2}{(250)^2} = 0.357,$$

$$n \cong 1$$

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

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

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

$$z_0 = \frac{102.2 - 100}{4/\sqrt{8}} = 1.56$$

8) Since 1.56> −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) $\beta = \Phi\left(-z_{0.05} - \dfrac{(95-100)\sqrt{8}}{4}\right)$ = Φ(-1.65 - −3.54) = Φ(1.89) = 0.97062

Power = 1-β = 1-0.97062 = 0.02938

c) n = $\dfrac{(z_\alpha + z_\beta)^2 \sigma^2}{\delta^2} = \dfrac{(z_{0.05} + z_{0.15})^2 \sigma^2}{(95-100)^2} = \dfrac{(1.65+1.03)^2 (4)^2}{(5)^2} = 4.597,$

$n \cong 5$

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

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

$$99.866 \le \mu$$

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

9-28 a) 1) The parameter of interest is the true mean hole diameter, μ.

2) $H_0 : \mu = 1.50$

3) $H_1 : \mu \neq 1.50$

4) $\alpha = 0.01$

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

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) $\overline{x} = 1.4975$, $\sigma = 0.01$

$$z_0 = \frac{1.4975 - 1.50}{0.01/\sqrt{25}} = -1.25$$

8) Since $-2.58 < -1.25 < 2.58$, do not reject the null hypothesis and conclude the true mean hole diameter is not significantly different from 1.5 in. at $\alpha = 0.01$.

b)

$$\beta = \Phi\left(z_{0.005} - \frac{\delta\sqrt{n}}{\sigma}\right) - \Phi\left(-z_{0.005} - \frac{\delta\sqrt{n}}{\sigma}\right)$$

$$= \Phi\left(2.58 - \frac{(1.495-1.5)\sqrt{25}}{0.01}\right) - \Phi\left(-2.58 - \frac{(1.495-1.5)\sqrt{25}}{0.01}\right)$$

$= \Phi(5.08) - \Phi(-0.08) = 1 - .46812 = 0.53188$

power=1-β=0.46812.

c) Set $\beta = 1 - 0.90 = 0.10$

$$n = \frac{(z_{\alpha/2} + z_{\beta})^2\sigma^2}{\delta^2} = \frac{(z_{0.005} + z_{0.10})^2\sigma^2}{(1.495-1.50)^2} \cong \frac{(2.58+1.29)^2(0.01)^2}{(-0.005)^2} = 59.908,$$

$n \cong 60$.

d) For $\alpha = 0.01$, $z_{\alpha/2} = z_{0.005} = 2.58$

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

$$1.4975 - 2.58\left(\frac{0.01}{\sqrt{25}}\right) \le \mu \le 1.4975 + 2.58\left(\frac{0.01}{\sqrt{25}}\right)$$

$1.4923 \le \mu \le 1.5027$

The confidence interval constructed contains the value 1.5, thus the true mean hole diameter could possibly be 1.5 in. using a 99% level of confidence. Since a two-sided 99% confidence interval is equivalent to a two-sided hypothesis test at $\alpha = 0.01$, the conclusions necessarily must be consistent.

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 there is sufficient evidence to conclude that the true average battery life exceeds 4 hours at $\alpha = 0.05$.

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

Power = 1-β = 1-0 = 1

c) $n = \dfrac{(z_\alpha + z_\beta)^2\sigma^2}{\delta^2} = \dfrac{(z_{0.05} + z_{0.1})^2\sigma^2}{(4.5-4)^2} = \dfrac{(1.65+1.29)^2(0.2)^2}{(0.5)^2} = 1.38,$

$n \cong 2$

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

9-30 a. 1) The parameter of interest is the true mean interior temperature life, μ.

2) $H_0 : \mu = 22.5$

3) $H_1 : \mu \ne 22.5$

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 $t_{\alpha/2,n-1} = 2.776$

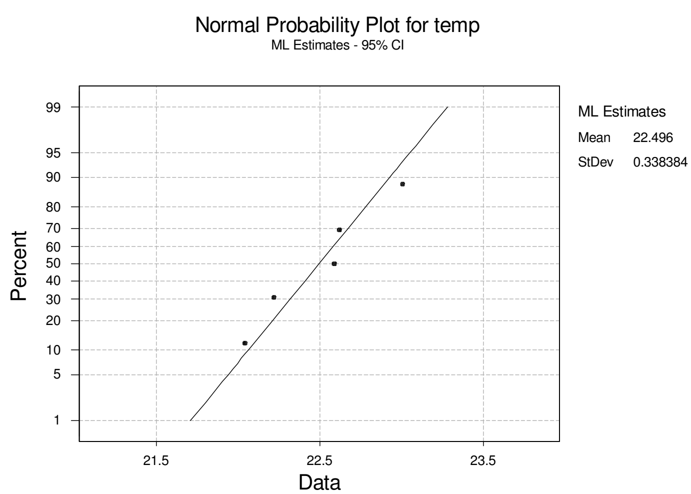
7) $\bar{x} = 22.496$, $s = 0.378$ n=5

$$t_0 = \frac{22.496 - 22.5}{0.378/\sqrt{5}} = -0.00237$$

8) Since –0.00237 >- 2.776, we cannot reject the null hypothesis. There is not sufficient evidence to conclude that the true mean interior temperature is not equal to 22.5 °C at $\alpha = 0.05$.

2*0.4 <*P-value* < 2* 0.5 ; 0..8 < *P*-value <1.0

b.) The points on the normal probability plot fall along the line. Therefore, there is no evidence to conclude that the interior temperature data is not normally distributed.



c.) $d = \frac{\delta}{\sigma} = \frac{|\mu - \mu_0|}{\sigma} = \frac{|22.75 - 22.5|}{0.378} = 0.66$

Using the OC curve, Chart VI e) for $\alpha = 0.05$, $d = 0.66$, and n = 5, we get $\beta \cong 0.8$ and power of 1–0.8 = 0.2.

d) $d = \frac{\delta}{\sigma} = \frac{|\mu - \mu_0|}{\sigma} = \frac{|22.75 - 22.5|}{0.378} = 0.66$

Using the OC curve, Chart VI e) for $\alpha = 0.05$, $d = 0.66$, and $\beta \cong 0.1$ (Power=0.9),
$n = 40$.

e) 95% two sided confidence interval

$$\bar{x} - t_{0.025,4}\left(\frac{s}{\sqrt{n}}\right) \le \mu \le \bar{x} + t_{0.025,4}\left(\frac{s}{\sqrt{n}}\right)$$

$$22.496 - 2.776\left(\frac{0.378}{\sqrt{5}}\right) \le \mu \le 22.496 + 2.776\left(\frac{0.378}{\sqrt{5}}\right)$$

$$22.027 \le \mu \le 22.965$$

We cannot conclude that the mean interior temperature is not equal to 22.5 since the value is included inside the confidence interval.

9-31 a. 1) The parameter of interest is the true mean female body temperature, μ.
2) $H_0 : \mu = 98.6$
3) $H_1 : \mu \neq 98.6$
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 e) for $\alpha = 0.05$, d = 0.83, and $\beta \cong 0.1$ (Power=0.9),
$n = 20$.

d) 95% two sided confidence interval

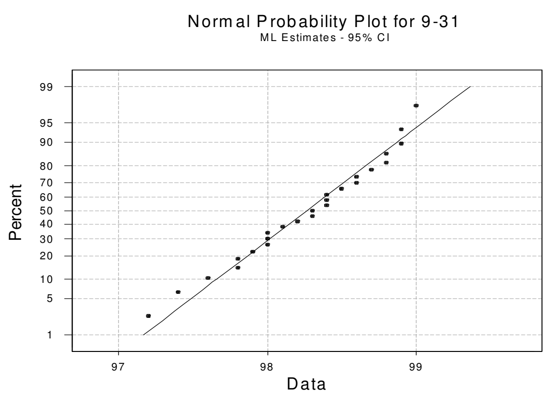
$$\bar{x} - t_{0.025,24}\left(\frac{s}{\sqrt{n}}\right) \le \mu \le \bar{x} + t_{0.025,24}\left(\frac{s}{\sqrt{n}}\right)$$

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

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

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.

9-32 a.) 1) The parameter of interest is the true mean rainfall, μ.

2) $H_0 : \mu = 25$

3) $H_1 : \mu > 25$

4) $\alpha = 0.01$

5) $t_0 = \dfrac{\bar{x} - \mu}{s/\sqrt{n}}$

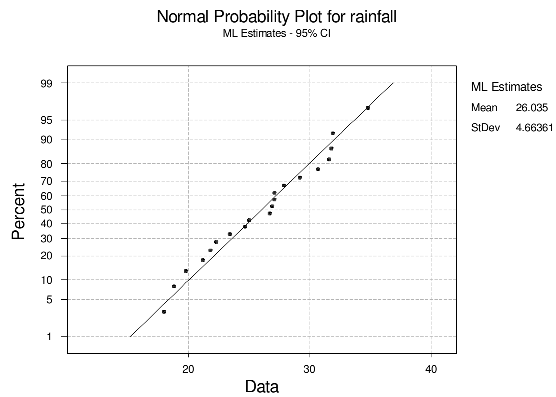
6) Reject H_0 if $t_0 > t_{\alpha,n-1}$ where $t_{0.01,19} = 2.539$

7) $\bar{x} = 26.04$ $s = 4.78$ $n = 20$

$$t_0 = \frac{26.04 - 25}{4.78/\sqrt{20}} = 0.97$$

8) Since $0.97 < 2.539$, do not reject the null hypothesis and conclude there is insufficient evidence to indicate that the true mean rainfall is greater than 25 acre-feet at $\alpha = 0.01$. The $0.10 < P\text{-value} < 0.25$.

b.) the data on the normal probability plot fall along the straight line. Therefore there is evidence that the data are normally distributed.



c.) $d = \dfrac{\delta}{\sigma} = \dfrac{|\mu - \mu_0|}{\sigma} = \dfrac{|27 - 25|}{4.78} = 0.42$

Using the OC curve, Chart VI h) for $\alpha = 0.01$, d = 0.42, and n = 20, we get $\beta \cong 0.7$ and power of 1−0.7 = 0.3.

d) $d = \dfrac{\delta}{\sigma} = \dfrac{|\mu - \mu_0|}{\sigma} = \dfrac{|27.5 - 25|}{4.78} = 0.52$

Using the OC curve, Chart VI h) for $\alpha = 0.05$, d = 0.42, and $\beta \cong 0.1$ (Power=0.9),

$n = 75$.

e) 95% two sided confidence interval

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

$$26.03 - 2.776\left(\frac{4.78}{\sqrt{20}}\right) \le \mu$$

$$23.06 \le \mu$$

Since the lower limit of the CI is less than 25, we conclude that there is insufficient evidence to indicate that the true mean rainfall is not greater than 25 acre-feet at α=0.01.

9-33 a. 1) The parameter of interest is the true mean sodium content, μ.

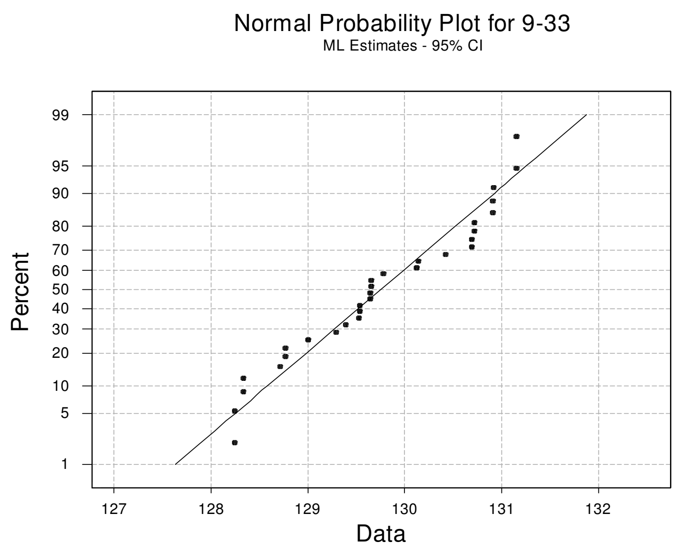
2) $H_0 : \mu = 130$
3) $H_1 : \mu \neq 130$
4) $\alpha = 0.05$
5) $t_0 = \dfrac{\overline{x} - \mu}{s / \sqrt{n}}$
6) Reject H_0 if $|t_0| > t_{\alpha/2,n-1}$ where $t_{\alpha/2,n-1} = 2.045$
7) $\overline{x} = 129.753$, s = 0.929 n=30

$$t_0 = \frac{129.753 - 130}{0.929\sqrt{30}} = -1.456$$

8) Since 1.456 < 2.064, do not reject the null hypothesis and conclude that the there is not sufficient evidence that the true mean sodium content is different from 130mg at $\alpha = 0.05$.

From table IV the t0 value is found between the values of 0.05 and 0.1 with 29 degrees of freedom, so 2*0.05<*P-value* = 2* 0.1 Therefore, 0.1< P-value<0.2.

b)



The assumption of normality appears to be reasonable.

c) $d = \dfrac{\delta}{\sigma} = \dfrac{|\mu - \mu_0|}{\sigma} = \dfrac{|130.5 - 130|}{0.929} = 0.538$

Using the OC curve, Chart VI e) for $\alpha = 0.05$, d = 0.53, and n = 30, we get $\beta \cong 0.2$ and power of 1−0.20 = 0.80.

d) $d = \frac{\delta}{\sigma} = \frac{|\mu - \mu_0|}{\sigma} = \frac{|130.1 - 130|}{0.929} = 0.11$

Using the OC curve, Chart VI e) for $\alpha = 0.05$, d = 0.11, and $\beta \cong 0.25$ (Power=0.75),
$n = 100$.

d) 95% two sided confidence interval

$$\bar{x} - t_{0.025,29}\left(\frac{s}{\sqrt{n}}\right) \le \mu \le \bar{x} + t_{0.025,29}\left(\frac{s}{\sqrt{n}}\right)$$

$$129.753 - 2.045\left(\frac{0.929}{\sqrt{30}}\right) \le \mu \le 129.753 + 2.045\left(\frac{0.929}{\sqrt{30}}\right)$$

$$129.406 \le \mu \le 130.100$$

We can conclude that the mean sodium content is equal to 130 because that value is inside the confidence interval.

9-34 a.)1) The parameter of interest is the true mean tire life, μ.

2) $H_0 : \mu = 60000$

3) $H_1 : \mu > 60000$

4) $\alpha = 0.05$

5) $t_0 = \frac{\bar{x} - \mu}{s / \sqrt{n}}$

6) Reject H_0 if $t_0 > t_{\alpha,n-1}$ where $t_{0.05,15} = 1.753$

7) $n = 16 \quad \bar{x} = 60{,}139.7 \quad s = 3645.94$

$$t_0 = \frac{60139.7 - 60000}{3645.94/\sqrt{16}} = 0.15$$

8) Since 0.15 < 1.753., do not reject the null hypothesis and conclude there is insufficient evidence to indicate that the true mean tire life is greater than 60,000 kilometers at $\alpha = 0.05$. The P-value > 0.40.

b.) $d = \frac{\delta}{\sigma} = \frac{|\mu - \mu_0|}{\sigma} = \frac{|61000 - 60000|}{3645.94} = 0.27$

Using the OC curve, Chart VI g) for $\alpha = 0.05$, d = 0.27, and $\beta \cong 0.1$ (Power=0.9),
$n = 4$.

Yes, the sample size of 16 was sufficient.

9-35. 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 impact strength, μ.

2) $H_0 : \mu = 1.0$

3) $H_1 : \mu > 1.0$

4) $\alpha = 0.05$

5) $t_0 = \frac{\bar{x} - \mu}{s / \sqrt{n}}$

6) Reject H_0 if $t_0 > t_{\alpha,n-1}$ where $t_{0.05,19} = 1.729$

7) $\bar{x} = 1.25 \quad s = 0.25 \quad n = 20$

$$t_0 = \frac{1.25 - 1.0}{0.25/\sqrt{20}} = 4.47$$

8) Since 4.47 > 1.729, reject the null hypothesis and conclude there is sufficient evidence to indicate that the true mean impact strength is greater than 1.0 ft-lb/in at $\alpha = 0.05$. The P-value < 0.0005
.

9-36. In order to use t statistic in hypothesis testing, we need to assume that the underlying distribution is normal.

1) The parameter of interest is the true mean current, μ.
2) $H_0 : \mu = 300$
3) $H_1 : \mu > 300$
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 $t_{0.05,9} = 1.833$
7) $n = 10 \quad \bar{x} = 317.2 \quad s = 15.7$

$$t_0 = \frac{317.2 - 300}{15.7/\sqrt{10}} = 3.46$$

8) Since 3.46 > 1.833, reject the null hypothesis and conclude there is sufficient evidence to indicate that the true mean current is greater than 300 microamps at $\alpha = 0.05$. The $0.0025 < P$-value < 0.005

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

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

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), $n = 40$.

9-38 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 oxygen concentration, μ.

2) $H_0 : \mu = 4$

3) $H_1 : \mu \neq 4$

4) $\alpha = 0.01$

5) $t_0 = \dfrac{\bar{x} - \mu}{s / \sqrt{n}}$

6) Reject H_0 if $|t_0| > t_{a/2, n-1} = t_{0.005, 19} = 2.861$

7) $\bar{x} = 3.265$, s = 2.127, n = 20

$$t_0 = \frac{3.265 - 4}{2.127 / \sqrt{20}} = -1.55$$

8) Since -2.861<-1.55 <1.48, do not reject the null hypothesis and conclude that there is insufficient evidence to indicate that the true mean oxygen not equal 4 at $\alpha = 0.01$.

b.) The *P*-Value: 2*0.05<*P*-value<2*0.10 therefore 0.10< *P*-value<0.20

c.) $d = \dfrac{\delta}{\sigma} = \dfrac{|\mu - \mu_0|}{\sigma} = \dfrac{|3 - 4|}{2.127} = 0.47$

Using the OC curve, Chart VI f) for $\alpha = 0.01$, d = 0.47, and n = 20, we get $\beta \cong 0.70$ and power of 1−0.70 = 0.30.

d.) $d = \dfrac{\delta}{\sigma} = \dfrac{|\mu - \mu_0|}{\sigma} = \dfrac{|2.5 - 4|}{2.127} = 0.71$

Using the OC curve, Chart VI f) for $\alpha = 0.01$, d = 0.71, and $\beta \cong 0.10$ (Power=0.90), $n = 40$.

9-39 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 cigar tar content, μ.

2) $H_0 : \mu = 1.5$

3) $H_1 : \mu > 1.5$

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 $t_{0.05,29} = 1.699$

7) $\bar{x} = 1.529$ s = 0.0566 n = 30

$$t_0 = \frac{1.529 - 1.5}{0.0566 / \sqrt{30}} = 2.806$$

8) Since 2.806 > 1.699, reject the null hypothesis and conclude that there is sufficient evidence to indicate that the true mean tar content is greater than 1.5 at $\alpha = 0.05$.

b.) From table IV the t_0 value is found between the values of 0.0025 and 0.005 with 29 degrees of freedom. Therefore, 0.0025< P-value<0.005.

Minitab gives *P-value* = 0.004.

c) $d = \dfrac{\delta}{\sigma} = \dfrac{|\mu - \mu_0|}{\sigma} = \dfrac{|1.6 - 1.5|}{0.0566} = 1.77$

Using the OC curve, Chart VI g) for $\alpha = 0.05$, d = 1.77, and n = 30, we get $\beta \cong 0$ and power of 1−0 = 1.

e.) $d = \dfrac{\delta}{\sigma} = \dfrac{|\mu - \mu_0|}{\sigma} = \dfrac{|1.6 - 1.5|}{0.0566} = 1.77$

Using the OC curve, Chart VI g) for $\alpha = 0.05$, d = 1.77, and $\beta \cong 0.20$ (Power=0.80), $n = 4$.

9-40 a) 1) The parameter of interest is the true mean height of female engineering students, μ.

2) $H_0 : \mu = 65$

3) $H_1 : \mu \neq 65$

4) $\alpha = 0.05$

5) $t_0 = \dfrac{\bar{x} - \mu}{s/\sqrt{n}}$

6) Reject H_0 if $|t_0| > t_{\alpha/2,n-1}$ where $t_{0.025,59} = 2.0281$

7) $\bar{x} = 65.811$ inches $s = 2.106$ inches n = 37

$$t_0 = \frac{65.811 - 65}{2.11/\sqrt{37}} = 2.34$$

8) Since 2.34 > 2.0281, reject the null hypothesis and conclude that there is sufficient evidence to indicate that the true mean height of female engineering students is not equal to 65 at $\alpha = 0.05$.

b.)P-value: $0.02 < P\text{-value} < 0.05$.

c) $d = \dfrac{|62 - 65|}{2.11} = 1.42$, n=37 so, from the OC Chart VI e) for $\alpha = 0.05$, we find that $\beta \cong 0$. Therefore, the power $\cong 1$.

d.) $d = \dfrac{|64 - 65|}{2.11} = 0.47$ for $\alpha = 0.05$, and $\beta \cong 0.2$ (Power=0.8).

$n^* = 40$.

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) $t_0 = \dfrac{\bar{x} - \mu}{s/\sqrt{n}}$

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), $n = 75$.

9-42 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 distance, μ.
2) $H_0 : \mu = 280$
3) $H_1 : \mu > 280$
4) $\alpha = 0.05$
5) $t_0 = \frac{\bar{x}-\mu}{s/\sqrt{n}}$
6) Reject H_0 if $t_0 > t_{\alpha,n-1}$ where $t_{0.05,99} = 1.6604$
7) $\bar{x} = 260.3$ s = 13.41 n = 100

$$t_0 = \frac{260.3-280}{13.41/\sqrt{100}} = -14.69$$

8) Since –14.69 < 1.6604, do not reject the null hypothesis and conclude that there is insufficient evidence to indicate that the true mean distance is greater than 280 at $\alpha = 0.05$.

b.) From table IV the t_0 value is found above the value 0.005, therefore, the P-value is greater than 0.995.

c) $d = \frac{\delta}{\sigma} = \frac{|\mu-\mu_0|}{\sigma} = \frac{|290-280|}{13.41} = 0.75$

Using the OC curve, Chart VI g) for $\alpha = 0.05$, d = 0.75, and n = 100, we get $\beta \cong 0$ and power of 1–0 = 1.

f.) $d = \frac{\delta}{\sigma} = \frac{|\mu-\mu_0|}{\sigma} = \frac{|290-280|}{13.41} = 0.75$

Using the OC curve, Chart VI g) for $\alpha = 0.05$, d = 0.75, and $\beta \cong 0.20$ (Power=0.80), $n = 15$.

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) $\chi_0^2 = \dfrac{(n-1)s^2}{\sigma^2}$

6) Reject H_0 if $\chi_0^2 > \chi_{\alpha,n-1}^2$ where $\chi_{0.01,14}^2 = 29.14$

7) $n = 15$, $s^2 = 0.008$

$$\chi_0^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{14(0.008)^2}{0.0001} = 8.96$$

8) Since $8.96 < 29.14$ do not reject H_0 and conclude there is insufficient evidence to indicate the true standard deviation of the diameter exceeds 0.01 at $\alpha = 0.01$.

b) P-value = $P(\chi^2 > 8.96)$ for 14 degrees of freedom: $0.5 < \text{P-value} < 0.9$

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

using chart VIk the required sample size is 50

9-44 a.) In order to use χ^2 statistic in hypothesis testing and confidence interval construction, we need to assume that the underlying distribution is normal.

1) The parameter of interest is the true variance of sugar content, σ^2. However, the answer can be found by performing a hypothesis test on σ^2.

2) $H_0: \sigma^2 = 18$

3) $H_1: \sigma^2 \neq 18$

4) $\alpha = 0.05$

5) $\chi_0^2 = \dfrac{(n-1)s^2}{\sigma^2}$

6) Reject H_0 if $\chi_0^2 < \chi_{1-\alpha/2,n-1}^2$ where $\chi_{0.975,9}^2 = 2.70$ or $\chi_0^2 > \chi_{\alpha,2,n-1}^2$ where $\chi_{0.025,9}^2 = 19.02$

7) $n = 10$, $s = 4.8$

$$\chi_0^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{9(4.8)^2}{18} = 11.52$$

8) Since $11.52 < 19.02$ do not reject H_0 and conclude there is insufficient evidence to indicate the true variance of sugar content is significantly different from 18 at $\alpha = 0.01$.

b.) *P*-value: The χ_0^2 is between 0.50 and 0.10. Therefore, $0.2 < P\text{-value} < 1$

c.) The 95% confidence interval includes the value 18, therefore, we could not be able to conclude that the variance was not equal to 18.

$$\frac{9(4.8)^2}{19.02} \le \sigma^2 \le \frac{9(4.8)^2}{2.70}$$

$$10.90 \le \sigma^2 \le 76.80$$

9-45 a) In order to use χ^2 statistic in hypothesis testing and confidence interval construction, we need to assume that the underlying distribution is normal.

1) The parameter of interest is the standard deviation of tire life, σ. However, the answer can be found by performing a hypothesis test on σ^2.
2) $H_0: \sigma^2 = 40,000$
3) $H_1: \sigma^2 > 40,000$
4) $\alpha = 0.05$
5) $\chi_0^2 = \dfrac{(n-1)s^2}{\sigma^2}$
6) Reject H_0 if $\chi_0^2 > \chi_{\alpha,n-1}^2$ where $\chi_{0.05,15}^2 = 25.00$
7) $n = 16$, $s^2 = (3645.94)^2$

$$\chi_0^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{15(3645.94)^2}{40000} = 4984.83$$

8) Since 4984.83 > 25.00 reject H_0 and conclude there is strong evidence to indicate the true standard deviation of tire life exceeds 200 km at $\alpha = 0.05$.

b) P-value = $P(\chi^2 > 4984.83)$ for 15 degrees of freedom P-value < 0.005

9-46 a) In order to use χ^2 statistic in hypothesis testing and confidence interval construction, we need to assume that the underlying distribution is normal.

1) The parameter of interest is the true standard deviation of Izod impact strength, σ. However, the answer can be found by performing a hypothesis test on σ^2.
2) $H_0: \sigma^2 = (0.10)^2$
3) $H_1: \sigma^2 \neq (0.10)^2$
4) $\alpha = 0.01$
5) $\chi_0^2 = \dfrac{(n-1)s^2}{\sigma^2}$
6) Reject H_0 if $\chi_0^2 < \chi_{1-\alpha/2,n-1}^2$ where $\chi_{0.995,19}^2 = 6.84$ 27 or $\chi_0^2 > \chi_{\alpha,2,n-1}^2$ where $\chi_{0.005,19}^2 = 38.58$
7) $n = 20$, $s = 0.25$

$$\chi_0^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{19(0.25)^2}{(0.10)^2} = 118.75$$

8) Since 118.75 > 38.58 reject H_0 and conclude there is sufficient evidence to indicate the true standard deviation of Izod impact strength is significantly different from 0.10 at $\alpha = 0.01$.

b.) *P*-value: The P-value<0.005

c.) 99% confidence interval for σ:

First find the confidence interval for σ^2 :

For $\alpha = 0.01$ and $n = 20$, $\chi_{\alpha/2,n-1}^2 = \chi_{0.995,19}^2 = 6.84$ and $\chi_{1-\alpha/2,n-1}^2 = \chi_{0.005,19}^2 = 38.58$

$$\frac{19(0.25)^2}{38.58} \le \sigma^2 \le \frac{19(0.25)^2}{6.84}$$

$$0.03078 \le \sigma^2 \le 0.1736$$

Since 0.01 falls below the lower confidence bound we would conclude that the population standard deviation is not equal to 0.01.

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 $\chi^2_{1-\alpha/2,n-1} = \chi^2_{0.975,50} = 32.36$

$$\frac{50(0.37)^2}{71.42} \leq \sigma^2 \leq \frac{50(0.37)^2}{32.36}$$

$$0.096 \leq \sigma^2 \leq 0.2115$$

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.

9-48 Using the chart in the Appendix, with $\lambda = \sqrt{\dfrac{0.012}{0.008}} = 1.22$ and n = 15 we find $\beta = 0.80$.

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

9-50 a) The parameter of interest is the true proportion of engine crankshaft bearings exhibiting surface roughness.

2) $H_0 : p = 0.10$

3) $H_1 : p > 0.10$

4) $\alpha = 0.05$

5) $z_0 = \dfrac{x - np_0}{\sqrt{np_0(1-p_0)}}$ or $z_0 = \dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}}$; Either approach will yield the same conclusion

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

7) $x = 10$ $n = 85$ $\hat{p} = \dfrac{13}{300} = 0.043$

$$z_0 = \frac{x - np_0}{\sqrt{np_0(1-p_0)}} = \frac{10 - 85(0.10)}{\sqrt{85(0.10)(0.90)}} = 0.54$$

8) Since 0.53 < 1.65, do not reject the null hypothesis and conclude the true proportion of crankshaft bearings exhibiting surface roughness is not significantly greater than 0.10, at $\alpha = 0.05$.

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

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

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

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

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

$$= \left(\frac{1.96\sqrt{0.10(1-0.10)} - 1.28\sqrt{0.15(1-0.15)}}{0.15 - 0.10}\right)^2$$

$$= (10.85)^2 = 117.63 \cong 118$$

9-52 a) Using the information from Exercise 8-48, test

1) The parameter of interest is the true fraction defective integrated circuits
2) $H_0 : p = 0.05$
3) $H_1 : p \neq 0.05$
4) $\alpha = 0.05$
5) $z_0 = \dfrac{x - np_0}{\sqrt{np_0(1-p_0)}}$ or $z_0 = \dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}}$; Either approach will yield the same conclusion
6) Reject H_0 if $z_0 < -z_{\alpha/2}$ where $-z_{\alpha/2} = -z_{0.025} = -1.96$ or $z_0 > z_{\alpha/2}$ where $z_{\alpha/2} = z_{0.025} = 1.96$
7) $x = 13$ $n = 300$ $\hat{p} = \dfrac{13}{300} = 0.043$

$$z_0 = \frac{x - np_0}{\sqrt{np_0(1-p_0)}} = \frac{13 - 300(0.05)}{\sqrt{300(0.05)(0.95)}} = -0.53$$

8) Since $-0.53 > -1.65$, do not reject null hypothesis and conclude the true fraction of defective integrated circuits is not significantly less than 0.05, at $\alpha = 0.05$.

b.) The P-value: $2(1\text{-}\Phi(0.53))=2(1\text{-}0.70194)=0.59612$

9-53. a) Using the information from Exercise 8-48, test

1) The parameter of interest is the true fraction defective integrated circuits
2) $H_0 : p = 0.05$
3) $H_1 : p < 0.05$
4) $\alpha = 0.05$
5) $z_0 = \dfrac{x - np_0}{\sqrt{np_0(1-p_0)}}$ or $z_0 = \dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}}$; Either approach will yield the same conclusion
6) Reject H_0 if $z_0 < -z_{\alpha}$ where $-z_{\alpha} = -z_{0.05} = -1.65$
7) $x = 13$ $n = 300$ $\hat{p} = \dfrac{13}{300} = 0.043$

$$z_0 = \frac{x - np_0}{\sqrt{np_0(1-p_0)}} = \frac{13 - 300(0.05)}{\sqrt{300(0.05)(0.95)}} = -0.53$$

8) Since $-0.53 > -1.65$, 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 = $1 - \Phi(0.53) = 0.29806$

9-54 a) 1) The parameter of interest is the true proportion of engineering students planning graduate studies
2) $H_0 : p = 0.50$
3) $H_1 : p \neq 0.50$
4) $\alpha = 0.05$
5) $z_0 = \dfrac{x - np_0}{\sqrt{np_0(1-p_0)}}$ or $z_0 = \dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}}$; Either approach will yield the same conclusion
6) Reject H_0 if $z_0 < -z_{\alpha/2}$ where $-z_{\alpha/2} = -z_{0.025} = -1.96$ or $z_0 > z_{\alpha/2}$ where $z_{\alpha/2} = z_{0.025} = 1.96$
7) $x = 117$ $n = 484$ $\hat{p} = \dfrac{13}{300} = 0.043$

$$z_0 = \frac{x - np_0}{\sqrt{np_0(1-p_0)}} = \frac{117 - 484(0.5)}{\sqrt{484(0.5)(0.5)}} = -11.36$$

8) Since $-11.36 > -1.65$, reject the null hypothesis and conclude the true proportion of engineering students planning graduate studies is significantly different from 0.5, at $\alpha = 0.05$.

b.) P-value $=2[1 - \Phi(11.36)] \cong 0$

c.) $\hat{p} = \dfrac{117}{484} = 0.242$

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

$$0.242 - 1.96\sqrt{\frac{0.242(0.758)}{484}} \le p \le 0.242 - 1.96\sqrt{\frac{0.242(0.758)}{484}}$$

$$0.204 \le p \le 0.280$$

Since the 95% confidence interval does not contain the value 0.5, then conclude that the true proportion of engineering students planning graduate studies is significantly different from 0.5.

9-55. a) 1) The parameter of interest is the true percentage of polished lenses that contain surface defects, p.
2) $H_0 : p = 0.02$
3) $H_1 : p < 0.02$
4) $\alpha = 0.05$
5) $z_0 = \dfrac{x - np_0}{\sqrt{np_0(1-p_0)}}$ or $z_0 = \dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}}$; Either approach will yield the same conclusion
6) Reject H_0 if $z_0 < -z_\alpha$ where $-z_\alpha = -z_{0.05} = -1.65$
7) $x = 6$ $n = 250$ $\hat{p} = \dfrac{6}{250} = 0.024$

$$z_0 = \frac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}} = \frac{0.024 - 0.02}{\sqrt{\dfrac{0.02(1-0.02)}{250}}} = 0.452$$

8) Since $0.452 > -1.65$ do not reject the null hypothesis and conclude the machine cannot be qualified at the 0.05 level of significance.

b) P-value $= \Phi(0.452) = 0.67364$

9-56 . a) 1) The parameter of interest is the true percentage of football helmets that contain flaws, p.

2) $H_0 : p = 0.1$

3) $H_1 : p > 0.1$

4) $\alpha = 0.01$

5) $z_0 = \dfrac{x - np_0}{\sqrt{np_0(1-p_0)}}$ or $z_0 = \dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}}$; Either approach will yield the same conclusion

6) Reject H_0 if $z_0 > z_\alpha$ where $z_\alpha = z_{0.01} = 2.33$

7) x = 16 n = 200 $\hat{p} = \dfrac{16}{200} = 0.08$

$$z_0 = \frac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}} = \frac{0.08 - 0.10}{\sqrt{\dfrac{0.10(1-0.10)}{200}}} = -0.94$$

8) Since -0.452 < 2.33 do not reject the null hypothesis and conclude the proportion of football helmets with flaws does not exceed 10%.

b) P-value = $1-\Phi(0.94) = 1-.8264 = 0.67364$

9-57. The problem statement implies that H_0: p = 0.6, H_1: p > 0.6 and defines an acceptance region as $\hat{p} \le \dfrac{315}{500} = 0.63$ and rejection region as $\hat{p} > 0.63$

a) The probability of a type 1 error is

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

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

9-58 1) The parameter of interest is the true proportion of batteries that fail before 48 hours, p.

2) $H_0 : p = 0.002$

3) $H_1 : p < 0.002$

4) $\alpha = 0.01$

5) $z_0 = \dfrac{x - np_0}{\sqrt{np_0(1-p_0)}}$ or $z_0 = \dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}}$; Either approach will yield the same conclusion

6) Reject H_0 if $z_0 < -z_\alpha$ where $-z_\alpha = -z_{0.01} = -2.33$

7) x = 15 n = 5000 $\hat{p} = \dfrac{15}{5000} = 0.003$

$$z_0 = \frac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}} = \frac{0.003 - 0.002}{\sqrt{\dfrac{0.002(1-0.998)}{5000}}} = 1.58$$

8) Since 1.58 > -2.33 do not reject the null hypothesis and conclude the proportion of proportion of cell phone batteries that fail is not less than 0.2% at α=0.01.

Section 9-7

9-59. Expected Frequency is found by using the Poisson distribution

$$P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!} \text{ where } \lambda = [0(24) + 1(30) + 2(31) + 3(11) + 4(4)]/100 = 1.41$$

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)

9-60. Expected Frequency is found by using the Poisson distribution

$$P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!} \text{ where } \lambda = [1(1) + 2(11) + \ldots + 7(10) + 8(9)]/75 = 4.907$$

Estimated mean = 4.907

Value	1	2	3	4	5	6	7	8
Observed Frequency	1	11	8	13	11	12	10	9
Expected Frequency	2.7214	6.6770	10.9213	13.3977	13.1485	10.7533	7.5381	4.6237

Since the first category has an expected frequency less than 3, combine it with the next category:

Value	1-2	3	4	5	6	7	8
Observed Frequency	12	8	13	11	12	10	9
Expected Frequency	9.3984	10.9213	13.3977	13.1485	10.7533	7.5381	4.6237

The degrees of freedom are k − p − 1 = 7 − 1 − 1 = 5

a) 1) The variable of interest is the form of the distribution for the number of flaws.
2) H_0: The form of the distribution is Poisson
3) H_1: The form of the distribution is not Poisson
4) $\alpha = 0.01$
5) The test statistic is

$$\chi_0^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$$

6) Reject H_0 if $\chi_o^2 > \chi_{0.01,5}^2 = 15.09$
7)

$$\chi_0^2 = \frac{(12 - 9.3984)^2}{9.3984} + \cdots + \frac{(9 - 4.6237)^2}{4.6237} = 6.955$$

8) Since 6.955 < 15.09 do not reject H_0. We are unable to reject the null hypothesis that the distribution of the number of flaws is Poisson.

b) P-value = 0.2237 (found using Minitab)

9-61. Estimated mean = 10.131

Value	5	6	8	9	10	11	12	13	14	15
Rel. Freq	0.067	0.067	0.100	0.133	0.200	0.133	0.133	0.067	0.033	0.067
Observed (Days)	2	2	3	4	6	4	4	2	1	2
Expected (Days)	1.0626	1.7942	3.2884	3.7016	3.7501	3.4538	2.9159	2.2724	1.6444	1.1106

Since there are several cells with expected frequencies less than 3, the revised table would be:

Value	5-8	9	10	11	12-15
Observed (Days)	7	4	6	4	9
Expected (Days)	6.1452	3.7016	3.7501	3.4538	7.9433

The degrees of freedom are k − p − 1 = 5 − 1 − 1 = 3

a) 1) The variable of interest is the form of the distribution for the number of calls arriving to a switchboard from noon to 1pm during business days.
2) H_0: The form of the distribution is Poisson
3) H_1: The form of the distribution is not Poisson
4) $\alpha = 0.05$
5) The test statistic is

$$\chi_0^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$$

6) Reject H_0 if $\chi_o^2 > \chi_{0.05,3}^2 = 7.81$
7)

$$\chi_0^2 = \frac{(7-6.1452)^2}{6.1452} + \frac{(4-3.7016)^2}{3.7016} + \frac{(6-3.7501)^2}{3.7501} + \frac{(4-3.4538)^2}{3.4538} + \frac{(9-7.9433)^2}{7.9433} = 1.72$$

8) Since 1.72 < 7.81 do not reject H_0. We are unable to reject the null hypothesis that the distribution for the number of calls is Poisson.

b) The P-value is between 0.9 and 0.5 using Table III. P-value = 0.6325 (found using Minitab)

9-62 Use the binomial distribution to get the expected frequencies with the mean = np = 6(0.25) = 1.5

Value	0	1	2	3	4
Observed	4	21	10	13	2
Expected	8.8989	17.7979	14.8315	6.5918	1.6479

The expected frequency for value 4 is less than 3. Combine this cell with value 3:

Value	0	1	2	3-4
Observed	4	21	10	15
Expected	8.8989	17.7979	14.8315	8.2397

The degrees of freedom are $k - p - 1 = 4 - 0 - 1 = 3$

a) 1) The variable of interest is the form of the distribution for the random variable X.
2) H_0: The form of the distribution is binomial with n = 6 and p = 0.25
3) H_1: The form of the distribution is not binomial with n = 6 and p = 0.25
4) $\alpha = 0.05$
5) The test statistic is

$$\chi_0^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$$

6) Reject H_0 if $\chi_o^2 > \chi_{0.05,3}^2 = 7.81$
7)

$$\chi_0^2 = \frac{(4 - 8.8989)^2}{8.8989} + \cdots + \frac{(15 - 8.2397)^2}{8.2397} = 10.39$$

8) Since 10.39 > 7.81 reject H_0. We can conclude that the distribution is not binomial with n = 6 and p = 0.25 at $\alpha = 0.05$.

b) P-value = 0.0155 (found using Minitab)

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 underfilled cartons, X.

2) H_0: The form of the distribution is binomial

3) H_1: The form of the distribution is not binomial

4) $\alpha = 0.05$

5) The test statistic is

$$\chi_0^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$$

6) Reject H_0 if $\chi_o^2 > \chi_{0.05,1}^2 = 3.84$

7) $$\chi_0^2 = \frac{(39-38.1426)^2}{38.1426} + \frac{(23-26.1571)^2}{26.1571} + \frac{(13-10.3962)^2}{10.39} = 1.053$$

8) Since 1.053 < 3.84 do not reject H_0. We are unable to reject the null hypothesis that the distribution of the number of underfilled cartons is binomial at $\alpha = 0.05$.

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

9-64 Estimated mean = 49.6741 use Poisson distribution with λ=49.674

All expected frequencies are greater than 3.

The degrees of freedom are k − p − 1 = 26 − 1 − 1 = 24

a) 1) The variable of interest is the form of the distribution for the number of cars passing through the intersection.

2) H_0: The form of the distribution is Poisson

3) H_1: The form of the distribution is not Poisson

4) $\alpha = 0.05$

5) The test statistic is

$$\chi_0^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$$

6) Reject H_0 if $\chi_o^2 > \chi_{0.05,24}^2 = 36.42$

7) Estimated mean = 49.6741

$$\chi_0^2 = 769.57$$

8) Since 769.57 >>> 36.42, reject H_0. We can conclude that the distribution is not Poisson at $\alpha = 0.05$.

b) P-value = 0 (found using Minitab)

Section 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)

9-66 1. The variable of interest is calls by surgical-medical patients.
2. H_0:Calls by surgical-medical patients are independent of Medicare status.
3. H_1:Calls by surgical-medical patients are not independent of Medicare status.
4. $\alpha = 0.01$
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 $\chi^2_{.01,1} = 6.637$
7. The calculated test statistic is $\chi^2_0 = 0.033$
8. $\chi^2_0 \not> \chi^2_{0.01,1}$, do not reject H_0 and conclude that the evidence is not sufficient to claim that surgical-medical patients and Medicare status are dependent. *P*-value = 0.85

9-67. 1. The variable of interest is statistics grades and OR grades.
2. H_0: Statistics grades are independent of OR grades.
3. H_1: Statistics and OR grades are not independent.
4. $\alpha = 0.01$
5. The test statistic is:

$$\chi_0^2 = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{(O_{ij}-E_{ij})^2}{E_{ij}}$$

6. The critical value is $\chi^2_{.01,9} = 21.665$
7. The calculated critical value is $\chi^2_0 = 25.55$
8. $\chi^2_0 > \chi^2_{0.01,9}$ Therefore, reject H_0 and conclude that the grades are not independent at $\alpha = 0.01$.
P-value = 0.002

9-68 1. The variable of interest is characteristic among deflections and ranges.

2. H_0: Deflection and range are independent.
3. H_1: Deflection and range are not independent.
4. $\alpha = 0.05$
5. The test statistic is:

$$\chi_0^2 = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{(O_{ij}-E_{ij})^2}{E_{ij}}$$

6. The critical value is $\chi^2_{0.05,4} = 9.488$

7. The calculated test statistic is $\chi^2_0 = 2.46$

8. $\chi_0^2 \not> \chi^2_{0.05,4}$, do not reject H_0 and conclude that the evidence is not sufficient to claim that the data are not independent at $\alpha = 0.05$. *P*-value = 0.652

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 $\chi^2_{0.01,3} = 11.344$

7. The calculated test statistic is $\chi^2_0 = 10.71$

8. $\chi_0^2 \not> \chi^2_{0.01,3}$, 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

9-70 1. The variable of interest is opinion on core curriculum change.
2. H_0: Opinion of the change is independent of the class standing.
3. H_1: Opinion of the change is not independent of the class standing.
4. $\alpha = 0.05$
5. The test statistic is:

$$\chi_0^2 = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{(O_{ij}-E_{ij})^2}{E_{ij}}$$

6. The critical value is $\chi^2_{0.05,3} = 7.815$

7. The calculated test statistic is $\chi^2_{.0} = 26.97$.

8. $\chi_0^2 >>> \chi^2_{0.05,3}$, reject H_0 and conclude that the opinions on the change are not independent of class standing. *P*-value ≈ 0

Supplemental Exercises

9-71 Sample Mean = $\hat{p}$ Sample Variance = $\frac{\hat{p}(1-\hat{p})}{n}$

	Sample Size, n	Sampling Distribution	Sample Mean	Sample Variance
a.	50	Normal	p	$\frac{p(1-p)}{50}$
b.	80	Normal	p	$\frac{p(1-p)}{80}$
c.	100	Normal	p	$\frac{p(1-p)}{100}$

d) As the sample size increases, the variance of the sampling distribution decreases.

9-72

	n	Test statistic	P-value	conclusion
a.	50	$z_0 = \frac{0.095-0.10}{\sqrt{0.10(1-0.10)/50}} = -0.12$	0.4522	Do not reject H_0
b.	100	$z_0 = \frac{0.095-0.10}{\sqrt{0.10(1-0.10)/100}} = -0.15$	0.4404	Do not reject H_0
c.	500	$z_0 = \frac{0.095-0.10}{\sqrt{0.10(1-0.10)/500}} = -0.37$	0.3557	Do not reject H_0
d.	1000	$z_0 = \frac{0.095-0.10}{\sqrt{0.10(1-0.10)/1000}} = -0.53$	0.2981	Do not reject H_0

e. The P-value decreases as the sample size increases.

9-73. $\sigma = 12$, $\delta = 205 - 200 = 5$, $\frac{\alpha}{2} = 0.025$, $z_{0.025} = 1.96$,

a) n = 20: $$\beta = \Phi\left(1.96 - \frac{5\sqrt{20}}{12}\right) = \Phi(0.163) = 0.564$$

b) n = 50: $$\beta = \Phi\left(1.96 - \frac{5\sqrt{50}}{12}\right) = \Phi(-0.986) = 1 - \Phi(0.986) = 1 - 0.839 = 0.161$$

c) n = 100: $$\beta = \Phi\left(1.96 - \frac{5\sqrt{100}}{12}\right) = \Phi(-2.207) = 1 - \Phi(2.207) = 1 - 0.9884 = 0.0116$$

d) β, which is the probability of a Type II error, decreases as the sample size increases because the variance of the sample mean decreases. Consequently, the probability of observing a sample mean in the acceptance region centered about the incorrect value of 200 ml/h decreases with larger n.

9-74 $\sigma = 14,\ \delta = 205 - 200 = 5,\ \frac{\alpha}{2} = 0.025,\ z_{0.025} = 1.96,$

a) n = 20: $$\beta = \Phi\left(1.96 - \frac{5\sqrt{20}}{14}\right) = \Phi(0.362) = 0.6406$$

b) n = 50: $$\beta = \Phi\left(1.96 - \frac{5\sqrt{50}}{14}\right) = \Phi(-0.565) = 1 - \Phi(0.565) = 1 - 0.7123 = 0.2877$$

c) n = 100: $$\beta = \Phi\left(1.96 - \frac{5\sqrt{100}}{14}\right) = \Phi(-1.611) = 1 - \Phi(1.611) = 1 - 0.9463 = 0.0537$$

d) The probability of a Type II error increases with an increase in the standard deviation.

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.

9-76 $\alpha=0.01$

a.) n=25 $$\beta = \Phi\left(z_{0.01} + \frac{85-86}{16/\sqrt{25}}\right) = \Phi(2.33 - 0.31) = \Phi(2.02) = 0.9783$$

n=100 $$\beta = \Phi\left(z_{0.01} + \frac{85-86}{16/\sqrt{100}}\right) = \Phi(2.33 - 0.63) = \Phi(1.70) = 0.9554$$

n=400 $$\beta = \Phi\left(z_{0.01} + \frac{85-86}{16/\sqrt{400}}\right) = \Phi(2.33 - 1.25) = \Phi(1.08) = 0.8599$$

n=2500 $$\beta = \Phi\left(z_{0.01} + \frac{85-86}{16/\sqrt{2500}}\right) = \Phi(2.33 - 3.13) = \Phi(-0.80) = 0.2119$$

b.) n=25 $$z_0 = \frac{86-85}{16/\sqrt{25}} = 0.31$$ *P*-value: $1 - \Phi(0.31) = 1 - 0.6217 = 0.3783$

n=100 $z_0 = \dfrac{86-85}{16/\sqrt{100}} = 0.63$ *P*-value: $1-\Phi(0.63) = 1-0.7357 = 0.2643$

n=400 $z_0 = \dfrac{86-85}{16/\sqrt{400}} = 1.25$ *P*-value: $1-\Phi(1.25) = 1-0.8944 = 0.1056$

n=2500 $z_0 = \dfrac{86-85}{16/\sqrt{2500}} = 3.13$ *P*-value: $1-\Phi(3.13) = 1-0.9991 = 0.0009$

The data would be statistically significant when n=2500 at α=0.01

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:

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

6) Since no critical value is given, we will calculate the P-value
7) $\bar{x} = 153.7$, 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 α = 0.01 or 0.05, we would not reject the null hypothesis, thus the claim would not be supported. If we used α = 0.10, we would reject the null in favor of the alternative and conclude the weld strength exceeds 150 psi.

9-78 a.) α=0.05

n=100 $\beta = \Phi\left(z_{0.05} + \dfrac{0.5-0.6}{\sqrt{0.5(0.5)/100}}\right) = \Phi(1.65-2.0) = \Phi(-0.35) = 0.3632$

$$Power = 1-\beta = 1-0.3632 = 0.6368$$

n=150 $\beta = \Phi\left(z_{0.05} + \dfrac{0.5-0.6}{\sqrt{0.5(0.5)/100}}\right) = \Phi(1.65-2.45) = \Phi(-0.8) = 0.2119$

$$Power = 1-\beta = 1-0.2119 = 0.7881$$

n=300 $\beta = \Phi\left(z_{0.05} + \dfrac{0.5-0.6}{\sqrt{0.5(0.5)/300}}\right) = \Phi(1.65-3.46) = \Phi(-1.81) = 0.03515$

$$Power = 1-\beta = 1-0.03515 = 0.96485$$

b.) α=0.01

n=100 $\beta = \Phi\left(z_{0.01} + \frac{0.5 - 0.6}{\sqrt{0.5(0.5)/100}} \right) = \Phi(2.33 - 2.0) = \Phi(0.33) = 0.6293$

$Power = 1 - \beta = 1 - 0.6293 = 0.3707$

n=150 $\beta = \Phi\left(z_{0.01} + \frac{0.5 - 0.6}{\sqrt{0.5(0.5)/100}} \right) = \Phi(2.33 - 2.45) = \Phi(-0.12) = 0.4522$

$Power = 1 - \beta = 1 - 0.4522 = 0.5478$

n=300 $\beta = \Phi\left(z_{0.01} + \frac{0.5 - 0.6}{\sqrt{0.5(0.5)/300}} \right) = \Phi(2.33 - 3.46) = \Phi(-1.13) = 0.1292$

$Power = 1 - \beta = 1 - 0.1292 = 0.8702$

Decreasing the value of α decreases the power of the test for the different sample sizes.

c.) α=0.05

n=100 $\beta = \Phi\left(z_{0.05} + \frac{0.5 - 0.8}{\sqrt{0.5(0.5)/100}} \right) = \Phi(1.65 - 6.0) = \Phi(-4.35) \cong 0.0$

$Power = 1 - \beta = 1 - 0 \cong 1$

The true value of *p* has a large effect on the power. The further *p* is away from p_0 the larger the power of the test.

d.)

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

$$= \left(\frac{2.58\sqrt{0.5(1-0.50)} - 1.65\sqrt{0.6(1-0.6)}}{0.6 - 0.5} \right)^2 = (4.82)^2 = 23.2 \cong 24$$

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

$$= \left(\frac{2.58\sqrt{0.5(1-0.50)} - 1.65\sqrt{0.8(1-0.8)}}{0.8 - 0.5} \right)^2 = (2.1)^2 = 4.41 \cong 5$$

The true value of *p* has a large effect on the power. The further *p* is away from p_0 the smaller the sample size that is required.

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: $\chi_0^2 = \dfrac{(n-1)s^2}{\sigma^2}$

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 = $P\left(\chi^2 < 5.546\right)$; $\quad 0.1 < P - \text{value} < 0.5$

8) The P-value is greater than any acceptable significance level, α, therefore we do not reject the null hypothesis. There is insufficient evidence to support the claim that the standard deviation is less than 20 microamps.

b) 7) n = 51, s = 20

$$\chi_0^2 = \frac{50(15.7)^2}{400} = 30.81$$

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

8) The P-value is less than 0.05, therefore we reject the null hypothesis and conclude that the standard deviation is significantly less than 20 microamps.

c) Increasing the sample size increases the test statistic χ_0^2 and therefore decreases the P-value, providing more evidence against the null hypothesis.

9-80 a) 1) the parameter of interest is the variance of fatty acid measurements, σ^2
2) $H_0 : \sigma^2 = 1.0$
3) $H_1 : \sigma^2 \neq 1.0$
4) α=0.01

5) The test statistic is: $\chi_0^2 = \dfrac{(n-1)s^2}{\sigma^2}$

6) $\chi^2_{0.995,5} = 0.41$ reject H_0 if $\chi_0^2 < 0.41$ or $\chi^2_{0.005,5} = 16.75$ reject H_0 if $\chi_0^2 > 16.75$

7) n = 6, s = 0.319

$$\chi_0^2 = \frac{5(0.319)^2}{1^2} = 0.509$$

P-value = $P\left(\chi^2 < 0.509\right)$; $\quad 0.01 < P - value < 0.02$

8) Since 0.509>0.41, do not reject the null hypothesis and conclude that there is insufficient evidence to conclude that the variance is not equal to 1.0. The *P*-value is greater than any acceptable significance level, α, therefore we do not reject the null hypothesis.

b) 1) the parameter of interest is the variance of fatty acid measurements, σ^2 (now n=51)
2) $H_0 : \sigma^2 = 1.0$

3) $H_1 : \sigma^2 \neq 1.0$
4) α=0.01

5) The test statistic is: $\chi_0^2 = \frac{(n-1)s^2}{\sigma^2}$

6) $\chi^2_{0.995,50} = 27.99$ reject H_0 if $\chi_0^2 < 27.99$ or $\chi^2_{0.005,5} = 79.49$ reject H_0 if $\chi_0^2 > 79.49$

7) n = 51, s = 0.319

$$\chi_0^2 = \frac{50(0.319)^2}{1^2} = 5.09$$

P-value = $P(\chi^2 < 5.09)$; $\quad P - value < 0.01$

8) Since 5.09<27.99, reject the null hypothesis and conclude that there is sufficient evidence to conclude that the variance is not equal to 1.0. The *P*-value is smaller than any acceptable significance level, α, therefore we do reject the null hypothesis.

c.) The sample size changes the conclusion tha t is drawn. With a small sample size, the results are inconclusive. A larger sample size helps to make sure that the correct conclusion is drawn.

9-81. Assume the data follow a normal distribution.

a) 1) The parameter of interest is the standard deviation, σ.
2) $H_0 : \sigma^2 = (0.00002)^2$
3) $H_1 : \sigma^2 < (0.00002)^2$
4) $\alpha = 0.01$

5) The test statistic is: $\chi_0^2 = \frac{(n-1)s^2}{\sigma^2}$

6) $\chi^2_{0.99,7} = 1.24$ reject H_0 if $\chi_0^2 < 1.24$

7) s = 0.00001 and $\alpha = 0.01$

$$\chi_0^2 = \frac{7(0.00001)^2}{(0.00002)^2} = 1.75$$

1.75 > 1.24, do not reject the null hypothesis; that is, there is insufficient evidence to conclude the standard deviation is at most 0.00002 mm.

b) Although the sample standard deviation is less than the hypothesized value of 0.00002, it is *not significantly less* (when $\alpha = 0.01$) than 0.00002 to conclude the standard deviation is at most 0.00002 mm. The value of 0.00001 could have occurred as a result of sampling variation.

9-82 Assume the data follow a normal distribution.

1) The parameter of interest is the standard deviation of the concentration, σ.
2) $H_0 : \sigma^2 = 4^2$
3) $H_1 : \sigma^2 < 4^2$
4) not given

5) The test statistic is: $\chi_0^2 = \frac{(n-1)s^2}{\sigma^2}$

6) will be determined based on the *P*-value
7) s = 0.004 and n = 10

$$\chi_0^2 = \frac{9(0.004)^2}{(4)^2} = 0.000009$$

P-value = $P(\chi^2 < 0.00009)$; $\quad P - value \cong 0.$

The P-value is approximately 0, therefore we reject the null hypothesis and conclude that the standard deviation of the concentration is less than 4 grams per liter.

9-83. Create a table for the number of nonconforming coil springs (value) and the observed number of times the number appeared. One possible table is:

Value	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Obs	0	0	0	1	4	3	4	6	4	3	0	3	3	2	1	1	0	2	1	2

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

$$\text{sample mean} = \frac{0(0)+1(0)+2(0)+\cdots+19(2)}{40} = 9.325$$

$$\hat{p}_{sample} = \frac{sample \quad mean}{n} = \frac{9.325}{50} = 0.1865$$

Value	Observed	Expected
0	0	0.00165
1	0	0.01889
2	0	0.10608
3	1	0.38911
4	4	1.04816
5	3	2.21073
6	4	3.80118
7	6	5.47765
8	4	6.74985
9	3	7.22141
10	0	6.78777
11	3	5.65869
12	3	4.21619
13	2	2.82541
14	1	1.71190
15	1	0.94191
16	0	0.47237
17	2	0.21659
18	1	0.09103
19	2	0.03515

Since several of the expected values are less than 3, some cells must be combined resulting in the following table:

Value	Observed	Expected
0-5	8	3.77462
6	4	3.80118
7	6	5.47765
8	4	6.74985
9	3	7.22141
10	0	6.78777
11	3	5.65869
12	3	4.21619
≥13	9	6.29436

The degrees of freedom are k − p − 1 = 9 − 1 − 1 = 7

a) 1) The variable of interest is the form of the distribution for the number of nonconforming coil springs.
2) H_0: The form of the distribution is binomial
3) H_1: The form of the distribution is not binomial

4) $\alpha = 0.05$
5) The test statistic is

$$\chi_0^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$$

6) Reject H_0 if $\chi_0^2 > \chi_{0.05,7}^2 = 14.07$
7)

$$\chi_0^2 = \frac{(8 - 3.77462)^2}{3.77462} + \frac{(4 - 3.8.011)^2}{3.8011} + \cdots + \frac{(9 - 6.29436)^2}{6.29436} = 17.929$$

8) Since 17.929 > 14.07 reject H_0. We are able to conclude the distribution of nonconforming springs is not binomial at $\alpha = 0.05$.

b) P-value = 0.0123 (found using Minitab)

9-84 Create a table for the number of errors in a string of 1000 bits (value) and the observed number of times the number appeared. One possible table is:

Value	0	1	2	3	4	5
Obs	3	7	4	5	1	0

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

$$\text{sample mean} = \frac{0(3) + 1(7) + 2(4) + 3(5) + 4(1) + 5(0)}{20} = 1.7$$

$$\hat{p}_{sample} = \frac{sample \quad mean}{n} = \frac{1.7}{1000} = 0.0017$$

Value	0	1	2	3	4	5
Observed	3	7	4	5	1	0
Expected	3.64839	6.21282	5.28460	2.99371	1.27067	0.43103

Since several of the expected values are less than 3, some cells must be combined resulting in the following table:

Value	0	1	2	≥3
Observed	3	7	4	6
Expected	3.64839	6.21282	5.28460	4.69541

The degrees of freedom are $k - p - 1 = 4 - 1 - 1 = 2$

a) 1) The variable of interest is the form of the distribution for the number of errors in a string of 1000 bits.
2) H_0: The form of the distribution is binomial
3) H_1: The form of the distribution is not binomial
4) $\alpha = 0.05$
5) The test statistic is

$$\chi_0^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$$

6) Reject H_0 if $\chi_0^2 > \chi_{0.05,2}^2 = 5.99$
7)

$$\chi_0^2 = \frac{(3 - 3.64839)^2}{3.64839} + \cdots + \frac{(6 - 4.69541)^2}{4.69541} = 0.88971$$

8) Since 0.88971 < 9.49 do not reject H_0. We are unable to reject the null hypothesis that the distribution of the number of errors is binomial at $\alpha = 0.05$.

b) *P*-value = 0.6409 (found using Minitab)

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:

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

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

9-86 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) $H_0 : \mu = 50$
3) $H_1 : \mu < 50$
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) $\bar{x} = 59.87$ s = 12.50 n = 60

$$z_0 = \frac{59.87 - 50}{12.50 / \sqrt{60}} = 6.12$$

8) Since 6.12>-1.65, do not reject the null hypothesis and conclude there is insufficient evidence to indicate that the true mean concentration of suspended solids is less than 50 ppm at α = 0.05.

b) The P-value = $\Phi(6.12) \cong 1$.

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 = (60) (0.125) = 7.5. The table of ranges and their corresponding frequencies is completed as follows.

Interval	Obs. Frequency.	Exp. Frequency.
$x \le 45.509$		7.5
$45.50 < x \le 51.435$		7.5
$51.43 < x \le 55.877$		7.5
$55.87 < x \le 59.8711$		7.5
$59.87 < x \le 63.874$		7.5
$63.87 < x \le 68.319$		7.5
$68.31 < x \le 74.248$		7.5
$x \ge 74.24$	6	7.5

The test statistic is:

$$\chi^2_o = \frac{(9-7.5)^2}{7.5} + \frac{(5-7.5)^2}{7.5} + \cdots + \frac{(8-7.5)^2}{7.5} + \frac{(6-7.5)^2}{7.5} = 5.06$$

and we would reject if this value exceeds $\chi^2_{0.05,5} = 11.07$. Since it does not, we cannot 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 = 100

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

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:

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

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.

9-88 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) $\alpha = 0.01$
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 $z_{0.05} = 2.33$
7) $\bar{x} = 0.324$ s = 0.0131 n = 40

$$z_0 = \frac{0.624 - 0.635}{0.0131/\sqrt{40}} = -5.31$$

8) Since –5.31< 2.33, do not reject the null hypothesis and conclude there is insufficient evidence to indicate that the true mean coefficient of restitution is greater than 0.635 at α = 0.01.

b) The P-value $\Phi(5.31) \cong 1$..

c.) If the lower bound of the CI was above the value 0.635 then we could conclude that the mean coefficient of restitution was greater than 0.635.

9-89 a) In order to use t statistics in hypothesis testing, we need to assume that the underlying distribution is normal. Use the t-test to test the hypothesis that the true mean is 2.5 mg/L.

1) State the parameter of interest: The parameter of interest is the true mean dissolved oxygen level, μ.
2) State the null hypothesis $H_0 : \mu = 2.5$
3) State the alternative hypothesis $H_1 : \mu \neq 2.5$
4) Give the significance level $\alpha = 0.05$
5) Give the statistic

$$t_0 = \frac{\bar{x} - \mu}{s / \sqrt{n}}$$

6) Reject H_0 if $|t_0| < t_{\alpha/2,n-1}$
7) find the sample statistic $\bar{x} = 3.265$ $s = 2.127$ $n = 20$

and calculate the t-statistic $t_0 = \dfrac{\bar{x} - \mu}{s / \sqrt{n}}$

8) Draw your conclusion and find the P-value.

b) Assume the data are normally distributed.

1) The parameter of interest is the true mean dissolved oxygen level, μ.
2) $H_0 : \mu = 2.5$
3) $H_1 : \mu \neq 2.5$
4) $\alpha = 0.05$
5)Test statistic

$$t_0 = \frac{\bar{x} - \mu}{s / \sqrt{n}}$$

6) Reject H_0 if $|t_0| > t_{\alpha/2,n-1}$ where $t_{\alpha/2,n-1} = t_{0.025,19b} = 2.093$
7) $\bar{x} = 3.265$ $s = 2.127$ $n = 20$

$$t_0 = \frac{3.265 - 2.5}{2.127 / \sqrt{20}} = 1.608$$

8) Since 1.608 < 2.093, do not reject the null hypotheses and conclude that the true mean is not significantly different from 2.5 mg/L

c.) The value of 1.608 is found between the columns of 0.05 and 0.1 of table IV. Therefore the P-value is between 0.1 and 0.2. Minitab gives a value of 0.124

d.) The confidence interval found in exercise 8-81 b. agrees with the hypothesis test above. The value of 2.5 is within the 95% confidence limits. The confidence interval shows that the interval is quite wide due to the large sample standard deviation value.

$$\bar{x} - t_{0.025,19} \frac{s}{\sqrt{n}} \le \mu \le \bar{x} + t_{0.025,19} \frac{s}{\sqrt{n}}$$

$$3.265 - 2.093 \frac{2.127}{\sqrt{20}} \le \mu \le 3.265 + 2.093 \frac{2.127}{\sqrt{20}}$$

$$2.270 \le \mu \le 4.260$$

9-90 a) In order to use t statistics in hypothesis testing, we need to assume that the underlying distribution is normal.

$$d = \frac{\delta}{\sigma} = \frac{|\mu - \mu_0|}{\sigma} = \frac{|73-75|}{1} = 2$$

Using the OC curve for $\alpha = 0.05$, $d = 2$, and n = 10, we get $\beta \cong 0.0$ and power of $1-0.0 \cong 1$.

$$d = \frac{\delta}{\sigma} = \frac{|\mu - \mu_0|}{\sigma} = \frac{|72-75|}{1} = 3$$

Using the OC curve for $\alpha = 0.05$, $d = 3$, and n = 10, we get $\beta \cong 0.0$ and power of $1-0.0 \cong 1$.

b) $d = \dfrac{\delta}{\sigma} = \dfrac{|\mu - \mu_0|}{\sigma} = \dfrac{|73-75|}{1} = 2$

Using the OC curve, Chart VI e) for $\alpha = 0.05$, $d = 2$, and $\beta \cong 0.1$ (Power=0.9),

$n^* = 5$. Therefore, $n = \dfrac{n^*+1}{2} = \dfrac{5+1}{2} = 3$

$$d = \frac{\delta}{\sigma} = \frac{|\mu - \mu_0|}{\sigma} = \frac{|72-75|}{1} = 3$$

Using the OC curve, Chart VI e) for $\alpha = 0.05$, $d = 3$, and $\beta \cong 0.1$ (Power=0.9),

$n^* = 3$. Therefore, $n = \dfrac{n^*+1}{2} = \dfrac{3+1}{2} = 2$

c) $\sigma = 2$.

$$d = \frac{\delta}{\sigma} = \frac{|\mu - \mu_0|}{\sigma} = \frac{|73 - 75|}{2} = 1$$

Using the OC curve for $\alpha = 0.05$, $d = 1$, and n = 10, we get $\beta \cong 0.10$ and power of $1-0.10 \cong 0.90$.

$$d = \frac{\delta}{\sigma} = \frac{|\mu - \mu_0|}{\sigma} = \frac{|72 - 75|}{2} = 1.5$$

Using the OC curve for $\alpha = 0.05$, $d = 1.5$, and n = 10, we get $\beta \cong 0.04$ and power of $1-0.04 \cong 0.96$.

$$d = \frac{\delta}{\sigma} = \frac{|\mu - \mu_0|}{\sigma} = \frac{|73 - 75|}{2} = 1$$

Using the OC curve, Chart VI e) for $\alpha = 0.05$, $d = 1$, and $\beta \cong 0.1$ (Power=0.9),

$n^* = 10$. Therefore, $n = \frac{n^* + 1}{2} = \frac{10 + 1}{2} = 5.5$ $n \cong 6$

$$d = \frac{\delta}{\sigma} = \frac{|\mu - \mu_0|}{\sigma} = \frac{|72 - 75|}{2} = 1.5$$

Using the OC curve, Chart VI e) for $\alpha = 0.05$, $d = 3$, and $\beta \cong 0.1$ (Power=0.9),

$n^* = 7$. Therefore, $n = \frac{n^* + 1}{2} = \frac{7 + 1}{2} = 4$

Increasing the standard deviation lowers the power of the test and increases the sample size required to obtain a certain power.

Mind Expanding Exercises

9-91 The parameter of interest is the true,μ.

$H_0 : \mu = \mu_0$

$H_1\ \mu \neq \mu_0$

9-92 a.) Reject H_0 if $z_0 < -z_{\alpha-\varepsilon}$ or $z_0 > z_\varepsilon$

$$P(\frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} < -\frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} = |\mu = \mu_0) + P(\frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} > \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} | \mu = \mu_0)\ \mathrm{P}$$

$$P(z_0 < -z_{\alpha-\varepsilon}) + P(z_0 > z_\varepsilon) = \Phi(-z_{\alpha-\varepsilon}) + 1 - \Phi(z_\varepsilon)$$

$$= ((\alpha - \varepsilon)) + (1 - (1 - \varepsilon)) = \alpha$$

b.) $\beta = P(z_\varepsilon \leq \overline{X} \leq z_\varepsilon$ when $\mu_1 = \mu_0 + d$)

or $\beta = P(-z_{\alpha-\varepsilon} < Z_0 < z_\varepsilon | \mu_1 = \mu_0 + \delta)$

$$\beta = P(-z_{\alpha-\varepsilon} < \frac{\overline{x} - \mu_0}{\sqrt{\sigma^2/n}} < z_\varepsilon | \mu_1 = \mu_0 + \delta)$$

$$= P(-z_{\alpha-\varepsilon} - \frac{\delta}{\sqrt{\sigma^2/n}} < Z < z_\varepsilon - \frac{\delta}{\sqrt{\sigma^2/n}})$$

$$= \Phi(z_\varepsilon - \frac{\delta}{\sqrt{\sigma^2/n}}) - \Phi(-z_{\alpha-\varepsilon} - \frac{\delta}{\sqrt{\sigma^2/n}})$$

]

9-93 1) The parameter of interest is the true mean number of open circuits, λ.

2) $H_0 : \lambda = 2$

3) $H_1 : \lambda > 2$

4) $\alpha = 0.05$

5) Since n>30 we can use the normal distribution

$$z_0 = \frac{\overline{X} - \lambda}{\sqrt{\lambda / n}}$$

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

7) $\bar{x} = 1038/500 = 2.076$ n = 500

$$z_0 = \frac{2.076 - 2}{2/\sqrt{500}} = 0.85$$

8) Since 0.85< 1.65, do not reject the null hypothesis and conclude there is insufficient evidence to indicate that the true mean number of open circuits is greater than 2 at $\alpha = 0.01$

9-94 1) The parameter of interest is the true standard deviation of the golf ball distance, λ.

2) $H_0 : \sigma = 10$

3) $H_1 : \sigma < 10$

4) $\alpha = 0.05$

5) Since n>30 we can use the normal distribution

$$z_0 = \frac{S - \sigma_0}{\sqrt{\sigma_0^2 /(2n)}}$$

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

7) $s = 13.41$ n = 100

$$z_0 = \frac{13.41 - 10}{\sqrt{10^2 /(200)}} = 4.82$$

8) Since 4.82 > -1.65, do not reject the null hypothesis and conclude there is insufficient evidence to indicate that the true standard deviation is less than 10 at $\alpha = 0.05$

9-95 95% percentile $\theta = \mu + 1.645\sigma$

using $z_0 = \dfrac{S - \sigma_0}{\sqrt{\sigma_0^2 /(2n)}}$

95% percentile: $\overline{X} + 1.645\sqrt{s^2 /(2n)}$

$$S.E.(\theta) = \sigma / \sqrt{n} = \frac{s}{\sqrt{2n}\sqrt{n}} = s / \sqrt{3n} =$$

9-96 1) The parameter of interest is the true standard deviation of the golf ball distance, λ.
2) $H_0 : \theta = 285$
3) $H_1 : \sigma > 285$
4) α=0.05
5) Since n>30 we can use the normal distribution

$$z_0 = \frac{\hat{\Theta} - \vartheta_0}{\sqrt{\sigma^2 /(3n)}}$$

6) Reject H_0 if $z_0 > z_\alpha$ where $z_{0.05}$ =1.65
7) $\hat{\Theta}$= 282.36 n = 100

$$z_0 = \frac{282.36 - 285}{\sqrt{10^2 /(300)}} = -4.57$$

8) Since -4.82 > 1.65, do not reject the null hypothesis and conclude there is insufficient evidence to indicate that the true 95% is greater than 285 at $\alpha = 0.05$

9-97 1) The parameter of interest is the true mean number of open circuits, λ.
2) $H_0 : \lambda = \lambda_0$
3) $H_1 : \lambda \neq \lambda_0$
4) $\alpha = 0.05$
5) test statistic

$$\chi_0^2 = \frac{2\lambda\sum_{i=1}^{n} X_i - \lambda_0}{\sqrt{2\lambda\sum_{i=1}^{n} X_i}}$$

6) Reject H_0 if $\chi_0^2 > \chi^2_{a/2,2n}$ or $\chi_0^2 < \chi^2_{1-a/2,2n}$

7) compute $2\lambda\sum_{i=1}^{n} X_i$ and plug into

$$\chi_0^2 = \frac{2\lambda\sum_{i=1}^{n} X_i - \lambda_0}{\sqrt{2\lambda\sum_{i=1}^{n} X_i}}$$

8) make conclusions
alternative hypotheses
1) $H_0 : \lambda = \lambda_0$
$H_1 : \lambda > \lambda_0$
Reject H_0 if $\chi_0^2 > \chi^2_{a,2n}$

2) $H_0 : \lambda = \lambda_0$
$H_1 : \lambda < \lambda_0$
Reject H_0 if $\chi_0^2 < \chi^2_{a,2n}$

CHAPTER 10

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

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

$n_1 = 10$ $n_2 = 10$

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

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 $\Delta = 0.04$

$$n \cong \frac{\left(z_{\alpha/2}+z_{\beta}\right)^2\left(\sigma_1^2+\sigma_2^2\right)}{\delta^2}=\frac{\left(1.96+1.645\right)^2\left((0.02)^2+(0.025)^2\right)}{(0.04)^2}=8.35,\quad n = 9,$$

use $n_1 = n_2 = 9$

10-2. 1) The parameter of interest is the difference in breaking strengths, $\mu_1 - \mu_2$ and $\Delta_0 = 10$

2) $H_0: \mu_1 - \mu_2 = 10$

3) $H_1: \mu_1 - \mu_2 > 10$

4) $\alpha = 0.05$

5) The test statistic is

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

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

7) $\overline{x}_1 = 162.5$ $\quad \overline{x}_2 = 155.0$ $\quad \delta = 10$

$\sigma_1 = 1.0$ $\quad \sigma_2 = 1.0$

$n_1 = 10$ $\quad n_2 = 12$

$$z_0 = \frac{(162.5 - 155.0) - 10}{\sqrt{\frac{(1.0)^2}{10}+\frac{(1.0)^2}{12}}} = -5.84$$

8) Since -5.84 < 1.645 do not reject the null hypothesis and conclude there is insufficient evidence to support the use of plastic 1 at $\alpha = 0.05$.

10-3 $\beta = \Phi\left(1.645 - \frac{(12-10)}{\sqrt{\frac{1}{10}+\frac{1}{12}}}\right) = \Phi(-3.03) = 0.0012$, Power = 1 – 0.0012 = 0.9988 ≈ 1

$$n = \frac{(z_{\alpha/2}+z_\beta)^2(\sigma_1^2+\sigma_2^2)}{(\Delta-\Delta_0)^2} = \frac{(1.645+1.645)^2(1+1)}{(12-10)^2} = 5.42 \cong 6$$

Yes, the sample size is adequate

10-4. a) 1) The parameter of interest is the difference in mean burning rate, $\mu_1 - \mu_2$

2) $H_0: \mu_1 - \mu_2 = 0$ or $\mu_1 = \mu_2$

3) $H_1: \mu_1 - \mu_2 \neq 0$ or $\mu_1 \neq \mu_2$
4) $\alpha = 0.05$
5) The test statistic is

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

6) Reject H_0 if $z_0 < -z_{\alpha/2} = -1.96$ or $z_0 > z_{\alpha/2} = 1.96$
7) $\bar{x}_1 = 18 \quad \bar{x}_2 = 24$
$\sigma_1 = 3 \quad \sigma_2 = 3$
$n_1 = 20 \quad n_2 = 20$

$$z_0 = \frac{(18 - 24)}{\sqrt{\frac{(3)^2}{20} + \frac{(3)^2}{20}}} = -6.32$$

8) Since $-6.32 < -1.96$ reject the null hypothesis and conclude the mean burning rates differ significantly at $\alpha = 0.05$.

b) P-value = $2(1 - \Phi(6.32)) = 2(1-1) = 0$

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

$$= \Phi\left(1.96 - \frac{2.5}{\sqrt{\frac{(3)^2}{20} + \frac{(3)^2}{20}}}\right) - \Phi\left(-1.96 - \frac{2.5}{\sqrt{\frac{(3)^2}{20} + \frac{(3)^2}{20}}}\right)$$

$$= \Phi(1.96 - 2.64) - \Phi(-1.96 - 2.64) = \Phi(-0.68) - \Phi(-4.6)$$

$$= 0.24825 - 0$$
$$= 0.24825$$

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

$$(18 - 24) - 1.96\sqrt{\frac{(3)^2}{20} + \frac{(3)^2}{20}} \le \mu_1 - \mu_2 \le (18 - 24) + 1.96\sqrt{\frac{(3)^2}{20} + \frac{(3)^2}{20}}$$

$$-7.86 \le \mu_1 - \mu_2 \le -4.14$$

We are 95% confident that the mean burning rate for solid fuel propellant 2 exceeds that of propellant 1 by between 4.14 and 7.86 cm/s.

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.

10-6. a) 1) The parameter of interest is the difference in mean fill volume, $\mu_1 - \mu_2$

2) $H_0: \mu_1 - \mu_2 = 0$ or $\mu_1 = \mu_2$

3) $H_1: \mu_1 - \mu_2 \neq 0$ 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) $\bar{x}_1 = 30.87$ $\quad \bar{x}_2 = 30.68$

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

$n_1 = 12$ $\quad n_2 = 10$

$$z_0 = \frac{(30.87 - 30.68)}{\sqrt{\frac{(0.10)^2}{12} + \frac{(0.15)^2}{10}}} = 3.42$$

8) Since 3.42 > 1.96 reject the null hypothesis and conclude the mean fill volumes of machine 1 and machine 2 differ significantly at $\alpha = 0.05$.

b) P-value = $2(1 - \Phi(3.42)) = 2(1 - 0.99969) = 0.00062$

c) Assume the sample sizes are to be equal, use $\alpha = 0.05$, $\beta = 0.10$, and $\Delta = 0.20$

$$n \cong \frac{\left(z_{\alpha/2} + z_\beta\right)^2\left(\sigma_1^2 + \sigma_2^2\right)}{\left(\Delta - \Delta_0\right)^2} = \frac{(1.96 + 1.28)^2\left((0.10)^2 + (0.15)^2\right)}{(-0.20)^2} = 8.53, \quad n = 9, \text{ use } n_1 = n_2 = 9$$

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

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

$n_1 = 15$ $n_2 = 20$

$$z_0 = \frac{(89.6 - 92.5)}{\sqrt{\frac{1.5}{15} + \frac{1.2}{20}}} = -7.25$$

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

10-8. 99% level of confidence, $E = 4$, and $z_{0.005} = 2.575$

$$n \cong \left(\frac{z_{0.005}}{E}\right)^2 \left(\sigma_1^2 + \sigma_2^2\right) = \left(\frac{2.575}{4}\right)^2 (9+9) = 7.46,$$ n = 8, use $n_1 = n_2 = 8$

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

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

10-10.

Case 1: Before Process Change

μ_1 = mean batch viscosity before change

$\bar{x}_1 = 750.2$

$\sigma_1 = 20$

$n_1 = 15$

Case 2: After Process Change

μ_2 = mean batch viscosity after change

$\bar{x}_2 = 756.88$

$\sigma_2 = 20$

$n_2 = 8$

90% confidence on $\mu_1 - \mu_2$, the difference in mean batch viscosity before and after process change:

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

$$(750.2 - 756.88) - 1.645\sqrt{\frac{(20)^2}{15} + \frac{(20)^2}{8}} \le \mu_1 - \mu_2 \le \left(750.2 - 756/88\right) + 1.645\sqrt{\frac{(20)^2}{15} + \frac{(20)^2}{8}}$$

$$-21.08 \le \mu_1 - \mu_2 \le 7.72$$

We are 90% confident that the difference in mean batch viscosity before and after the process change lies within -21.08 and 7.72. Since 0 is contained in this interval we can conclude with 90% confidence that the mean batch viscosity was unaffected by the process change.

10-11. Catalyst 1 Catalyst 2

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

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

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

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

$$-5.83 \le \mu_1 - \mu_2 \le -0.57$$

We are 95% confident that the mean active concentration of catalyst 2 exceeds that of catalyst 1 by between 0.57 and 5.83 g/l.

b) Yes, since the 95% confidence interval did not contain the value 0, we would conclude that the mean active concentration depends on the choice of catalyst.

10-12. a) 1) The parameter of interest is the difference in mean batch viscosity before and after the process change, $\mu_1 - \mu_2$

2) $H_0 : \mu_1 - \mu_2 = 10$

3) $H_1 : \mu_1 - \mu_2 < 10$

4) $\alpha = 0.10$

5) The test statistic is

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

6) Reject H_0 if $z_0 < -z_\alpha$ where $z_{0.1} = -1.28$

7) $\bar{x}_1 = 750.2$ $\quad$ $\bar{x}_2 = 756.88$ $\quad$ $\Delta_0 = 10$
$\sigma_1 = 20$ $\quad$ $\sigma_2 = 20$
$n_1 = 15$ $\quad$ $n_2 = 8$

$$z_0 = \frac{(750.2 - 756.88) - 10}{\sqrt{\frac{(20)^2}{15} + \frac{(20)^2}{8}}} = -1.90$$

8) Since $-1.90 < -1.28$ reject the null hypothesis and conclude the process change has increased the mean by less than 10.

b) P-value = $P(z \le -1.90) = 1 - P(z \le 1.90) = 1 - 0.97128 = 0.02872$

c) Parts a and b above give evidence that the mean batch viscosity change is less than 10. This conclusion is also seen by the confidence interval given in a previous problem since the interval did not contain the value 10. Since the upper endpoint is 7.72, then this also gives evidence that the difference is less than 10.

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

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

6) Reject H_0 if $z_0 < -z_{\alpha/2} = -1.96$ or $z_0 > z_{\alpha/2} = 1.96$
7) 7) $\overline{x}_1 = 65.22$ $\overline{x}_2 = 68.42$ $\delta = 0$
$\sigma_1 = 3$ $\sigma_2 = 3$
$n_1 = 10$ $n_2 = 10$

$$z_0 = \frac{(65.22 - 68.42) - 0}{\sqrt{\frac{9}{10} + \frac{9}{10}}} = -2.385$$

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.

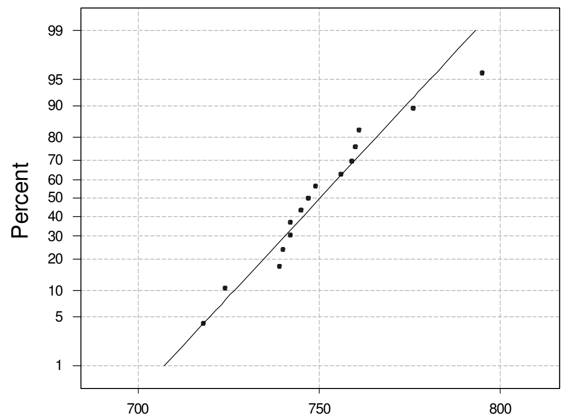
10-14.

$$\beta = \Phi\left(1.96 - \frac{(5)}{\sqrt{\frac{3^2}{10} + \frac{3^2}{10}}}\right) - \Phi\left(-1.96 - \frac{(5)}{\sqrt{\frac{3^2}{10} + \frac{3^2}{10}}}\right)$$
$$= \Phi(-1.77) - \Phi(-5.69) = 0.038364 - 0$$
$$= 0.038364$$

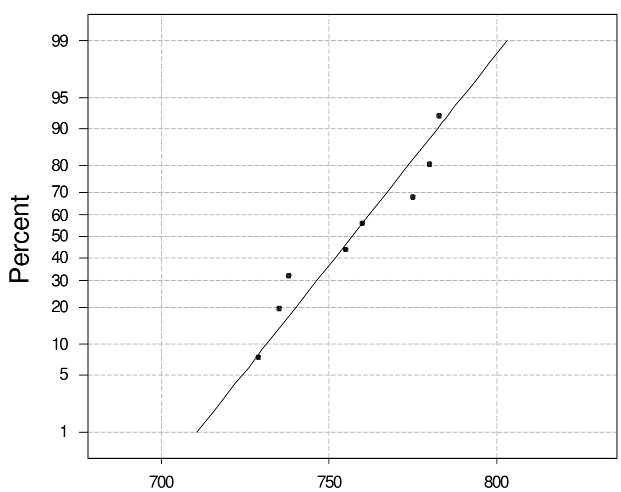
Power = $1 - \beta = 1 - 0.038364 = 09616$. it would appear that the sample sizes are adequate to detect the difference of 5, based on the power. Calculate the value of n using α and β.

$$n \cong \frac{(z_{\alpha/2} + z_\beta)^2(\sigma_1^2 + \sigma_2^2)}{(\Delta - \Delta_0)^2} = \frac{(1.96 + 1.77)^2(9+9)}{(5)^2} = 10.02,$$ Therefore, 10 is just slightly too few samples.

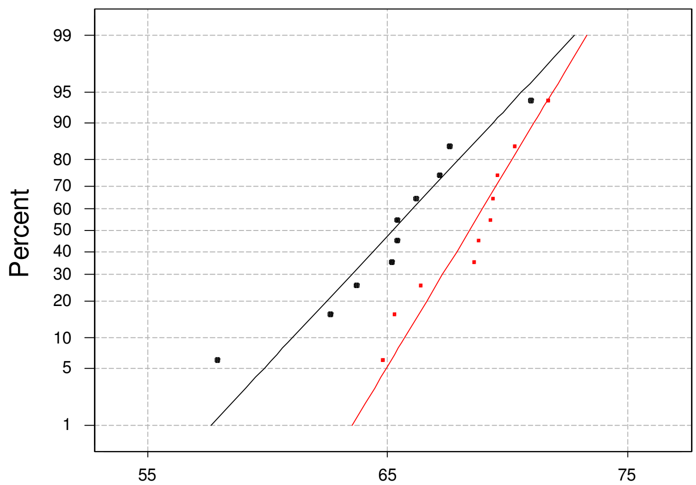
10-15 The data from the first sample n=15 appear to be normally distributed.



The data from the second sample n=8 appear to be normally distributed



10-16 The data all appear to be normally distributed based on the normal probability plot below.



Section 10-3

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

2) $H_0: \mu_1 - \mu_2 = 0$ or $\mu_1 = \mu_2$

3) $H_1: \mu_1 - \mu_2 \neq 0$ or $\mu_1 \neq \mu_2$

4) $\alpha = 0.05$

5) The test statistic is

$$t_0 = \frac{(\overline{x}_1 - \overline{x}_2) - \Delta_0}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

6) Reject the null hypothesis if $t_0 < -t_{\alpha/2,n_1+n_2-2}$ where $-t_{0.025,30} = -2.042$ or $t_0 > t_{\alpha/2,n_1+n_2-2}$ where $t_{0.025,30} = 2.042$

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

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

$n_1 = 15$ $n_2 = 17$

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

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

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: $t_{0.025,30} = 2.042$

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

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

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

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.

10-18. Assume the populations follow normal distributions and $\sigma_1^2 = \sigma_2^2$. The assumption of equal variances may be permitted in this case since it is known that the t-test and confidence intervals involving the t-distribution are robust to this assumption of equal variances when sample sizes are equal.

Case 1: AFCC

μ_1 = mean foam expansion for AFCC

$\bar{x}_1 = 4.7$

$s_1 = 0.6$

$n_1 = 5$

Case 2: ATC

μ_2 = mean foam expansion for ATC

$\bar{x}_2 = 6.9$

$s_2 = 0.8$

$n_2 = 5$

95% confidence interval: $t_{0.025,8} = 2.306$

$$s_p = \sqrt{\frac{4(0.60)+4(0.80)}{8}} = 0.7071$$

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

$$(4.7 - 6.9) - 2.306(0.7071)\sqrt{\frac{1}{5}+\frac{1}{5}} \le \mu_1 - \mu_2 \le \left(4.7 - 6.9\right) + 2.306(0.7071)\sqrt{\frac{1}{5}+\frac{1}{5}}$$

$$-3.23 \le \mu_1 - \mu_2 \le -1.17$$

Yes, with 95% confidence, we believe the mean foam expansion for ATC exceeds that of AFCC by between 1.17 and 3.23.

10-19. a) 1) The parameter of interest is the difference in mean catalyst yield, $\mu_1 - \mu_2$, with $\Delta_0 = 0$

2) $H_0: \mu_1 - \mu_2 = 0$ or $\mu_1 = \mu_2$

3) $H_1: \mu_1 - \mu_2 < 0$ or $\mu_1 < \mu_2$

4) $\alpha = 0.01$

5) The test statistic is

$$t_0 = \frac{(\overline{x}_1 - \overline{x}_2) - \Delta_0}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

6) Reject the null hypothesis if $t_0 < -t_{\alpha,n_1+n_2-2}$ where $-t_{0.01,25} = -2.485$

7) $\overline{x}_1 = 86$ $\quad \overline{x}_2 = 89$

$s_1 = 3$ $\quad s_2 = 2$

$n_1 = 12$ $\quad n_2 = 15$

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

$$= \sqrt{\frac{11(3)^2 + 14(2)^2}{25}} = 2.4899$$

$$t_0 = \frac{(86-89)}{2.4899\sqrt{\frac{1}{12}+\frac{1}{15}}} = -3.11$$

8) Since $-3.11 < -2.787$, reject the null hypothesis and conclude that the mean yield of catalyst 2 significantly exceeds that of catalyst 1 at $\alpha = 0.01$.

b) 99% confidence interval: $t_{0.005,19} = 2.861$

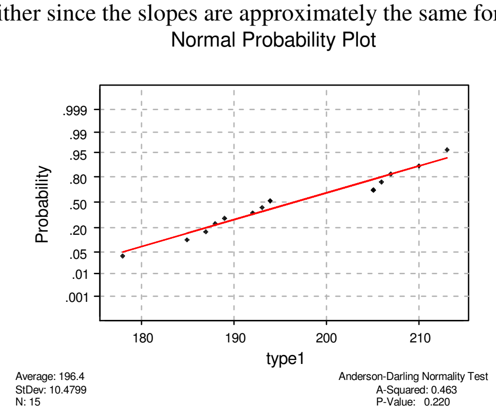
$$(\overline{x}_1 - \overline{x}_2) - t_{\alpha/2,n_1+n_2-2}(s_p)\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} \le \mu_1 - \mu_2 \le (\overline{x}_1 - \overline{x}_2) + t_{\alpha/2,n_1+n_2-2}(s_p)\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$

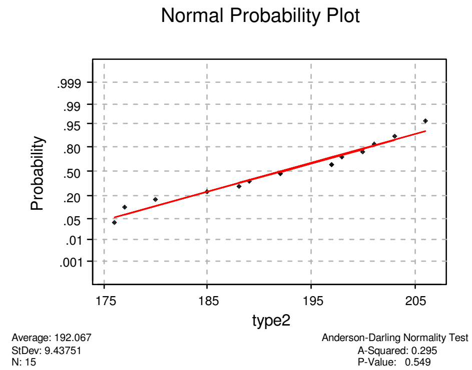
$$(86-89) - 2.787\,(2.4899)\sqrt{\frac{1}{12}+\frac{1}{15}} \le \mu_1 - \mu_2 \le (86-89) + 2.787\,(2.4899)\sqrt{\frac{1}{12}+\frac{1}{15}}$$

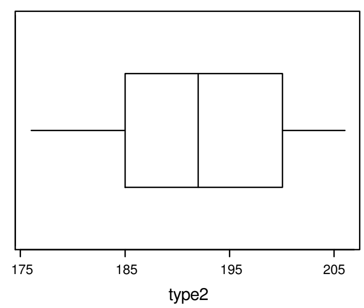
$$-5.688 \le \mu_1 - \mu_2 \le -0.3122$$

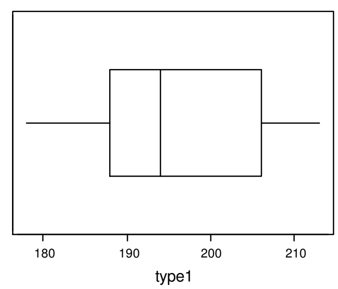
We are 95% confident that the mean yield of catalyst 2 exceeds that of catalyst 1 by between 0.3122 and 5.688

10-20. a) According to the normal probability plots, the assumption of normality appears to be met since the data fall approximately along a straight line. The equality of variances does not appear to be severely violated either since the slopes are approximately the same for both samples.









b) 1) The parameter of interest is the difference in deflection temperature under load, $\mu_1 - \mu_2$, with $\Delta_0 = 0$

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

$$t_0 = \frac{(\overline{x}_1 - \overline{x}_2) - \Delta_0}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

6) Reject the null hypothesis if $t_0 < -t_{\alpha, n_1+n_2-2}$ where $-t_{0.05,28} = -1.701$

7) Type 1 Type 2

$\bar{x}_1 = 196.4$ $\bar{x}_2 = 192.067$

$s_1 = 10.48$ $s_2 = 9.44$

$n_1 = 15$ $n_2 = 15$

$$s_p = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}} = \sqrt{\frac{14(10.48)^2 + 14(9.44)^2}{28}} = 9.97$$

$$t_0 = \frac{(196.4 - 192.067)}{9.97\sqrt{\frac{1}{15}+\frac{1}{15}}} = 1.19$$

8) Since $1.19 > -1.701$ do not reject the null hypothesis and conclude the mean deflection temperature under load for type 2 does not significantly exceed the mean deflection temperature under load for type 1 at the 0.05 level of significance.

c) P-value = $2P(t < 1.19)$ $0.75 < \text{p-value} < 0.90$

d) $\Delta = 5$ Use s_p as an estimate of σ:

$$d = \frac{\mu_2 - \mu_1}{2s_p} = \frac{5}{2(9.97)} = 0.251$$

Using Chart VI g) with $\beta = 0.10$, $d = 0.251$ we get $n \cong 100$. So, since $n^* = 2n-1$, $n_1 = n_2 = 51$; Therefore, the sample sizes of 15 are inadequate.

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) $\bar{x}_1 = 9.97$ $\bar{x}_2 = 10.4$

$s_1 = 0.422$ $s_2 = 0.231$

$n_1 = 10$ $n_2 = 10$

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

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

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 = $2P(t < -2.83)$ $2(0.005) < \text{P-value} < 2(0.010) = 0.010 < \text{P-value} < 0.020$

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

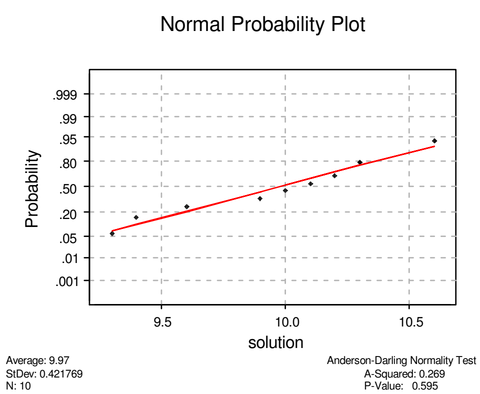
$$(\bar{x}_1 - \bar{x}_2) - t_{\alpha/2,n_1+n_2-2}(s_p)\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} \le \mu_1 - \mu_2 \le (\bar{x}_1 - \bar{x}_2) + t_{\alpha/2,n_1+n_2-2}(s_p)\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$

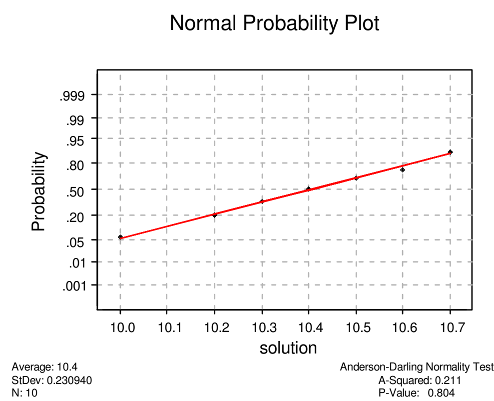
$$(9.97 - 10.4) - 2.101(.340)\sqrt{\frac{1}{10}+\frac{1}{10}} \le \mu_1 - \mu_2 \le (9.97 - 10.4) + 2.101(.340)\sqrt{\frac{1}{10}+\frac{1}{10}}$$

$$-0.7495 \le \mu_1 - \mu_2 \le -0.1105$$

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 0.7495.

d) According to the normal probability plots, the assumption of normality appears to be met since the data from both the samples fall approximately along a straight line. The equality of variances does not appear to be severely violated either since the slopes are approximately the same for both samples.





10-22. a) 1) The parameter of interest is the difference in mean impact strength, $\mu_1 - \mu_2$, with $\Delta_0 = 0$

2) $H_0: \mu_1 - \mu_2 = 0$ or $\mu_1 = \mu_2$

3) $H_1: \mu_1 - \mu_2 < 0$ or $\mu_1 < \mu_2$

4) $\alpha = 0.05$

5) The test statistic is

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

6) Reject the null hypothesis if $t_0 < -t_{\alpha,\nu}$ where $t_{0.05,23} = 1.714$ since

$$\nu = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)}{n_2 - 1}} = 23.72$$

$$\nu \cong 23$$

(truncated)

7) $\bar{x}_1 = 290 \quad \bar{x}_2 = 321$

$s_1 = 12 \quad s_2 = 22$

$n_1 = 10 \quad n_2 = 16$

$$t_0 = \frac{(290 - 321)}{\sqrt{\frac{(12)^2}{10} + \frac{(22)^2}{16}}} = -4.64$$

8) Since $-4.64 < -1.714$ reject the null hypothesis and conclude that supplier 2 provides gears with higher mean impact strength at the 0.05 level of significance.

b) P-value = $P(t < -4.64)$: P-value < 0.0005

c) 1) The parameter of interest is the difference in mean impact strength, $\mu_2 - \mu_1$

2) $H_0: \mu_2 - \mu_1 = 25$

3) $H_1: \mu_2 - \mu_1 > 25$ or $\mu_2 > \mu_1 + 25$

4) $\alpha = 0.05$

5) The test statistic is

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

6) Reject the null hypothesis if $t_0 > t_{\alpha,\nu} = 1.708$ where

$$\nu = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)}{n_2 - 1}} = 23.72$$

$$\nu \cong 23$$

7) $\bar{x}_1 = 290 \quad \bar{x}_2 = 321 \quad \Delta_0 = 25 \quad s_1 = 12 \quad s_2 = 22 \quad n_1 = 10 \quad n_2 = 16$

$$t_0 = \frac{(321 - 290) - 25}{\sqrt{\frac{(12)^2}{10} + \frac{(22)^2}{16}}} = 0.898$$

8) Since $0.898 < 1.714$, do not reject the null hypothesis and conclude that the mean impact strength from supplier 2 is not at least 25 ft-lb higher that supplier 1 using $\alpha = 0.05$.

10-23. Using the information provided in Exercise 9-20, and $t_{0.025,25} = 2.06$, we find a 95% confidence interval on the difference, $\mu_2 - \mu_1$:

$$(\bar{x}_2 - \bar{x}_1) - t_{0.025,25}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \le \mu_2 - \mu_1 \le (\bar{x}_2 - \bar{x}_1) + t_{0.025,25}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$$31 - 2.069(6.682) \le \mu_2 - \mu_1 \le 31 + 2.069(6.682)$$

$$17.175 \le \mu_2 - \mu_1 \le 44.825$$

Since the 95% confidence interval represents the differences that $\mu_2 - \mu_1$ could take on with 95% confidence, we can conclude that Supplier 2 does provide gears with a higher mean impact strength than Supplier 1. This is visible from the interval (17.175, 44.825) since zero is not contained in the interval and the differences are all positive, meaning that $\mu_2 - \mu_1 > 0$.

10-24 a) 1) The parameter of interest is the difference in mean speed, $\mu_1 - \mu_2$, $\Delta_0 = 0$

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) $\alpha = 0.10$

5) The test statistic is

$$t_0 = \frac{(\overline{x}_1 - \overline{x}_2) - \Delta_0}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

6) Reject the null hypothesis if $t_0 > t_{\alpha, n_1+n_2-2}$ where $t_{0.10,14} = 1.345$

7) Case 1: 25 mil   Case 2: 20 mil

$\overline{x}_1 = 1.15$ $\overline{x}_2 = 1.06$

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

$s_1 = 0.11$ $s_2 = 0.09$

$$= \sqrt{\frac{7(0.11)^2 + 7(0.09)^2}{14}} = 0.1005$$

$n_1 = 8$ $n_2 = 8$

$$t_0 = \frac{(1.15 - 1.06)}{0.1005\sqrt{\frac{1}{8} + \frac{1}{8}}} = 1.79$$

8) Since 1.79 > 1.345 reject the null hypothesis and conclude reducing the film thickness from 25 mils to 20 mils significantly increases the mean speed of the film at the 0.10 level of significance (Note: since increase in film speed will result in *lower* values of observations).

b) P-value = $P(t > 1.79)$ $0.025 < \text{P-value} < 0.05$

c) 90% confidence interval: $t_{0.025,14} = 2.145$

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

$$(1.15 - 1.06) - 2.145(.1005)\sqrt{\frac{1}{8} + \frac{1}{8}} \le \mu_1 - \mu_2 \le (1.15 - 1.06) + 2.145(.1005)\sqrt{\frac{1}{8} + \frac{1}{8}}$$

$$-0.0178 \le \mu_1 - \mu_2 \le 0.1978$$

We are 90% confident the mean speed of the film at 20 mil exceeds the mean speed for the film at 25 mil by between -0.0178 and 0.1978 $\mu J/in^2$.

10-25. 1) The parameter of interest is the difference in mean melting point, $\mu_1 - \mu_2$, 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) $\alpha = 0.02$

5) The test statistic is

$$t_0 = \frac{(\bar{x}_1 - \bar{x}_2) - \Delta_0}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

6) Reject the null hypothesis if $t_0 < -t_{\alpha/2, n_1+n_2-2}$ where $-t_{0.0025,40} = -2.021$ or $t_0 > t_{\alpha/2, n_1+n_2-2}$ where $t_{0.025,40} = 2.021$

7) $\bar{x}_1 = 420$ $\bar{x}_2 = 426$

$s_1 = 4$ $s_2 = 3$

$n_1 = 21$ $n_2 = 21$

$$s_p = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}} = \sqrt{\frac{20(4)^2 + 20(3)^2}{40}} = 3.536$$

$$t_0 = \frac{(420-426)}{3.536\sqrt{\frac{1}{21}+\frac{1}{21}}} = -5.498$$

8) Since $-5.498 < -2.021$ reject the null hypothesis and conclude that the data do not support the claim that both alloys have the same melting point at $\alpha = 0.02$

P-value = $2P\left(t < -5.498\right)$ P-value < 0.0010

10-26. $d = \frac{|\mu_1 - \mu_2|}{2\sigma} = \frac{3}{2(4)} = 0.375$

Using the appropriate chart in the Appendix, with $\beta = 0.10$ and $\alpha = 0.05$ we have: $n^* = 75$, so

$n = \frac{n^* + 1}{2} = 38$, $n_1 = n_2 = 38$

10-27. a) 1) The parameter of interest is the difference in mean wear amount, $\mu_1 - \mu_2$, 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

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

6) Reject the null hypothesis if $t_0 < -t_{0.025,26}$ where $-t_{0.025,26} = -2.056$ or $t_0 > t_{0.025,26}$ where $t_{0.025,26} =$ 2.056 since

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

$$\nu \cong 26$$

(truncated)

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

$s_1 = 2$ $s_2 = 8$

$n_1 = 25$ $n_2 = 25$

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

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) < P-value < 2(0.005)

0.005 < P-value < 0.010

c) 1) The parameter of interest is the difference in mean wear amount, $\mu_1 - \mu_2$

2) $H_0: \mu_1 - \mu_2 = 0$

3) $H_1: \mu_1 - \mu_2 > 0$

4) $\alpha = 0.05$

5) The test statistic is $$t_0 = \frac{(\bar{x}_1 - \bar{x}_2) - \Delta_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

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

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

$s_1 = 2$ $s_2 = 8$

$n_1 = 25$ $n_2 = 25$ $$t_0 = \frac{(20-15)}{\sqrt{\frac{(2)^2}{25} + \frac{(8)^2}{25}}} = 3.03$$

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.

10-28 1) The parameter of interest is the difference in mean coating thickness, $\mu_1 - \mu_2$, with $\Delta_0 = 0$.

2) $H_0: \mu_1 - \mu_2 = 0$

3) $H_1: \mu_1 - \mu_2 > 0$

4) $\alpha = 0.01$

5) The test statistic is

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

6) Reject the null hypothesis if $t_0 > t_{0.01,18}$ where $t_{0.01,18} = 2.552$ since

$$\nu = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)}{n_2 - 1}} 18.37$$

$\nu \cong 18$
(truncated)

7) $\bar{x}_1 = 103.5$ $\bar{x}_2 = 99.7$

$s_1 = 10.2$ $s_2 = 20.1$

$n_1 = 11$ $n_2 = 13$

$$t_0 = \frac{(103.5 - 99.7)}{\sqrt{\frac{(10.2)^2}{11} + \frac{(20.1)^2}{13}}} = 0.597$$

8) Since 0.597 < 2.552, do not reject the null hypothesis and conclude that increasing the temperature does not significantly reduce the mean coating thickness at $\alpha = 0.01$.

P-value = P(t > 0597), 0.25 < P-value < 0.40

10-29. If $\alpha = 0.01$, construct a 99% two-sided confidence interval on the difference to answer question 10-28.
$t_{0.005,19} = 2.878$

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

$$(103.5 - 99.7) - 2.878\sqrt{\frac{(10.2)^2}{11} + \frac{(20.1)^2}{13}} \le \mu_1 - \mu_2 \le (103.5 - 99.7) - 2.878\sqrt{\frac{(10.2)^2}{11} + \frac{(20.1)^2}{13}}$$

$$-14.52 \le \mu_1 - \mu_2 \le 22.12.$$

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.

10-30. 95% confidence interval:
$t_{0.025,26} = 2.056$

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

$$(20 - 15) - 2.056\sqrt{\frac{(2)^2}{25} + \frac{(8)^2}{25}} \le \mu_1 - \mu_2 \le (20 - 15) + 2.056\sqrt{\frac{(2)^2}{25} + \frac{(8)^2}{25}}$$

$$1.609 \le \mu_1 - \mu_2 \le 8.391$$

95% lower one-sided confidence interval:

$$t_{0.05,26} = 1.706$$

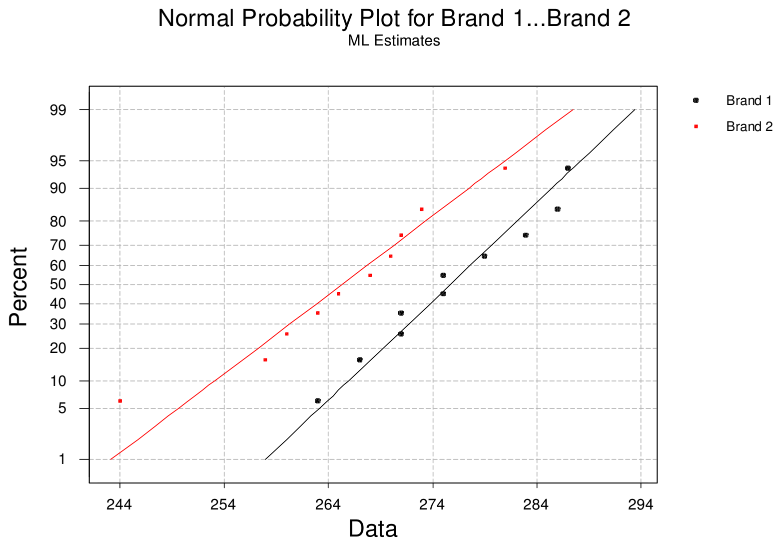
$$(\bar{x}_1 - \bar{x}_2) - t_{\alpha,\nu}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \le \mu_1 - \mu_2$$

$$(20 - 15) - 1.706\sqrt{\frac{(2)^2}{25} + \frac{(8)^2}{25}} \le \mu_1 - \mu_2$$

$$2.186 \le \mu_1 - \mu_2$$

For part a):
We are 95% confident the mean abrasive wear from company 1 exceeds the mean abrasive wear from company 2 by between 1.609 and 8.391 mg/1000.
For part c):
We are 95% confident the mean abrasive wear from company 1 exceeds the mean abrasive wear from company 2 by at least 2.19mg/1000.

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

$$t_0 = \frac{(\overline{x}_1 - \overline{x}_2) - \Delta_0}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

6) Reject the null hypothesis if $t_0 < -t_{\alpha/2,n_1+n_2-2}$ where $-t_{0.025,18} = -2.101$ or $t_0 > t_{\alpha/2,n_1+n_2-2}$ where $t_{0.025,18} = 2.101$

7) $\overline{x}_1 = 275.7$ $\overline{x}_2 = 265.3$

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

$s_1 = 8.03$ $s_2 = 10.04$

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

$n_1 = 10$ $n_2 = 10$

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

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

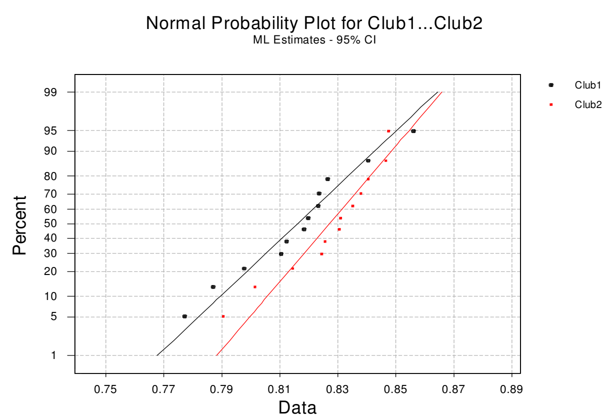
e.) 1-β=0.25 β=0..27 $d = \frac{3}{2(9.09)} = 0.165$ n*=100 $n = \frac{100 + 1}{2} = 50.5$

Therefore, n=51

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

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

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



10-32 a.)

The data appear to be normally distributed and the variances appear to be approximately equal. The slopes of the lines on the normal probability plots are almost the same.

b)

1) The parameter of interest is the difference in mean coefficient of restitution, $\mu_1 - \mu_2$

2) $H_0: \mu_1 - \mu_2 = 0$ or $\mu_1 = \mu_2$

3) $H_1: \mu_1 - \mu_2 \ne 0$ or $\mu_1 \ne \mu_2$

4) $\alpha = 0.05$

5) The test statistic is

$$t_0 = \frac{(\bar{x}_1 - \bar{x}_2) - \Delta_0}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

6) Reject the null hypothesis if $t_0 < -t_{\alpha/2, n_1+n_2-2}$ where $-t_{0.025,22} = -2.074$ or $t_0 > t_{\alpha/2, n_1+n_2-2}$ where

$t_{0.025,22} = 2.074$

7) $\overline{x}_1 = 0.8161$ $\overline{x}_2 = 0.8271$

$s_1 = 0.0217$ $s_2 = 0.0175$

$n_1 = 12$ $n_2 = 12$

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

$$= \sqrt{\frac{11(0.0217)^2 + 11(0.0175)^2}{22}} = 0.01971$$

$$t_0 = \frac{(0.8161 - 0.8271)}{0.01971\sqrt{\frac{1}{12}+\frac{1}{12}}} = -1.367$$

8) Since –1.367 > -2.074 do not reject the null hypothesis and conclude that the data do not support the claim that there is a difference in the mean coefficients of restitution for club1 and club2 at $\alpha = 0.05$

c.)P-value = 2P$\left(t < -1.36\right)$ P-value ≈ 2(0.1)=0.2

d.) $d = \dfrac{0.2}{2(0.01971)} = 5.07$ $\beta \cong 0$ Power ≅1

e.) 1-β= 0.8 β= 0.2 $d = \dfrac{0.1}{2(0.01971)} = 2.53$ n*=4 , $n = \dfrac{n^*+1}{2} = 2.5$ $n \cong 3$

f.) 95% confidence interval

$$\left(\overline{x}_1 - \overline{x}_2\right) - t_{\alpha,\nu} s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} \le \mu_1 - \mu_2 \le \left(\overline{x}_1 - \overline{x}_2\right) + t_{\alpha,\nu} s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$

$$(0.8161 - 0.8271) - 2.074(0.01971)\sqrt{\frac{1}{12}+\frac{1}{12}} \le \mu_1 - \mu_2 \le (0.8161 - 0.8271) + 2.074(0.01971)\sqrt{\frac{1}{12}+\frac{1}{12}}$$

$$-0.0277 \le \mu_1 - \mu_2 \le 0.0057$$

Zero is included in the confidence interval, so we would conclude that there is not a significant difference in the mean coefficient of restitution's for each club at α=0.05.

Section 10-4

10-33. $\overline{d} = 0.2736$ $s_d = 0.1356$, n = 9

95% confidence interval:

$$\overline{d} - t_{\alpha/2,n-1}\left(\frac{s_d}{\sqrt{n}}\right) \le \mu_d \le \overline{d} + t_{\alpha/2,n-1}\left(\frac{s_d}{\sqrt{n}}\right)$$

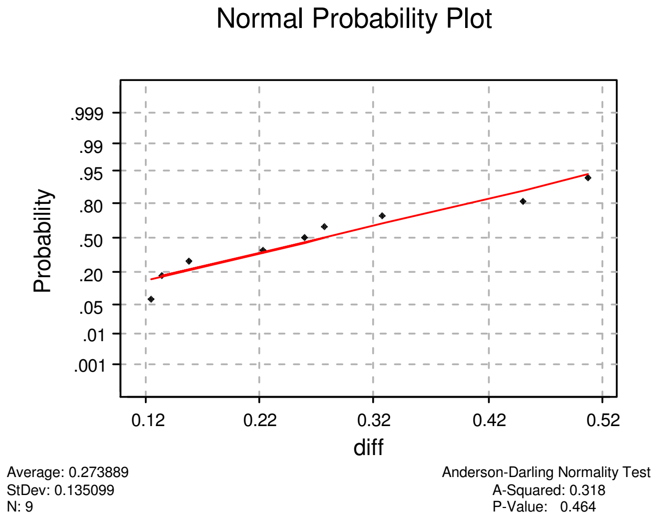
$$0.2736 - 2.306\left(\frac{0.1356}{\sqrt{9}}\right) \le \mu_d \le 0.2736 + 2.306\left(\frac{0.1356}{\sqrt{9}}\right)$$

$$0.1694 \le \mu_d \le 0.3778$$

With 95% confidence, we believe the mean shear strength of Karlsruhe method exceeds the mean shear strength of the Lehigh method by between 0.1694 and 0.3778. Since 0 is not included in this interval, the interval is consistent with rejecting the null hypothesis that the means are the same.

The 95% confidence interval is directly related to a test of hypothesis with 0.05 level of significance, and the conclusions reached are identical.

10-34. It is only necessary for the differences to be normally distributed for the paired t-test to be appropriate and reliable. Therefore, the t-test is appropriate.



10-35. 1) The parameter of interest is the difference between the mean parking times, μ_d.

2) $H_0: \mu_d = 0$

3) $H_1: \mu_d \neq 0$

4) $\alpha = 0.10$

5) The test statistic is

$$t_0 = \frac{\bar{d}}{s_d / \sqrt{n}}$$

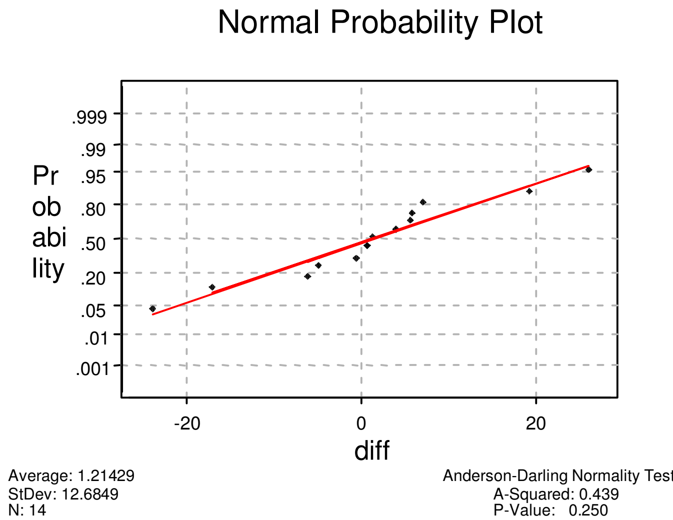
6) Reject the null hypothesis if $t_0 < -t_{0.05,13}$ where $-t_{0.05,13} = -1.771$ or $t_0 > t_{0.05,13}$ where $t_{0.05,13} = 1.771$

7) $\bar{d} = 1.21$
$s_d = 12.68$
$n = 14$

$$t_0 = \frac{1.21}{12.68 / \sqrt{14}} = 0.357$$

8) Since $-1.771 < 0.357 < 1.771$ do not reject the null and conclude the data do not support the claim that the two cars have different mean parking times at the 0.10 level of significance. The result is consistent with the confidence interval constructed since 0 is included in the 90% confidence interval.

10-36. According to the normal probability plots, the assumption of normality does not appear to be violated since the data fall approximately along a straight line.



10-37 $\bar{d} = 868.375$ $s_d = 1290$, n = 8 where d_i = brand 1 - brand 2

99% confidence interval:

$$\bar{d} - t_{\alpha/2,n-1}\left(\frac{s_d}{\sqrt{n}}\right) \le \mu_d \le \bar{d} + t_{\alpha/2,n-1}\left(\frac{s_d}{\sqrt{n}}\right)$$

$$868.375 - 3.499\left(\frac{1290}{\sqrt{8}}\right) \le \mu_d \le 868.375 + 3.499\left(\frac{1290}{\sqrt{8}}\right)$$

$-727.46 \le \mu_d \le 2464.21$

Since this confidence interval contains zero, we are 99% confident there is no significant difference between the two brands of tire.

10-38. a) $\bar{d} = 0.667$ $s_d = 2.964$, n = 12

95% confidence interval:

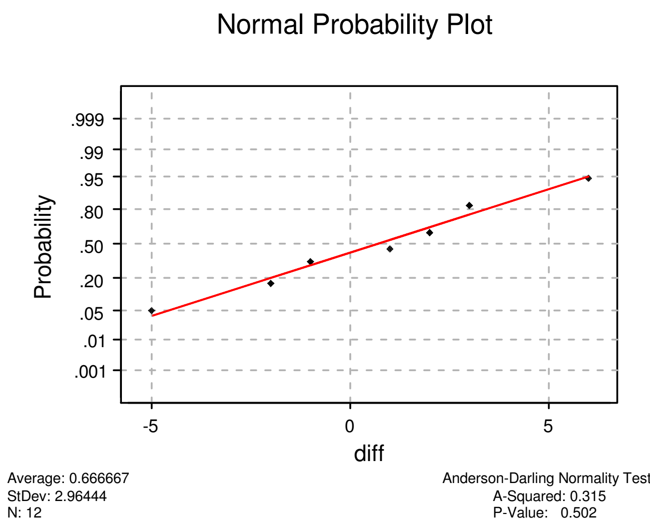
$$\bar{d} - t_{\alpha/2,n-1}\left(\frac{s_d}{\sqrt{n}}\right) \le \mu_d \le \bar{d} + t_{\alpha/2,n-1}\left(\frac{s_d}{\sqrt{n}}\right)$$

$$0.667 - 2.201\left(\frac{2.964}{\sqrt{12}}\right) \le \mu_d \le 0.667 + 2.201\left(\frac{2.964}{\sqrt{12}}\right)$$

$-1.216 \le \mu_d \le 2.55$

Since zero is contained within this interval, we are 95% confident there is no significant indication that one design language is preferable.

b) According to the normal probability plots, the assumption of normality does not appear to be violated since the data fall approximately along a straight line.



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.

10-40. a) 1) The parameter of interest is the mean difference in natural vibration frequencies, μ_d where d_i = finite element − Equivalent Plate.

2) $H_0: \mu_d = 0$

3) $H_1: \mu_d \neq 0$

4) $\alpha = 0.05$

5) The test statistic is

$$t_0 = \frac{\bar{d}}{s_d / \sqrt{n}}$$

6) Reject the null hypothesis if $t_0 < -t_{0.025,6}$ where $-t_{0.025,6} = -2.447$ or $t_0 > t_{0.005,6}$ where $t_{0.005,6} =$ 2.447

7) $\bar{d} = -5.49$
$s_d = 5.924$
$n = 7$

$$t_0 = \frac{-5.49}{5.924 / \sqrt{7}} = -2.45$$

8) Since −2.447< −2.45 < 2.447, do not reject the null and conclude the data suggest that the two methods do not produce significantly different mean values for natural vibration frequency at the 0.05 level of significance.

b) 95% confidence interval:

$$\bar{d} - t_{\alpha/2,n-1}\left(\frac{s_d}{\sqrt{n}}\right) \le \mu_d \le \bar{d} + t_{\alpha/2,n-1}\left(\frac{s_d}{\sqrt{n}}\right)$$

$$-5.49 - 2.447\left(\frac{5.924}{\sqrt{7}}\right) \le \mu_d \le -5.49 + 2.447\left(\frac{5.924}{\sqrt{7}}\right)$$

$-10.969 \le \mu_d \le -0.011$

With 95% confidence, we believe that the mean difference between the natural vibration frequency from the equivalent plate method and the natural vibration frequency from the finite element method is between −10.969 and -0.011 cycles.

10-41. 1) The parameter of interest is the difference in mean weight, μ_d where d_i =Weight Before − Weight After.

2) $H_0: \mu_d = 0$

3) $H_1: \mu_d > 0$

4) $\alpha = 0.05$

5) The test statistic is

$$t_0 = \frac{\bar{d}}{s_d / \sqrt{n}}$$

6) Reject the null hypothesis if $t_0 > t_{0.05,9}$ where $t_{0.05,9} = 1.833$

7) $\bar{d} = 17$
$s_d = 6.41$
$n = 10$

$$t_0 = \frac{17}{6.41/\sqrt{10}} = 8.387$$

8) Since 8.387 > 1.833 reject the null and conclude there is evidence to conclude that the mean weight loss is significantly greater than 0; that is, the data support the claim that this particular diet modification program is significantly effective in reducing weight at the 0.05 level of significance.

10-42. 1) The parameter of interest is the mean difference in impurity level, μ_d
where d_i = Test 1 − Test 2.
2) $H_0: \mu_d = 0$
3) $H_1: \mu_d \neq 0$
4) $\alpha = 0.01$
5) The test statistic is

$$t_0 = \frac{\bar{d}}{s_d / \sqrt{n}}$$

6) Reject the null hypothesis if $t_0 < -t_{0.005,7}$ where $-t_{0.005,7} = -3.499$ or $t_0 > t_{0.005,7}$ where $t_{0.005,7} = 3.499$

7) $\bar{d} = -0.2125$
$s_d = 0.1727$
$n = 8$

$$t_0 = \frac{-0.2125}{0.1727/\sqrt{8}} = -3.48$$

8) Since −3.48 > -3.499 cannot reject the null and conclude the tests give significantly different impurity levels at α=0.01.

10-43. 1) The parameter of interest is the difference in mean weight loss, μ_d
where d_i = Before − After.
2) $H_0: \mu_d = 10$
3) $H_1: \mu_d > 10$
4) $\alpha = 0.05$
5) The test statistic is

$$t_0 = \frac{\bar{d} - \Delta_0}{s_d / \sqrt{n}}$$

6) Reject the null hypothesis if $t_0 > t_{0.05,9}$ where $t_{0.05,9} = 1.833$
7) $\bar{d} = 17$
$s_d = 6.41$
$n = 10$

$$t_0 = \frac{17 - 10}{6.41/\sqrt{10}} = 3.45$$

8) Since 3.45 > 1.833 reject the null and conclude there is evidence to support the claim that this particular diet modification program is effective in producing a mean weight loss of at least 10 lbs at the 0.05 level of significance.

10-44. Use s_d as an estimate for σ:

$$n = \left(\frac{(z_\alpha + z_\beta)\sigma_d}{10}\right)^2 = \left(\frac{(1.645 + 1.29)6.41}{10}\right)^2 = 3.53, \ n = 4$$

Yes, the sample size of 10 is adequate for this test.

Section 10-5

10-45 a) $f_{0.25,5,10} = 1.59$ d) $f_{0.75,5,10} = \frac{1}{f_{0.25,10,5}} = \frac{1}{1.89} = 0.529$

b) $f_{0.10,24,9} = 2.28$ e) $f_{0.90,24,9} = \frac{1}{f_{0.10,9,24}} = \frac{1}{1.91} = 0.525$

c) $f_{0.05,8,15} = 2.64$ f) $f_{0.95,8,15} = \frac{1}{f_{0.05,15,8}} = \frac{1}{3.22} = 0.311$

10-46 a) $f_{0.25,7,15} = 1.47$ d) $f_{0.75,7,15} = \frac{1}{f_{0.25,15,7}} = \frac{1}{1.68} = 0.596$

b) $f_{0.10,10,12} = 2.19$ e) $f_{0.90,10,12} = \frac{1}{f_{0.10,12,10}} = \frac{1}{2.28} = 0.438$

c) $f_{0.01,20,10} = 4.41$ f) $f_{0.99,20,10} = \frac{1}{f_{0.01,10,20}} = \frac{1}{3.37} = 0.297$

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) $H_1 : \sigma_1^2 \neq \sigma_2^2$

4) $\alpha = 0.05$

5) The test statistic is

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

6) Reject the null hypothesis if $f_0 < f_{0.975,9,15}$ where $f_{0.975,9,15} = 0.265$ or $f_0 > f_{0.025,9,15}$ where $f_{0.025,9,15} = 3.12$

7) $n_1 = 10$ $n_2 = 16$

$s_1 = 4.7$ $s_2 = 5.8$

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

8) Since $0.265 < 0.657 < 3.12$ do not reject the null hypothesis and conclude there is insufficient evidence to indicate the two population variances differ significantly at the 0.05 level of significance.

10-48. 1) The parameters of interest are the etch-rate variances, σ_1^2, σ_2^2.

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

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

4) $\alpha = 0.05$

5) The test statistic is

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

6) Reject the null hypothesis if $f_0 < f_{0.975,9,9} = 0.248$ or $f_0 > f_{0.025,9,9} = 4.03$

7) $n_1 = 10$ $n_2 = 10$

$s_1 = 0.422$ $s_2 = 0.231$

$$f_0 = \frac{(0.422)^2}{(0.231)^2} = 3.337$$

8) Since 0.248 < 3.337 < 4.03 do not reject the null hypothesis and conclude the etch rate variances do not differ at the 0.05 level of significance.

10-49. With $\lambda = \sqrt{2} = 1.4\, \beta = 0.10, \quad and \quad \alpha = 0.05$, we find from Chart VI o that $n_1^* = n_2^* = 100$. Therefore, the samples of size 10 would not be adequate.

10-50. a) 90% confidence interval for the ratio of variances:

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

$$\left(\frac{(0.35)}{(0.40)}\right) 0.412 \leq \frac{\sigma_1^2}{\sigma_2^2} \leq \left(\frac{(0.35)}{(0.40)}\right) 2.33$$

$$0.3605 \leq \frac{\sigma_1^2}{\sigma_2^2} \leq 2.039$$

$$0.6004 \leq \frac{\sigma_1}{\sigma_2} \leq 1.428$$

b) 95% confidence interval:

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

$$\left(\frac{(0.35)}{(0.40)}\right) 0.342 \leq \frac{\sigma_1^2}{\sigma_2^2} \leq \left(\frac{(0.35)}{(0.40)}\right) 2.82$$

$$0.299 \leq \frac{\sigma_1^2}{\sigma_2^2} \leq 2.468$$

$$0.5468 \leq \frac{\sigma_1}{\sigma_2} \leq 1.5710$$

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

c) 90% lower-sided confidence interval:

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

$$\left(\frac{(0.35)}{(0.40)}\right)0.500 \le \frac{\sigma_1^2}{\sigma_2^2}$$

$$0.438 \le \frac{\sigma_1^2}{\sigma_2^2}$$

$$0.661 \le \frac{\sigma_1}{\sigma_2}$$

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:

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

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

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

10-52 1) The parameters of interest are the thickness variances, σ_1^2, σ_2^2

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

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

4) $\alpha = 0.02$

5) The test statistic is

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

6) Reject the null hypothesis if $f_0 < f_{0.99,7,7}$ where $f_{0.99,7,7} = 0.143$ or $f_0 > f_{0.01,7,7}$ where $f_{0.01,7,7} = 6.99$

7) $n_1 = 8$ $n_2 = 8$
$s_1 = 0.11$ $s_2 = 0.09$

$$f_0 = \frac{(0.11)^2}{(0.09)^2} = 1.49$$

8) Since 0.143 < 1.49 < 6.99 do not reject the null hypothesis and conclude the thickness variances do not significantly differ at the 0.02 level of significance.

10-53 1) The parameters of interest are the strength variances, σ_1^2, σ_2^2

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

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

4) $\alpha = 0.05$

5) The test statistic is

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

6) Reject the null hypothesis if $f_0 < f_{0.975,9,15}$ where $f_{0.975,9,15} = 0.265$ or $f_0 > f_{0.025,9,15}$ where $f_{0.025,9,15} = 3.12$

7) $n_1 = 10$ $n_2 = 16$
$s_1 = 12$ $s_2 = 22$

$$f_0 = \frac{(12)^2}{(22)^2} = 0.297$$

8) Since 0.265 < 0.297 < 3.12 do not reject the null hypothesis and conclude the population variances do not significantly differ at the 0.05 level of significance.

10-54 1) The parameters of interest are the melting variances, σ_1^2, σ_2^2

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

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

4) $\alpha = 0.05$

5) The test statistic is

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

6) Reject the null hypothesis if $f_0 < f_{0.975,20,20}$ where $f_{0.975,20,20} = 0.4058$ or $f_0 > f_{0.025,20,20}$ where $f_{0.025,20,20} = 2.46$

7) $n_1 = 21$ $n_2 = 21$
$s_1 = 4$ $s_2 = 3$

$$f_0 = \frac{(4)^2}{(3)^2} = 1.78$$

8) Since 0.4058 < 1.78 < 2.46 do not reject the null hypothesis and conclude the population variances do not significantly differ at the 0.05 level of significance.

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) $H_1: \sigma_1^2 \neq \sigma_2^2$

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 $f_{0.005,10,12} = 5.0855$

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.

10-56. 1) The parameters of interest are the time to assemble standard deviations, σ_1, σ_2

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

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

4) $\alpha = 0.02$

5) The test statistic is

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

6) Reject the null hypothesis if $f_0 < f_{1-\alpha/2, n_1-1, n_2-1}$ =0..365 or $f_0 > f_{\alpha/2, n_1-1, n_2-1} = 2.86$

7) $n_1 = 25$ $n_2 = 21$ $s_1 = 0.98$ $s_2 = 1.02$

$$f_0 = \frac{(0.98)^2}{(1.02)^2} = 0.923$$

8) Since 0.365 < 0.923 < 2.86 do not reject the null hypothesis and conclude there is no evidence to support the claim that men and women differ significantly in repeatability for this assembly task at the 0.02 level of significance.

10-57. 98% confidence interval:

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

$$(0.923)0.365 \le \frac{\sigma_1^2}{\sigma_2^2} \le (0.923)2.86$$

$$0.3369 \le \frac{\sigma_1^2}{\sigma_2^2} \le 2.640$$

Since the value 1 is contained within this interval, we can conclude that there is no significant difference between the variance of the repeatability of men and women for the assembly task at a 98% confidence level.

10-58 For one population standard deviation being 50% larger than the other, then $\lambda = 2$. Using $n = 8$, $\alpha = 0.01$ and Chart VI *p*, we find that $\beta \cong 0.85$. Therefore, we would say that $n = n_1 = n_2 = 8$ is not adequate to detect this difference with high probability.

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 can conclude that there is no significant difference in the variance of the distances at a 95% significance level.

10-60 1) The parameters of interest are the time to assemble standard deviations, σ_1, σ_2

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

3) $H_1: \sigma_1^2 \neq \sigma_2^2$
4) $\alpha = 0.05$
5) The test statistic is

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

6) Reject the null hypothesis if $f_0 < f_{.0.975,11,911} = 0.288$ or $f_0 > f_{0.025,11,11} = 3.474$

7) $n_1 = 12$ $n_2 = 12$ $s_1 = 0.0217$ $s_2 = 0.0175$

$$f_0 = \frac{(0.0217)^2}{(0.0175)^2} = 1.538$$

8) Since 0.288<1.538<3.474 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 coefficient of restitution between the two clubs at the 0.05 level of significance.

95% confidence interval:

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

$$(1.538)0.288 \le \frac{\sigma_1^2}{\sigma_2^2} \le (1.538)3.474$$

$$0.443 \le \frac{\sigma_1^2}{\sigma_2^2} \le 5.343$$

Since the value 1 is contained within this interval, we can conclude that there is no significant difference in the variances in the variances of the coefficient of restitution at a 95% significance level.

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) $H_1: p_1 \neq p_2$
4) $\alpha = 0.05$
5) Test statistic is

$$z_0 = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \quad \text{where}$$

$$\hat{p} = \frac{x_1 + x_2}{n_1 + n_2}$$

6) Reject the null hypothesis if $z_0 < -z_{0.025}$ where $-z_{0.025} = -1.96$ or $z_0 > z_{0.025}$ where $z_{0.025} = 1.96$

7) $n_1 = 300$ $n_2 = 300$
$x_1 = 15$ $x_2 = 8$

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

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

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

10-62. 1) the parameters of interest are the proportion of satisfactory lenses, p_1 and p_2

2) $H_0 : p_1 = p_2$

3) $H_1 : p_1 \neq p_2$

4) $\alpha = 0.05$

5) Test statistic is

$$z_0 = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} \quad \text{where} \quad \hat{p} = \frac{x_1 + x_2}{n_1 + n_2}$$

6) Reject the null hypothesis if $z_0 < -z_{0.005}$ where $-z_{0.005}$ = −258 or $z_0 > z_{0.005}$ where $z_{0.005}$ = 2.58

7) $n_1 = 300$ $n_2 = 300$
$x_1 = 253$ $x_2 = 196$

$\hat{p}_1 = 0.843$ $\hat{p}_2 = 0.653$ $\hat{p} = \frac{253+196}{300+300} = 0.748$

$$z_0 = \frac{0.843 - 0.653}{\sqrt{0.748(1-0.748)\left(\frac{1}{300}+\frac{1}{300}\right)}} = 5.36$$

8) Since 5.36 > 2.58 reject the null hypothesis and conclude that yes the evidence indicates that there is significant difference in the fraction of polishing-induced defects produced by the two polishing solutions the 0.01 level of significance.

P-value = 2(1−P(z < 5.36)) = 0

By constructing a 99% confidence interval on the difference in proportions, the same question can be answered by considering whether or not 0 is contained in the interval.

10-63. a) Power = 1 − β

$$\beta = \Phi\left(\frac{z_{\alpha/2}\sqrt{\overline{p}\,\overline{q}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)} - (p_1 - p_2)}{\hat{\sigma}_{\hat{p}_1 - \hat{p}_2}}\right) - \Phi\left(\frac{-z_{\alpha/2}\sqrt{\overline{p}\,\overline{q}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)} - (p_1 - p_2)}{\hat{\sigma}_{\hat{p}_1 - \hat{p}_2}}\right)$$

$$\overline{p} = \frac{300(0.05)+300(0.01)}{300+300} = 0.03 \qquad \overline{q} = 0.97$$

$$\hat{\sigma}_{\hat{p}_1-\hat{p}_2} = \sqrt{\frac{0.05(1-0.05)}{300}+\frac{0.01(1-0.01)}{300}} = 0.014$$

$$\beta = \Phi\left(\frac{1.96\sqrt{0.03(0.97)\left(\frac{1}{300}+\frac{1}{300}\right)} - (0.05-0.01)}{0.014}\right) - \Phi\left(\frac{-1.96\sqrt{0.03(0.97)\left(\frac{1}{300}+\frac{1}{300}\right)} - (0.05-0.01)}{0.014}\right)$$

$$= \Phi(-0.91) - \Phi(-4.81) = 0.18141 - 0 = 0.18141$$

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

10-64. a) $$\beta = \Phi\left(\frac{z_{\alpha/2}\sqrt{\overline{p}\,\overline{q}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)} - (p_1 - p_2)}{\hat{\sigma}_{\hat{p}_1 - \hat{p}_2}}\right) - \Phi\left(\frac{-z_{\alpha/2}\sqrt{\overline{p}\,\overline{q}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)} - (p_1 - p_2)}{\hat{\sigma}_{\hat{p}_1 - \hat{p}_2}}\right)$$

$$\overline{p} = \frac{300(0.05)+300(0.02)}{300+300} = 0.035 \qquad \overline{q} = 0.965$$

$$\hat{\sigma}_{\hat{p}_1-\hat{p}_2} = \sqrt{\frac{0.05(1-0.05)}{300}+\frac{0.02(1-0.02)}{300}} = 0.015$$

$$\beta = \Phi\left(\frac{1.96\sqrt{0.035(0.965)\left(\frac{1}{300}+\frac{1}{300}\right)} - (0.05-0.02)}{0.015}\right) - \Phi\left(\frac{-1.96\sqrt{0.035(0.965)\left(\frac{1}{300}+\frac{1}{300}\right)} - (0.05-0.02)}{0.015}\right)$$

$$= \Phi(-0.04) - \Phi(-3.96) = 0.48405 - 0.00004 = 0.48401$$

Power = 1 − 0.48401 = 0.51599

b) $$n = \frac{\left(z_{\alpha/2}\sqrt{\frac{(p_1+p_2)(q_1+q_2)}{2}} + z_\beta\sqrt{p_1q_1+p_2q_2}\right)^2}{(p_1-p_2)^2}$$

$$= \frac{\left(1.96\sqrt{\frac{(0.05+0.02)(0.95+0.98)}{2}} + 1.29\sqrt{0.05(0.95)+0.02(0.98)}\right)^2}{(0.05-0.02)^2} = 790.67$$

n = 791

10-65. 1) the parameters of interest are the proportion of residents in favor of an increase, p_1 and p_2

2) $H_0 : p_1 = p_2$

3) $H_1 : p_1 \neq p_2$

4) $\alpha = 0.05$

5) Test statistic is $$z_0 = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$ where $$\hat{p} = \frac{x_1+x_2}{n_1+n_2}$$

6) Reject the null hypothesis if $z_0 < -z_{0.025}$ where $-z_{0.025} = -1.96$ or $z_0 > z_{0.025}$ where $z_{0.025} = 1.96$

7) $n_1 = 500$ $n_2 = 400$

$x_1 = 385$ $x_2 = 267$

$\hat{p}_1 = 0.77$ $\hat{p}_2 = 0.6675$ $$\hat{p} = \frac{385+267}{500+400} = 0.724$$

$$z_0 = \frac{0.77-0.6675}{\sqrt{0.724(1-0.724)\left(\frac{1}{500}+\frac{1}{400}\right)}} = 3.42$$

8) Since 3.42 > 1.96 reject the null hypothesis and conclude that yes the data do indicate a significant difference in the proportions of support for increasing the speed limit between residents of the two counties at the 0.05 level of significance.

P-value = 2(1−P(z < 3.42)) = 0.00062

10-66. 95% confidence interval on the difference:

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

$$(0.05-0.0267)-1.96\sqrt{\frac{0.05(1-0.05)}{300}+\frac{0.0267(1-0.0267)}{300}} \leq p_1-p_2 \leq (0.05-0.0267)+1.96\sqrt{\frac{0.05(1-0.05)}{300}+\frac{0.0267(1-0.0267)}{300}}$$

$-0.0074 \leq p_1 - p_2 \leq 0.054$

Since this interval contains the value zero, we are 95% confident there is no significant difference in the fraction of defective parts produced by the two machines and that the difference in proportions is between −0.0074 and 0.054.

10-67 95% confidence interval on the difference:

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

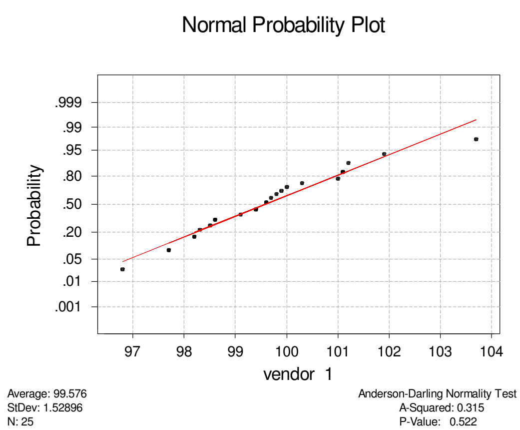
$$(0.77 - 0.6675) - 1.96\sqrt{\frac{0.77(1-0.77)}{500} + \frac{0.6675(1-0.6675)}{400}} \le p_1 - p_2 \le (0.77 - 0.6675) + 1.96\sqrt{\frac{0.77(1-0.77)}{500} + \frac{0.6675(1-0.6675)}{400}}$$

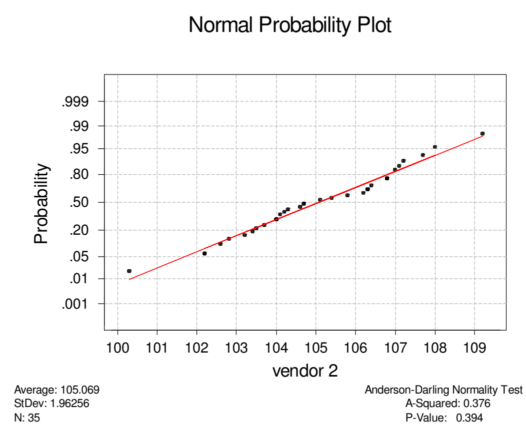
$$0.0434 \le p_1 - p_2 \le 0.1616$$

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

10-68 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) 1) the parameters of interest are the variances of resistance of products, σ_1^2, σ_2^2

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

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

4) $\alpha = 0.05$

5) The test statistic is

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

6) Reject H_0 if $f_0 < f_{0.975,24,34}$ where $f_{0.975,24,34} = \frac{1}{f_{0.025,34,24}} = \frac{1}{2.18} = 0.459$

or $f_0 > f_{0.025,24,34}$ where $f_{0.025,24,34} = 2.07$

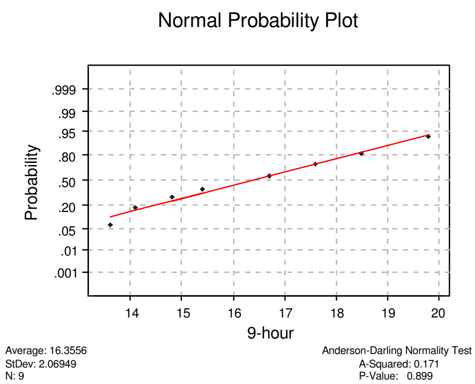
7) $s_1 = 1.53$ $s_2 = 1.96$
$n_1 = 25$ $n_2 = 35$

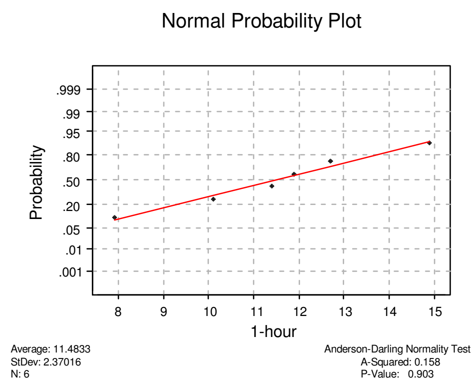
$$f_0 = \frac{(1.53)^2}{(1.96)^2} = 0.609$$

8) Since 0.609 > 0.459, cannot reject H_0 and conclude the variances are significantly different at $\alpha = 0.05$.

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) $\bar{x}_1 = 16.36$ $\quad \bar{x}_2 = 11.483$

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

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

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

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

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

$$(16.36 - 11.483) - 3.012(2.19)\sqrt{\frac{1}{9} + \frac{1}{6}} \le \mu_1 - \mu_2 \le (16.36 - 11.483) + 3.012(2.19)\sqrt{\frac{1}{9} + \frac{1}{6}}$$

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

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).

10-70. a) 95% confidence interval for σ_1^2 / σ_2^2

95% confidence interval on $\frac{\sigma_1^2}{\sigma_2^2}$:

$$f_{0.975,8,5} = \frac{1}{f_{0.025,5,8}} = \frac{1}{4.82} = 0.2075, \qquad f_{0.025,8,5} = 6.76$$

$$\frac{s_1^2}{s_2^2} f_{0.975,8,5} \le \frac{\sigma_1^2}{\sigma_2^2} \le \frac{s_1^2}{s_2^2} f_{0.025,8,5}$$

$$\left(\frac{4.285}{5.617}\right)(0.2075) \le \frac{\sigma_1^2}{\sigma_2^2} \le \left(\frac{4.285}{5.617}\right)(6.76)$$

$$0.1583 \le \frac{\sigma_1^2}{\sigma_2^2} \le 5.157$$

b) Since the value 1 is contained within this interval, with 95% confidence, the population variances do not differ significantly and can be assumed to be equal.

10-71 a) 1) The parameter of interest is the mean weight loss, μ_d
where d_i = Initial Weight − Final Weight.

2) $H_0 : \mu_d = 3$

3) $H_1 : \mu_d > 3$

4) $\alpha = 0.05$

5) The test statistic is

$$t_0 = \frac{\bar{d} - \Delta_0}{s_d / \sqrt{n}}$$

6) Reject H_0 if $t_0 > t_{\alpha,n-1}$ where $t_{0.05,7} = 1.895$.

7) $\bar{d} = 4.125$
$s_d = 1.246$
$n = 8$

$$t_0 = \frac{4.125 - 3}{1.246 / \sqrt{8}} = 2.554$$

8) Since 2.554 > 1.895, reject the null hypothesis and conclude the average weight loss is significantly greater than 3 at $\alpha = 0.05$.

b) 2) $H_0 : \mu_d = 3$

3) $H_1 : \mu_d > 3$

4) $\alpha = 0.01$

5) The test statistic is

$$t_0 = \frac{\bar{d} - \Delta_0}{s_d / \sqrt{n}}$$

6) Reject H_0 if $t_0 > t_{\alpha,n-1}$ where $t_{0.01,7} = 2.998$.

7) $\bar{d} = 4.125$
$s_d = 1.246$
$n = 8$

$$t_0 = \frac{4.125 - 3}{1.246 / \sqrt{8}} = 2.554$$

8) Since 2.554 <2.998, do not reject the null hypothesis and conclude the average weight loss is not significantly greater than 3 at $\alpha = 0.01$.

c) 2) $H_0 : \mu_d = 5$

3) $H_1 : \mu_d > 5$

4) $\alpha = 0.05$

5) The test statistic is

$$t_0 = \frac{\bar{d} - \Delta_0}{s_d / \sqrt{n}}$$

6) Reject H_0 if $t_0 > t_{\alpha,n-1}$ where $t_{0.05,7} = 1.895$.

7) $\bar{d} = 4.125$
$s_d = 1.246$
$n = 8$

$$t_0 = \frac{4.125 - 5}{1.246 / \sqrt{8}} = -1.986$$

8) Since −1.986 < 1.895, do not reject the null hypothesis and conclude the average weight loss is not significantly greater than 5 at α = 0.05.

Using α = 0.01

2) $H_0 : \mu_d = 5$

3) $H_1 : \mu_d > 5$

4) α = 0.01

5) The test statistic is

$$t_0 = \frac{\overline{d} - \Delta_0}{s_d / \sqrt{n}}$$

6) Reject H_0 if $t_0 > t_{\alpha,n-1}$ where $t_{0.01,7} = 2.998$.

7) $\overline{d} = 4.125$

$s_d = 1.246$

$n = 8$

$$t_0 = \frac{4.125 - 5}{1.246 / \sqrt{8}} = -1.986$$

8) Since −1.986 < 2.998, do not reject the null hypothesis and conclude the average weight loss is not significantly greater than 5 at α = 0.01.

10-72. $$(\bar{x}_1 - \bar{x}_2) - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \le \mu_1 - \mu_2 \le (\bar{x}_1 - \bar{x}_2) + z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

a) 90% confidence interval: $z_{\alpha/2} = 1.65$

$$(88 - 91) - 1.65\sqrt{\frac{5^2}{20} + \frac{4^2}{20}} \le \mu_1 - \mu_2 \le (88 - 91) + 1.65\sqrt{\frac{5^2}{20} + \frac{4^2}{20}}$$

$$-5.362 \le \mu_1 - \mu_2 \le -0.638$$

Yes, with 90% confidence, the data indicate that the mean breaking strength of the yarn of manufacturer 2 exceeds that of manufacturer 1 by between 5.362 and 0.638.

b) 98% confidence interval: $z_{\alpha/2} = 2.33$

$$(88 - 91) - 2.33\sqrt{\frac{5^2}{20} + \frac{4^2}{20}} \le \mu_1 - \mu_2 \le (88 - 91) + 2.33\sqrt{\frac{5^2}{20} + \frac{4^2}{20}}$$

$$-6.340 \le \mu_1 - \mu_2 \le 0.340$$

Yes, we are 98% confident manufacturer 2 produces yarn with higher breaking strength by between 0.340 and 6.340 psi.

c) The results of parts a) and b) are different because the confidence level or z-value used is different.. Which one is used depends upon the level of confidence considered acceptable.

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) $\hat{p} = \frac{x_1 + x_2}{n_1 + n_2} = 0.000356$

$\hat{p}_2 = \frac{x_2}{n_2} = \frac{33}{200745} = 0.00016$ (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 $z_{\alpha/2} = 2.58$

$z_0 = 6.55$

Since 6.55 > 2.58, 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.

10-74 a) $\alpha = 0.10$ $z_{\alpha/2} = 1.65$

$$n \cong \frac{(z_{\alpha/2})^2(\sigma_1^2 + \sigma_2^2)}{(E)^2} \cong \frac{(1.65)^2(25+16)}{(1.5)^2} = 49.61, \quad n = 50$$

b) $\alpha = 0.10$ $z_{\alpha/2} = 2.33$

$$n \cong \frac{(z_{\alpha/2})^2(\sigma_1^2 + \sigma_2^2)}{(E)^2} \cong \frac{(2.33)^2(25+16)}{(1.5)^2} = 98.93, \quad n = 99$$

c) As the confidence level increases, sample size will also increase.

d) $\alpha = 0.10$ $z_{\alpha/2} = 1.65$

$$n \cong \frac{(z_{\alpha/2})^2(\sigma_1^2 + \sigma_2^2)}{(E)^2} \cong \frac{(1.65)^2(25+16)}{(0.75)^2} = 198.44, \quad n = 199$$

e) $\alpha = 0.10$ $z_{\alpha/2} = 2.33$

$$n \cong \frac{(z_{\alpha/2})^2(\sigma_1^2 + \sigma_2^2)}{(E)^2} \cong \frac{(2.33)^2(25+16)}{(0.75)^2} = 395.70, \quad n = 396$$

f) As the error decreases, the required sample size increases.

10-75 $\hat{p}_1 = \frac{x_1}{n_1} = \frac{387}{1500} = 0.258$ $\hat{p}_2 = \frac{x_2}{n_2} = \frac{310}{1200} = 0.2583$

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

a) $z_{\alpha/2} = z_{0.025} = 1.96$

$$(0.258 - 0.2583) \pm 1.96\sqrt{\frac{0.258(0.742)}{1500} + \frac{0.2583(0.7417)}{1200}}$$

$$-0.0335 \le p_1 - p_2 \le 0.0329$$

Since zero is contained in this interval, we are 95% confident there is no significant difference between the proportion of unlisted numbers in the two cities.

b) $z_{\alpha/2} = z_{0.05} = 1.65$

$$(0.258 - 0.2583) \pm 1.65\sqrt{\frac{0.258(0.742)}{1500} + \frac{0.2583(0.7417)}{1200}}$$

$-0.0282 \le p_1 - p_2 \le 0.0276$

Again, the proportion of unlisted numbers in the two cities do not differ.

c) $\hat{p}_1 = \frac{x_1}{n_1} = \frac{774}{3000} = 0.258$

$\hat{p}_2 = \frac{x_2}{n_2} = \frac{620}{2400} = 0.2583$

95% confidence interval:

$$(0.258 - 0.2583) \pm 1.96\sqrt{\frac{0.258(0.742)}{3000} + \frac{0.2583(0.7417)}{2400}}$$

$-0.0238 \le p_1 - p_2 \le 0.0232$

90% confidence interval:

$$(0.258 - 0.2583) \pm 1.65\sqrt{\frac{0.258(0.742)}{3000} + \frac{0.2583(0.7417)}{2400}}$$

$$-0.0201 \le p_1 - p_2 \le 0.0195$$

Increasing the sample size decreased the error and width of the confidence intervals, but does not change the conclusions drawn. The conclusion remains that there is no significant difference.

10-76 a) 1) The parameters of interest are the proportions of those residents who wear a seat belt regularly, p_1 , p_2

2) $H_0 : p_1 = p_2$

3) $H_1 : p_1 \ne p_2$

4) $\alpha = 0.05$

5) The test statistic is

$$z_0 = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

6) Reject H_0 if $z_0 < -z_{\alpha/2}$ or $z_0 > z_{\alpha/2}$ where $z_{0.025} = 1.96$

7) $\hat{p}_1 = \frac{x_1}{n_1} = \frac{165}{200} = 0.825$ $\hat{p} = \frac{x_1 + x_2}{n_1 + n_2} = 0.807$

$\hat{p}_2 = \frac{x_2}{n_2} = \frac{198}{250} = 0.792$

$$z_0 = \frac{0.825 - 0.792}{\sqrt{0.807(1 - 0.807)\left(\frac{1}{200} + \frac{1}{250}\right)}} = 0.8814$$

8) Since $-1.96 < 0.8814 < 1.96$ do not reject H_0 and conclude that evidence is insufficient to claim that there is a difference in seat belt usage $\alpha = 0.05$.

b) $\alpha = 0.10$

Reject H_0 if $z_0 < -z_{\alpha/2}$ or $z_0 > z_{\alpha/2}$ where $z_{0.05} = 1.65$

$z_0 = 0.8814$

Since $-1.65 < 0.8814 < 1.65$, do not reject H_0 and conclude that evidence is insufficient to claim that there is a difference in seat belt usage $\alpha = 0.10$.

c) The conclusions are the same, but with different levels of confidence.

d) $n_1 = 400$, $n_2 = 500$

$\alpha = 0.05$

Reject H_0 if $z_0 < -z_{\alpha/2}$ or $z_0 > z_{\alpha/2}$ where $z_{0.025} = 1.96$

$$z_0 = \frac{0.825 - 0.792}{\sqrt{0.807(1-0.807)\left(\frac{1}{400}+\frac{1}{500}\right)}} = 1.246$$

Since $-1.96 < 1.246 < 1.96$ do not reject H_0 and conclude that evidence is insufficient to claim that there is a difference in seat belt usage $\alpha = 0.05$.

$\alpha = 0.10$

Reject H_0 if $z_0 < -z_{\alpha/2}$ or $z_0 > z_{\alpha/2}$ where $z_{0.05} = 1.65$

$z_0 = 1.012$

Since $-1.65 < 1.246 < 1.65$, do not reject H_0 and conclude that evidence is insufficient to claim that there is a difference in seat belt usage $\alpha = 0.10$.

As the sample size increased, the test statistic has also increased, since the denominator of z_0 decreased. However, the decrease (or sample size increase) was not enough to change our conclusion.

10-77. a) Yes, there could be some bias in the results due to the telephone survey.

b) If it could be shown that these populations are similar to the respondents, the results may be extended.

10-78 a) 1) The parameters of interest are the proportion of lenses that are unsatisfactory after tumble-polishing, p_1, p_2

2) $H_0 : p_1 = p_2$

3) $H_1 : p_1 \neq p_2$

4) $\alpha = 0.01$

5) The test statistic is

$$z_0 = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$

6) Reject H_0 if $z_0 < -z_{\alpha/2}$ or $z_0 > z_{\alpha/2}$ where $z_{\alpha/2} = 2.58$

7) x_1 = number of defective lenses

$$\hat{p}_1 = \frac{x_1}{n_1} = \frac{47}{300} = 0.1567 \qquad \hat{p} = \frac{x_1 + x_2}{n_1 + n_2} = 0.2517$$

$$\hat{p}_2 = \frac{x_2}{n_2} = \frac{104}{300} = 0..3467$$

$$z_0 = \frac{0.1567 - 0.3467}{\sqrt{0.2517(1-0.2517)\left(\frac{1}{300}+\frac{1}{300}\right)}} = -5.36$$

8) Since −5.36 < −2.58 reject H_0 and conclude there is strong evidence to support the claim that the two polishing fluids are different.

b) The conclusions are the same whether we analyze the data using the proportion unsatisfactory or proportion satisfactory. The proportion of defectives are different for the two fluids.

10-79.

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

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

$$n = 60$$

10-80 The parameter of interest is $\mu_1 - 2\mu_2$

$H_0: \mu_1 = 2\mu_2$ → $H_0: \mu_1 - 2\mu_2 = 0$

$H_1: \mu_1 > 2\mu_2$ $\quad$ $H_1: \mu_1 - 2\mu_2 > 0$

Let n_1 = size of sample 1 $\quad$ $\overline{X}_1$ estimate for μ_1

Let n_2 = size of sample 2 $\quad$ $\overline{X}_2$ estimate for μ_2

$\overline{X}_1 - 2\overline{X}_2$ is an estimate for $\mu_1 - 2\mu_2$

The variance is $V(\overline{X}_1 - 2\overline{X}_2) = V(\overline{X}_1) + V(2\overline{X}_2) = \frac{\sigma_1^2}{n_1} + \frac{4\sigma_2^2}{n_2}$

The test statistic for this hypothesis would then be:

$$Z_0 = \frac{(\overline{X}_1 - 2\overline{X}_2) - 0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{4\sigma_2^1}{n_2}}}$$

We would reject the null hypothesis if $z_0 > z_{\alpha/2}$ for a given level of significance.

The P-value would be $P(Z \ge z_0)$.

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, $n^* = 50$

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

$n_1 = n_2 = 26$

10-82 a) $\alpha = 0.05$, $\beta = 0.05$ $\Delta = 1.5$ Use $s_p = 0.7071$ to approximate σ in equation 10-19.

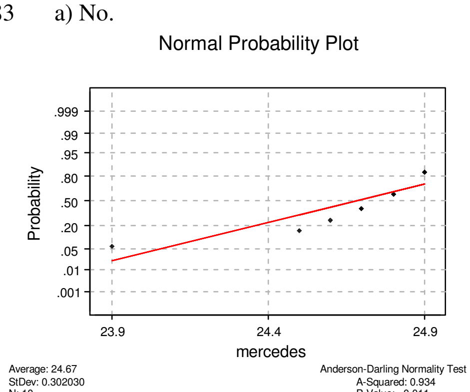
$$d = \frac{\Delta}{2(s_p)} = \frac{1.5}{2(.7071)} = 1.06 \cong 1$$

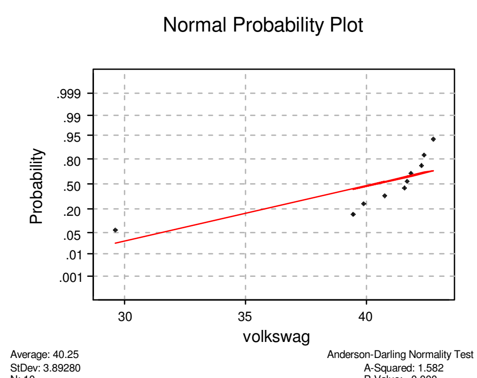
From Chart VI (e), $n^* = 20$ $n = \frac{n^* + 1}{2} = \frac{20 + 1}{2} = 10.5$

n= 11 would be needed to reject the null hypothesis that the two agents differ by 0.5 with probability of at least 0.95.

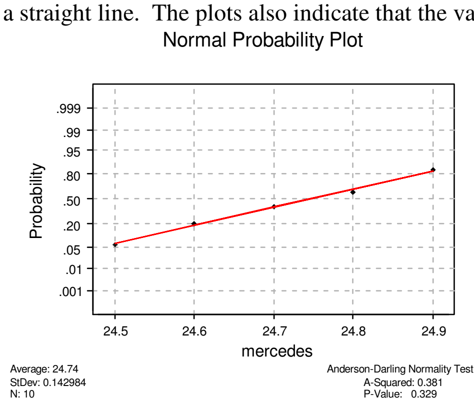
b) The original size of n = 5 in Exercise 10-18 was not appropriate to detect the difference since it is necessary for a sample size of 16 to reject the null hypothesis that the two agents differ by 1.5 with probability of at least 0.95.

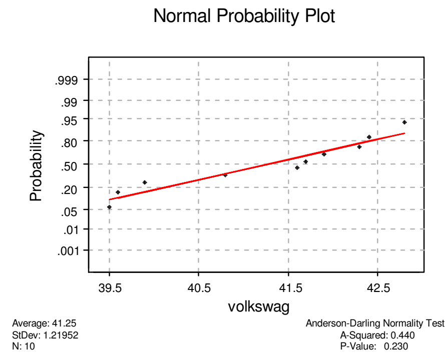
10-83 a) No.





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

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

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

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

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

$$18.114 < \frac{\sigma_V^2}{\sigma_M^2} < 294.35$$

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 the same for the two types of vehicles. There is evidence that the variability is greater for a Volkswagen than for a Mercedes.

10-84 1) the parameters of interest are the variances in mileage performance, σ_1^2, σ_2^2

2) $H_0 : \sigma_1^2 = \sigma_2^2$ Where Volkswagen is represented by variance 1, Mercedes by variance 2.

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

4) $\alpha = 0.05$

5) The test statistic is

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

6) Reject H_0 if $f_0 < f_{0.975,9,9}$ where $f_{0.975,9,9} = \dfrac{1}{f_{0.025,9,9}} = \dfrac{1}{4.03} = 0.248$

or $f_0 > f_{0.025,9,9}$ where $f_{0.025,9,9} = 4.03$

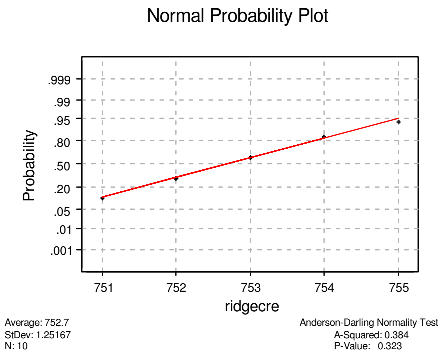
7) $s_1 = 1.22$ $s_2 = 0.143$

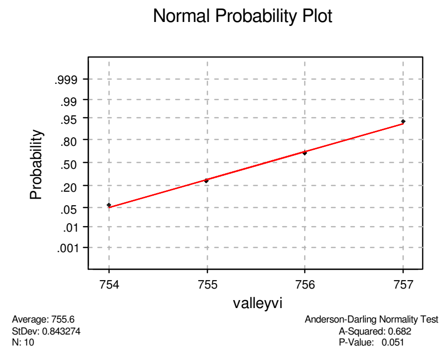
$n_1 = 10$ $n_2 = 10$

$$f_0 = \frac{(1.22)^2}{(0.143)^2} = 72.78$$

8) Since 72.78 > 4.03, reject H_0 and conclude that there is a significant difference between Volkswagen and Mercedes in terms of mileage variability. Same conclusions reached in 10-83d.

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

$s_1 = 1.252$ $s_2 = 0.843$

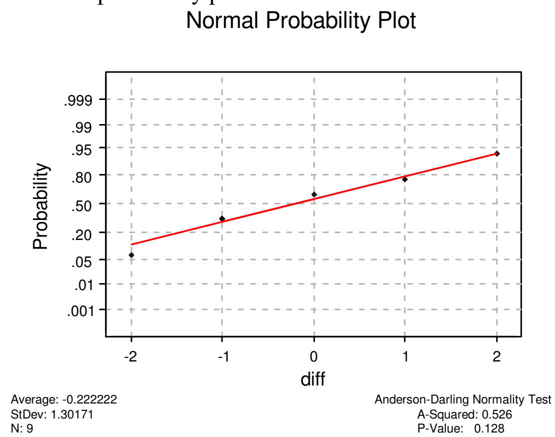
$n_1 = 10$ $n_2 = 10$

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

8) Since −6.06 < −2.101, reject H_0 and conclude there is a significant difference between the two winery's with respect to mean fill volumes.

10-86. d=2/2(1.07)=0.93, giving a power of just under 80%. Since the power is relatively low, an increase in the sample size would increase the power of the test.

10-87. a) The assumption of normality appears to be valid. This is evident by the fact that the data lie along a straight line in the normal probability plot.



b) 1) The parameter of interest is the mean difference in tip hardness, μ_d

2) $H_0 : \mu_d = 0$

3) $H_1 : \mu_d \neq 0$

4) No significance level, calculate P-value

5) The test statistic is

$$t_0 = \frac{\bar{d}}{s_d / \sqrt{n}}$$

6) Reject H_0 if the P-value is significantly small.

7) $\bar{d} = -0.222$

$s_d = 1.30$

$n = 9$

$$t_0 = \frac{-0.222}{1.30 / \sqrt{9}} = -0.512$$

8) P-value = 2P(T < -0.512) = 2P(T > 0.512) $\quad 2(0.25) < \text{P-value} < 2(0.40)$

$0.50 < \text{P-value} < 0.80$

Since the P-value is larger than any acceptable level of significance, do not reject H_0 and conclude there is no difference in mean tip hardness.

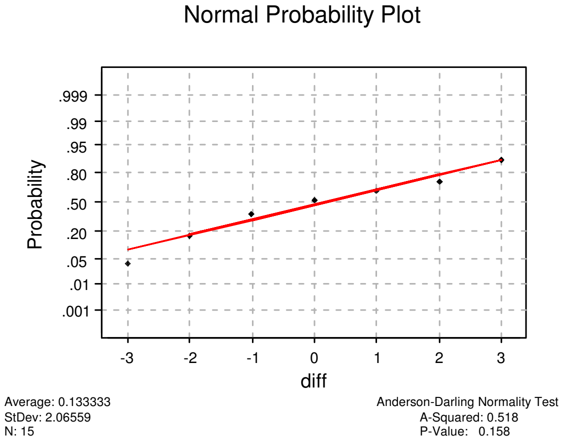
c) $\beta = 0.10$

$\mu_d = 1$

$$d = \frac{1}{\sigma_d} = \frac{1}{1.3} = 0.769$$

From Chart VI with $\alpha = 0.01$, $n = 30$

10-88. a) According to the normal probability plot the data appear to follow a normal distribution. This is evident by the fact that the data fall along a straight line.



b) 1) The parameter of interest is the mean difference in depth using the two gauges, μ_d

2) $H_0 : \mu_d = 0$

3) $H_1 : \mu_d \neq 0$

4) No significance level, calculate p-value

5) The test statistic is

$$t_0 = \frac{\bar{d}}{s_d / \sqrt{n}}$$

6) Reject H_0 if the P-value is significantly small.

7) $\bar{d} = 0.133$

$s_d = 2.065$

$n = 15$

$$t_0 = \frac{0.133}{2.065/\sqrt{15}} = 0.25$$

8) P-value = 2P(T > 0.25) $\quad$ 2(0.40) < P-value

0.80 < P-value

Since the P-value is larger than any acceptable level of significance, do not reject H_0 and conclude there is no difference in mean depth measurements for the two gauges.

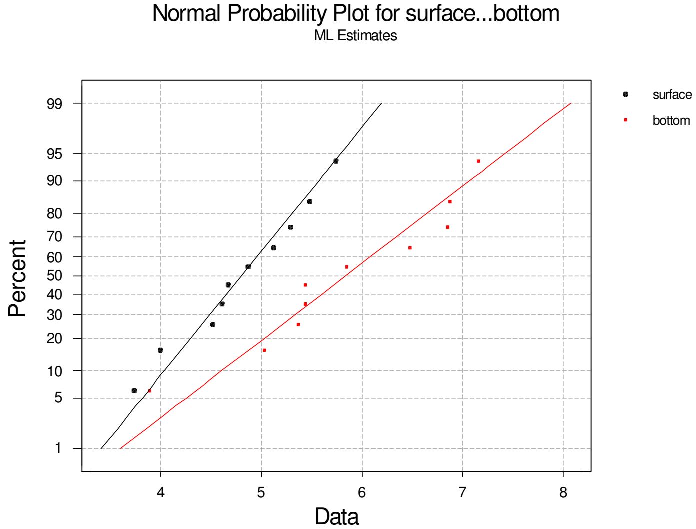
c) Power = 0.8, Therefore, since Power= 1-β , β= 0.20

$$\mu_d = 1.65$$

$$d = \frac{1.65}{\sigma_d} = \frac{1.65}{(2.065)} = 0.799$$

From Chart VI (f) with α = 0.01 and β = 0.20, we find n =30.

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

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

6) Reject the null hypothesis if $t_0 < -t_{0.025,15}$ where $-t_{0.025,15} = -2.131$ or $t_0 > t_{0.025,15}$ where $t_{0.025,15} =$ 2.131 since

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

$\nu \cong 15$

(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) Assuming the sample sizes were equal:

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

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

d.)Assuming the sample sizes were equal:

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

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

Mind-Expanding Exercises

10-90 The estimate of μ is given by $\bar{X}$. Therefore, $\bar{X} = \frac{1}{2}\left(\bar{X}_1 + \bar{X}_2\right) - \bar{X}_3$. The variance of $\bar{X}$ can be shown to be: $V(\bar{X}) = \frac{1}{4}\left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right) + \frac{\sigma_3^2}{n_3}$. Using s_1, s_2, and s_3 as estimates for σ_1 σ_2 and σ_3 respectively.

a) A 100(1-α)% confidence interval on μ is then:

$$\bar{X} - Z_{\alpha/2}\sqrt{\frac{1}{4}\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right) + \frac{s_3^2}{n_3}} \le \mu \le \bar{X} + Z_{\alpha/2}\sqrt{\frac{1}{4}\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right) + \frac{s_3^2}{n_3}}$$

b) A 95% confidence interval for μ is

$$\left(\frac{1}{2}(4.6+5.2)-6.1\right)-1.96\sqrt{\frac{1}{4}\left(\frac{0.7^2}{100}+\frac{0.6^2}{120}\right)+\frac{0.8^2}{130}} \le \mu \le \left(\frac{1}{2}(4.6+5.2)-6.1\right)+1.96\sqrt{\frac{1}{4}\left(\frac{0.7^2}{100}+\frac{0.6^2}{120}\right)+\frac{0.8^2}{130}}$$

$$-1.2-0.163 \le \mu \le -1.2+0.163$$

$$-1.363 \le \mu \le -1.037$$

Since zero is not contained in this interval, and because the possible differences (-1.363, -1.037) are negative, we can conclude that there is sufficient evidence to indicate that pesticide three is more effective.

10-91 The $V(\overline{X}_1 - \overline{X}_2) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$ and suppose this is to equal a constant k. Then, we are to minimize $C_1 n_1 + C_2 n_2$ subject to $\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} = k$. Using a Lagrange multiplier, we minimize by setting the partial derivatives of $f(n_1, n_2, \lambda) = C_1 n_1 + C_2 n_2 + \lambda\left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} - k\right)$ with respect to n_1, n_2 and λ equal to zero.

These equations are

$$\frac{\partial}{\partial n_1} f(n_1, n_2, \lambda) = C_1 - \frac{\lambda \sigma_1^2}{n_1^2} = 0 \qquad (1)$$

$$\frac{\partial}{\partial n_2} f(n_1, n_2, \lambda) = C_2 - \frac{\lambda \sigma_2^2}{n_2^2} = 0 \qquad (2)$$

$$\frac{\partial}{\partial \lambda} f(n_1, n_2, \lambda) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} = k \qquad (3)$$

Upon adding equations (1) and (2), we obtain $C_1 + C_2 - \lambda\left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right) = 0$

Substituting from equation (3) enables us to solve for λ to obtain $\frac{C_1 + C_2}{k} = \lambda$

Then, equations (1) and (2) are solved for n_1 and n_2 to obtain

$$n_1 = \frac{\sigma_1^2 (C_1 + C_2)}{kC_1} \qquad n_2 = \frac{\sigma_2^2 (C_1 + C_2)}{kC_2}$$

It can be verified that this is a minimum and that with these choices for n_1 and n_2.

$$V(\overline{X}_1 - \overline{X}_2) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}.$$

10-92 Maximizing the probability of rejecting H_0 is equivalent to minimizing

$$P\left(-z_{\alpha/2} < \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} < z_{\alpha/2} \mid \mu_1 - \mu_2 = \delta\right) = P\left(-z_{\alpha/2} - \frac{\delta}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} < Z < z_{\alpha/2} - \frac{\delta}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}\right)$$

where z is a standard normal random variable. This probability is minimized by maximizing $\frac{\delta}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}$.

Therefore, we are to minimize $\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$ subject to $n_1 + n_2 = N$.

From the constraint, $n_2 = N - n_1$, and we are to minimize $f(n_1) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{N - n_1}$. Taking the derivative of $f(n_1)$ with respect to n_1 and setting it equal to zero results in the equation $\frac{-\sigma_1^2}{n_1^2} + \frac{\sigma_2^2}{(N-n_1)^2} = 0$.

Upon solving for n_1, we obtain $n_1 = \frac{\sigma_1 N}{\sigma_1 + \sigma_2}$ and $n_2 = \frac{\sigma_2 N}{\sigma_1 + \sigma_2}$

Also, it can be verified that the solution minimizes $f(n_1)$.

10-93 a) $\alpha = P(Z > z_\varepsilon \ or \ Z < -z_{\alpha-\varepsilon})$ where Z has a standard normal distribution.

Then, $\alpha = P(Z > z_\varepsilon) + P(Z < -z_{\alpha-\varepsilon}) = \varepsilon + \alpha - \varepsilon = \alpha$

b) $\beta = P(-z_{\alpha-\varepsilon} < Z_0 < z_\varepsilon | \mu_1 = \mu_0 + \delta)$

$$\beta = P(-z_{\alpha-\varepsilon} < \frac{\bar{x} - \mu_0}{\sqrt{\sigma^2/n}} < z_\varepsilon | \mu_1 = \mu_0 + \delta)$$

$$= P(-z_{\alpha-\varepsilon} - \frac{\delta}{\sqrt{\sigma^2/n}} < Z < z_\varepsilon - \frac{\delta}{\sqrt{\sigma^2/n}})$$

$$= \Phi(z_\varepsilon - \frac{\delta}{\sqrt{\sigma^2/n}}) - \Phi(-z_{\alpha-\varepsilon} - \frac{\delta}{\sqrt{\sigma^2/n}})$$

10-94. The requested result can be obtained from data in which the pairs are very different. Example:

pair	1	2	3	4	5
sample 1	100	10	50	20	70
sample 2	110	20	59	31	80

$\bar{x}_1 = 50 \quad \bar{x}_2 = 60$

$s_1 = 36.74 \qquad s_2 = 36.54 \qquad s_{pooled} = 36.64$

Two-sample t-test : $t_0 = -0.43$ P-value = 0.68

$\bar{x}_d = -10 \qquad s_d = 0.707$

Paired t-test : $t_0 = -31.62$ P-value ≈ 0

10-95 a.) $\theta = \dfrac{p_1}{p_2}$ and $\hat{\theta} = \dfrac{\hat{p}_1}{\hat{p}_2}$ and $\ln(\hat{\theta}) \sim N[\ln(\theta), \sqrt{(n_1 - x_1)/n_1x_1 + (n_2 - x_2)/n_2x_2}\,]$

The (1-α) confidence Interval for ln(θ) will use the relationship $Z = \dfrac{\ln(\hat{\theta}) - \ln(\theta)}{\left(\left(\dfrac{n_1 - x_1}{n_1x_1}\right) + \left(\dfrac{n_2 - x_2}{n_2x_2}\right)\right)^{1/4}}$

$$\ln(\hat{\theta}) - Z_{\alpha/2}\left(\left(\frac{n_1 - x_1}{n_1x_1}\right) + \left(\frac{n_2 - x_2}{n_2x_2}\right)\right)^{1/4} \le \ln(\theta) \le \ln(\hat{\theta}) + Z_{\alpha/2}\left(\left(\frac{n_1 - x_1}{n_1x_1}\right) + \left(\frac{n_2 - x_2}{n_2x_2}\right)\right)^{1/4}$$

b.) The (1-α) confidence Interval for θ use the CI developed in part (a.) where θ = e^(ln(θ))

$$\hat{\theta} - e^{Z_{\alpha/2}\left(\left(\frac{n_1 - x_1}{n_1x_1}\right) + \left(\frac{n_2 - x_2}{n_2x_2}\right)\right)^{1/4}} \le \theta \le \hat{\theta} + e^{Z_{\alpha/2}\left(\left(\frac{n_1 - x_1}{n_1x_1}\right) + \left(\frac{n_2 - x_2}{n_2x_2}\right)\right)^{1/4}}$$

c.)

$$\hat{\theta} - e^{Z_{\alpha/2}\left(\left(\frac{n_1 - x_1}{n_1x_1}\right) + \left(\frac{n_2 - x_2}{n_2x_2}\right)\right)^{.25}} \le \theta \le \hat{\theta} + e^{Z_{\alpha/2}\left(\left(\frac{n_1 - x_1}{n_1x_1}\right) + \left(\frac{n_2 - x_2}{n_2x_2}\right)\right)^{.25}}$$

$$1.42 - e^{1.96\left(\left(\frac{100-27}{2700}\right) + \left(\frac{100-19}{1900}\right)\right)^{1/4}} \le \theta \le 1.42 + e^{1.96\left(\left(\frac{100-27}{2700}\right) + \left(\frac{100-19}{1900}\right)\right)^{1/4}}$$

$$-1.317 \le \theta \le 4.157$$

Since the confidence interval contains the value 1, we conclude that there is no difference in the proportions at the 95% level of significance

10-96 $H_0 : \sigma_1^2 = \sigma_2^2$

$H_1 : \sigma_1^2 \neq \sigma_2^2$

$$\beta = P\left(f^{2}_{1-\alpha/2, n_1-1, n_2-1} < \frac{S_1^2}{S_2^2} < f^{2}_{\alpha/2, n_1-1, n_2-1} \mid \frac{\sigma_1^2}{\sigma_2^2} = \delta \neq 1 \right)$$

$$= P\left(\frac{\sigma_2^2}{\sigma_1^2} f_{1-\alpha/2, n_1-1, n_2-1} < \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} < \frac{\sigma_2^2}{\sigma_1^2} f_{\alpha/2, n_1-1, n_2-1} \mid \frac{\sigma_1^2}{\sigma_2^2} = \delta \right)$$

where $\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}$ has an F distribution with $n_1 - 1$ and $n_2 - 1$ degrees of freedom.

CHAPTER 11

Section 11-2

11-1. a) $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$

$$S_{xx} = 157.42 - \tfrac{43^2}{14} = 25.348571$$

$$S_{xy} = 1697.80 - \tfrac{43(572)}{14} = -59.057143$$

$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = \frac{-59.057143}{25.348571} = -2.330$$

$$\hat{\beta}_0 = \overline{y} - \hat{\beta}_1 \overline{x} = \tfrac{572}{14} - (-2.3298017)(\tfrac{43}{14}) = 48.013$$

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

11-2. a) $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$

$$S_{xx} = 143215.8 - \tfrac{1478^2}{20} = 33991.6$$

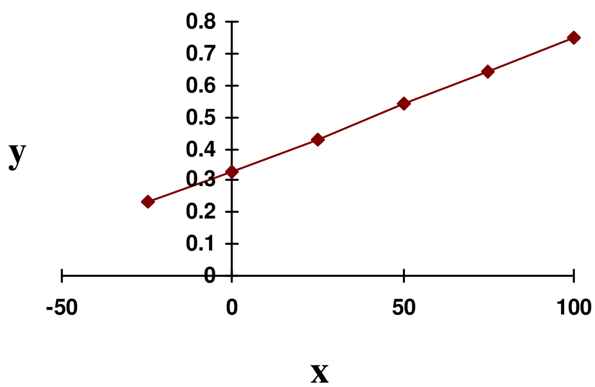
$$S_{xy} = 1083.67 - \tfrac{(1478)(12.75)}{20} = 141.445$$

$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = \frac{141.445}{33991.6} = 0.00416$$

$$\hat{\beta}_0 = \tfrac{12.75}{20} - (0.0041617512)(\tfrac{1478}{20}) = 0.32999$$

$$\hat{y} = 0.32999 + 0.00416x$$

$$\hat{\sigma}^2 = MS_E = \frac{SS_E}{n-2} = \frac{0.143275}{18} = 0.00796$$



b) $\hat{y} = 0.32999 + 0.00416(85) = 0.6836$

c) $\hat{y} = 0.32999 + 0.00416(90) = 0.7044$

d) $\hat{\beta}_1 = 0.00416$

11-3. a) $\hat{y} = 0.3299892 + 0.0041612(\frac{9}{5}x + 32)$

$\hat{y} = 0.3299892 + 0.0074902x + 0.1331584$

$\hat{y} = 0.4631476 + 0.0074902x$

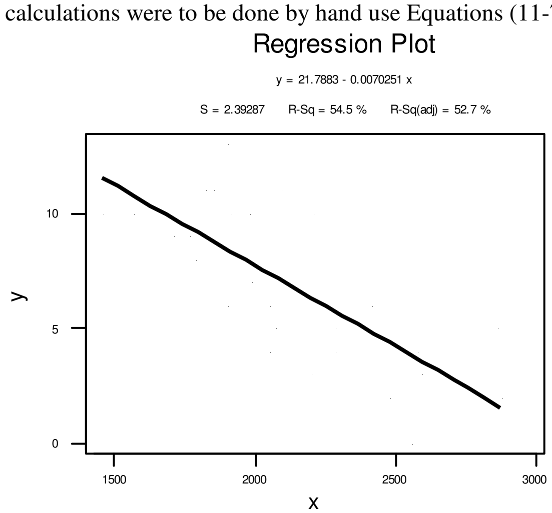
b) $\hat{\beta}_1 = 0.00749$

11-4. a)

```
Regression Analysis - Linear model: Y = a+bX
Dependent variable: Games                      Independent variable: Yards
--------------------------------------------------------------------------------
                               Standard             T              Prob.
Parameter       Estimate         Error            Value            Level
Intercept        21.7883        2.69623           8.081           .00000
Slope         -7.0251E-3     1.25965E-3        -5.57703           .00001
--------------------------------------------------------------------------------
                           Analysis of Variance
Source             Sum of Squares     Df  Mean Square    F-Ratio  Prob. Level
Model                   178.09231      1    178.09231    31.1032       .00001
Residual                148.87197     26      5.72585
--------------------------------------------------------------------------------
Total (Corr.)           326.96429     27
Correlation Coefficient = -0.738027          R-squared =  54.47 percent
Stnd. Error of Est. = 2.39287
```

$\hat{\sigma}^2 = 5.7258$

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



b) $\hat{y} = 21.7883 - 0.0070251(1800) = 9.143$

c) −0.0070251(-100) = 0.70251 games won.

d) $\frac{1}{0.0070251} = 142.35$ yds decrease required.

e) $\hat{y} = 21.7883 - 0.0070251(1917) = 8.321$

$e = y - \hat{y}$

$= 10 - 8.321 = 1.679$

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).

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

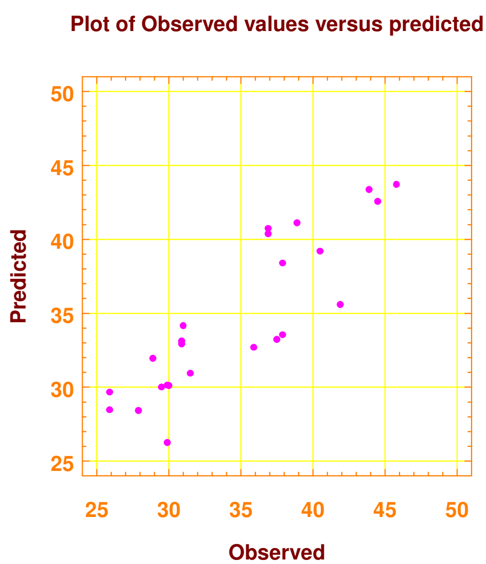
b) $\hat{y} = 13.3202 + 3.32437(7.5) = 38.253$

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.



11-6. a)

```
Regression Analysis - Linear model: Y = a+bX
Dependent variable: Usage              Independent variable: Temperature
--------------------------------------------------------------------------------
                               Standard            T            Prob.
Parameter       Estimate        Error            Value          Level
Intercept        -6.3355        1.66765        -3.79906         .00349
Slope            9.20836      0.0337744         272.643         .00000
--------------------------------------------------------------------------------
                           Analysis of Variance
Source             Sum of Squares     Df  Mean Square    F-Ratio  Prob. Level
Model                   280583.12      1    280583.12    74334.4       .00000
Residual                37.746089     10     3.774609
--------------------------------------------------------------------------------
Total (Corr.)           280620.87     11
Correlation Coefficient = 0.999933          R-squared =  99.99 percent
Stnd. Error of Est. = 1.94284
```

$$\hat{\sigma}^2 = 3.7746$$

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

$$\hat{y} = -6.3355 + 9.20836x$$

b) $\hat{y} = -6.3355 + 9.20836(55) = 500.124$

c) If monthly temperature increases by 1°F, $\hat{y}$ increases by 9.20836.

d) $\hat{y} = -6.3355 + 9.20836(47) = 426.458$

$\hat{y} = 426.458$

$$e = y - \hat{y} = 424.84 - 426.458 = -1.618$$

11-7. a)

```
Predictor       Coef       StDev          T        P
Constant      33.535       2.614      12.83    0.000
x           -0.03540     0.01663      -2.13    0.047

S = 3.660       R-Sq = 20.1%     R-Sq(adj) = 15.7%

Analysis of Variance
Source       DF          SS          MS         F        P
Regression    1       60.69       60.69      4.53    0.047
Error        18      241.06       13.39
Total        19      301.75
```

$$\hat{\sigma}^2 = 13.392$$

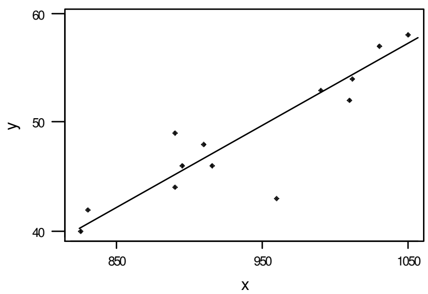
$$\hat{y} = 33.5348 - 0.0353971x$$

b) $\hat{y} = 33.5348 - 0.0353971(150) = 28.226$

c) $\hat{y} = 29.4995$

$$e = y - \hat{y} = 31.0 - 29.4995 = 1.50048$$

11-8. a)



Predictor	Coef	StDev	T	P
Constant	-16.509	9.843	-1.68	0.122
x	0.06936	0.01045	6.64	0.000

S = 2.706 R-Sq = 80.0% R-Sq(adj) = 78.2%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	322.50	322.50	44.03	0.000
Error	11	80.57	7.32		
Total	12	403.08			

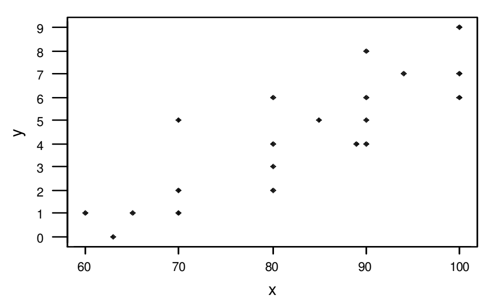
$\hat{\sigma}^2 = 7.3212$

$\hat{y} = -16.5093 + 0.0693554x$

b) $\hat{y} = 46.6041$ $\quad e = y - \hat{y} = 1.39592$

c) $\hat{y} = -16.5093 + 0.0693554(950) = 49.38$

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	-10.132	1.995	-5.08	0.000
x	0.17429	0.02383	7.31	0.000

S = 1.318 R-Sq = 74.8% R-Sq(adj) = 73.4%

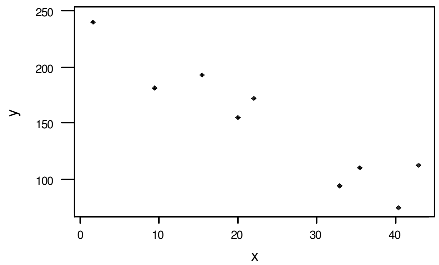
Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	92.934	92.934	53.50	0.000
Residual Error	18	31.266	1.737		
Total	19	124.200			

b) $\hat{\sigma}^2 = 1.737$ and $\hat{y} = -10.132 + 0.17429x$

c) $\hat{y} = 4.68265$ at $x = 85$

11-10. a)



Yes, a linear regression model appears to be plausable.

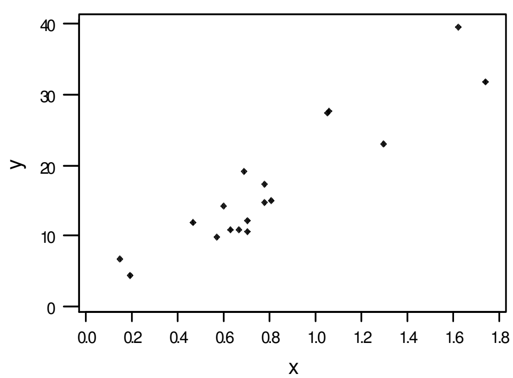
```
Predictor       Coef       StDev          T        P
Constant      234.07       13.75      17.03    0.000
x            -3.5086      0.4911      -7.14    0.000
S = 19.96       R-Sq = 87.9%     R-Sq(adj) = 86.2%
Analysis of Variance
Source        DF          SS          MS         F        P
Regression     1       20329       20329     51.04    0.000
Error          7        2788         398
Total          8       23117
```

b) $\hat{\sigma}^2 = 398.25$ and $\hat{y} = 234.071 - 3.50856x$

c) $\hat{y} = 234.071 - 3.50856(30) = 128.814$

d) $\hat{y} = 156.883 \quad e = 15.1175$

11-11. a)



Yes, a simple linear regression model seems appropriate for these data.

```
Predictor       Coef       StDev          T        P
Constant       0.470       1.936       0.24    0.811
x             20.567       2.142       9.60    0.000
S = 3.716       R-Sq = 85.2%     R-Sq(adj) = 84.3%

Analysis of Variance

Source        DF          SS          MS         F        P
Regression     1      1273.5      1273.5     92.22    0.000
Error         16       220.9        13.8
Total         17      1494.5
```

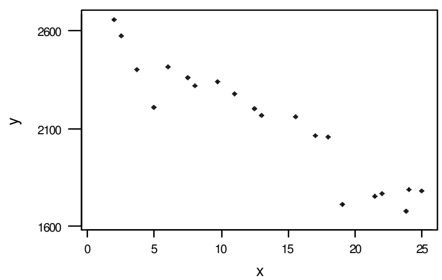
b) $\hat{\sigma}^2 = 13.81$

$\hat{y} = 0.470467 + 20.5673x$

c) $\hat{y} = 0.470467 + 20.5673(1) = 21.038$

d) $\hat{y} = 10.1371 \quad e = 1.6629$

11-12. a)



Yes, a simple linear regression (straight-line) model seems plausible for this situation.

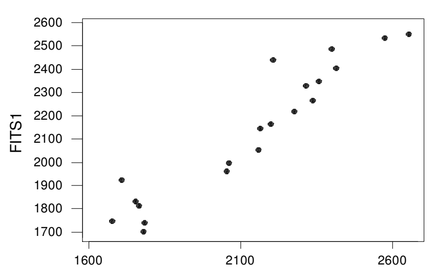
```
Predictor       Coef       StDev           T        P
Constant     2625.39       45.35       57.90    0.000
x            -36.962       2.967      -12.46    0.000

S = 99.05       R-Sq = 89.6%     R-Sq(adj) = 89.0%

Analysis of Variance
Source       DF          SS          MS         F        P
Regression    1     1522819     1522819    155.21    0.000
Error        18      176602        9811
Total        19     1699421
```

b) $\hat{\sigma}^2 = 9811.2$

$\hat{y} = 2625.39 - 36.962x$

c) $\hat{y} = 2625.39 - 36.962(20) = 1886.15$

d) If there were no error, the values would all lie along the 45° axis. The plot indicates age was reasonable regressor variable.



11-13. $\hat{\beta}_0 + \hat{\beta}_1\bar{x} = (\bar{y} - \hat{\beta}_1\bar{x}) + \hat{\beta}_1\bar{x} = \bar{y}$

11-14. a) The slopes of both regression models will be the same, but the intercept will be shifted.

b) $\hat{y} = 2132.41 - 36.9618x$

$$\hat{\beta}_0 = 2625.39 \qquad \text{vs.} \qquad \hat{\beta}_0^* = 2132.41$$
$$\hat{\beta}_1 = -36.9618 \qquad\qquad \hat{\beta}_1^* = -36.9618$$

11-15. Let $x_i^* = x_i - \bar{x}$. Then, the model is $Y_i^* = \beta_0^* + \beta_1^* x_i^* + \varepsilon_i$.

Equations 11-7 and 11-8 can be applied to the new variables using the facts that $\sum_{i=1}^{n} x_i^* = \sum_{i=1}^{n} y_i^* = 0$. Then,

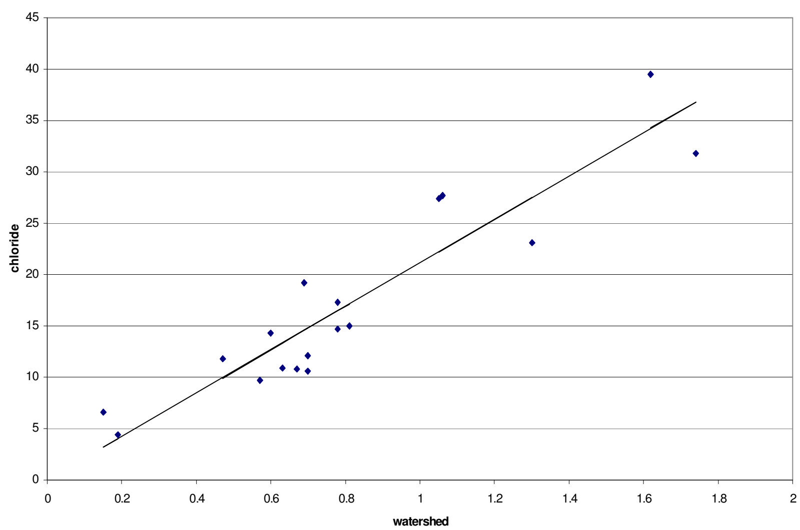
$\hat{\beta}_1^* = \hat{\beta}_1$ and $\hat{\beta}_0^* = 0$.

11-16. The least squares estimate minimizes $\sum (y_i - \beta x_i)^2$. Upon setting the derivative equal to zero, we obtain

$$2\sum (y_i - \beta x_i)\ (-x_i) = 2[\sum y_i x_i - \beta \sum x_i^2] = 0$$

Therefore, $\hat{\beta} = \dfrac{\sum y_i x_i}{\sum x_i^2}$.

11-17. $\hat{y} = 21.031461x$. The model seems very appropriate - an even better fit.



Section 11-5

11-18. a) 1) The parameter of interest is the regressor variable coefficient, β_1

2) $H_0: \beta_1 = 0$

3) $H_1: \beta_1 \neq 0$

4) $\alpha = 0.05$

5) The test statistic is

$$f_0 = \frac{MS_R}{MS_E} = \frac{SS_R / 1}{SS_E / (n-2)}$$

6) Reject H_0 if $f_0 > f_{\alpha,1,12}$ where $f_{0.05,1,12} = 4.75$

7) Using results from Exercise 11-1

$$SS_R = \hat{\beta}_1 S_{xy} = -2.3298017(-59.057143)$$
$$= 137.59$$
$$SS_E = S_{yy} - SS_R$$
$$= 159.71429 - 137.59143$$
$$= 22.123$$

$$f_0 = \frac{137.59}{22.123/12} = 74.63$$

8) Since 74.63 > 4.75 reject H_0 and conclude that compressive strength is significant in predicting intrinsic permeability of concrete at $\alpha = 0.05$. We can therefore conclude model specifies a useful linear relationship between these two variables.

P − value ≅ 0.000002

b) $\hat{\sigma}^2 = MS_E = \frac{SS_E}{n-2} = \frac{22.123}{12} = 1.8436$ and $se(\hat{\beta}_1) = \sqrt{\frac{\hat{\sigma}^2}{S_{xx}}} = \sqrt{\frac{1.8436}{25.3486}} = 0.2696$

c) $se(\hat{\beta}_0) = \sqrt{\hat{\sigma}^2\left[\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right]} = \sqrt{1.8436\left[\frac{1}{14} + \frac{3.0714^2}{25.3486}\right]} = 0.9043$

11-19. a) 1) The parameter of interest is the regressor variable coefficient, β_1.

2) $H_0: \beta_1 = 0$

3) $H_1: \beta_1 \neq 0$

4) $\alpha = 0.05$

5) The test statistic is

$$f_0 = \frac{MS_R}{MS_E} = \frac{SS_R / 1}{SS_E / (n-2)}$$

6) Reject H_0 if $f_0 > f_{\alpha,1,18}$ where $f_{0.05,1,18} = 4.414$

7) Using the results from Exercise 11-2

$$SS_R = \hat{\beta}_1 S_{xy} = (0.0041612)(141.445)$$
$$= 0.5886$$
$$SS_E = S_{yy} - SS_R$$
$$= (8.86 - \tfrac{12.75^2}{20}) - 0.5886$$
$$= 0.143275$$

$$f_0 = \frac{0.5886}{0.143275/18} = 73.95$$

8) Since 73.95 > 4.414, reject H_0 and conclude the model specifies a useful relationship at $\alpha = 0.05$.

$\text{P} - \text{value} \cong 0.000001$

b) $se(\hat{\beta}_1) = \sqrt{\dfrac{\hat{\sigma}^2}{S_{xx}}} = \sqrt{\dfrac{.00796}{33991.6}} = 4.8391x10^{-4}$

$$se(\hat{\beta}_0) = \sqrt{\hat{\sigma}^2\left[\frac{1}{n} + \frac{\bar{x}}{S_{xx}}\right]} = \sqrt{.00796\left[\frac{1}{20} + \frac{73.9^2}{33991.6}\right]} = 0.04091$$

11-20. a) Refer to ANOVA table of Exercise 11-4.

1) The parameter of interest is the regressor variable coefficient, β_1.

2) $H_0: \beta_1 = 0$

3) $H_1: \beta_1 \neq 0$

4) $\alpha = 0.01$

5) The test statistic is

$$f_0 = \frac{MS_R}{MS_E} = \frac{SS_R / 1}{SS_E / (n-2)}$$

6) Reject H_0 if $f_0 > f_{\alpha,1,26}$ where $f_{0.01,1,26} = 7.721$

7) Using the results of Exercise 10-4

$$f_0 = \frac{MS_R}{MS_E} = 31.1032$$

8) Since 31.1032 > 7.721 reject H_0 and conclude the model is useful at $\alpha = 0.01$. $\text{P} - \text{value} = 0.000007$

b) $se(\hat{\beta}_1) = \sqrt{\dfrac{\hat{\sigma}^2}{S_{xx}}} = \sqrt{\dfrac{5.7257}{3608611\ .43}} = .001259$

$$se(\hat{\beta}_0) = \sqrt{\hat{\sigma}^2\left[\frac{1}{n} + \frac{\bar{x}}{S_{xx}}\right]} = \sqrt{5.7257\left[\frac{1}{28} + \frac{2110.13^2}{3608611\ .43}\right]} = 2.6962$$

c) 1) The parameter of interest is the regressor variable coefficient, β_1.

2) $H_0 : \beta_1 = -0.01$

3) $H_1 : \beta_1 \neq -0.01$

4) $\alpha = 0.01$

5) The test statistic is $t_0 = \dfrac{\hat{\beta}_1 + .01}{se(\hat{\beta}_1)}$

6) Reject H_0 if $t_0 < -t_{\alpha/2,n-2}$ where $-t_{0.005,26} = -2.78$ or $t_0 > t_{0.005,26} = 2.78$

7) Using the results from Exercise 10-4

$$t_0 = \frac{-0.0070251 + .01}{0.00125965} = 2.3618$$

8) Since 2.3618 < 2.78 do not reject H_0 and conclude the intercept is not zero at $\alpha = 0.01$.

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

$$t_0 = \frac{13.3201}{2.5717} = 5.179$$

8) Since 5.179 > 2.074 reject H_0 and conclude the intercept is not zero at $\alpha = 0.05$.

11-22. Refer to ANOVA for Exercise 10-6

a) 1) The parameter of interest is the regressor variable coefficient, β_1.

2) $H_0: \beta_1 = 0$

3) $H_1: \beta_1 \neq 0$

4) $\alpha = 0.01$

5) The test statistic is $f_0 = \dfrac{MS_R}{MS_E} = \dfrac{SS_R / 1}{SS_E / (n-2)}$

6) Reject H_0 if $f_0 > f_{\alpha,1,22}$ where $f_{0.01,1,10} = 10.049$

7) Using the results from Exercise 10-6

$$f_0 = \frac{280583.12/1}{37.746089/10} = 74334.4$$

8) Since 74334.4 > 10.049, reject H_0 and conclude the model is useful $\alpha = 0.01$. P-value < 0.000001

b) $se(\hat{\beta}_1) = 0.0337744$, $se(\hat{\beta}_0) = 1.66765$

c) 1) The parameter of interest is the regressor variable coefficient, β_1.

2) $H_0: \beta_1 = 10$

3) $H_1: \beta_1 \neq 10$

4) $\alpha = 0.01$

5) The test statistic is $t_0 = \dfrac{\hat{\beta}_1 - \beta_{1,0}}{se(\hat{\beta}_1)}$

6) Reject H_0 if $t_0 < -t_{\alpha/2,n-2}$ where $-t_{0.005,10} = -3.17$ or $t_0 > t_{0.005,10} = 3.17$

7) Using the results from Exercise 10-6

$$t_0 = \frac{9.21 - 10}{0.0338} = -23.37$$

8) Since −23.37 < −3.17 reject H_0 and conclude the slope is not 10 at $\alpha = 0.01$. P-value = 0.

d) H_0: $\beta_0 = 0$ H_1: $\beta_0 \neq 0$

$$t_0 = \frac{-6.3355 - 0}{1.66765} = -3.8$$

P-value < 0.005; Reject H_0 and conclude that the intercept should be included in the model.

11-23. Refer to ANOVA table of Exercise 11-7

a) $H_0: \beta_1 = 0$

$H_1: \beta_1 \neq 0$ $\alpha = 0.01$

$$f_0 = 4.53158$$

$$f_{0.01,1,18} = 8.285$$

$$f_0 \not> f_{\alpha,1,18}$$

Therefore, do not reject H_0. P-value = 0.04734. Insufficient evidence to conclude that the model is a useful relationship.

b) $se(\hat{\beta}_1) = 0.0166281$

$se(\hat{\beta}_0) = 2.61396$

c) $H_0: \beta_1 = -0.05$

$H_1: \beta_1 < -0.05$

$\alpha = 0.01$

$$t_0 = \frac{-0.0354 - (-0.05)}{0.0166281} = 0.87803$$

$t_{.01,18} = 2.552$

$t_0 \not< -t_{\alpha,18}$

Therefore, do not rejectH_0. P-value = 0.804251. Insufficient evidence to conclude that β_1 is ≥ -0.05.

d) $H_0: \beta_0 = 0$ $H_1: \beta_0 \neq 0$ $\alpha = 0.01$

$t_0 = 12.8291$

$t_{.005,18} = 2.878$

$t_0 > t_{\alpha/2,18}$

Therefore, reject H_0. P-value ≅ 0

11-24. Refer to ANOVA of Exercise 11-8

a) $H_0: \beta_1 = 0$

$H_1: \beta_1 \neq 0$

$\alpha = 0.05$

$f_0 = 44.0279$

$f_{.05,1,11} = 4.84$

$f_0 > f_{\alpha,1,11}$

Therefore, rejectH_0. P-value = 0.00004.

b) $se(\hat{\beta}_1) = 0.0104524$

$se(\hat{\beta}_0) = 9.84346$

c) $H_0: \beta_0 = 0$

$H_1: \beta_0 \neq 0$

$\alpha = 0.05$

$t_0 = -1.67718$

$t_{.025,11} = 2.201$

$|t_0| \not< -t_{\alpha/2,11}$

Therefore, do not rejectH_0. P-value = 0.12166.

11-25. Refer to ANOVA of Exercise 11-9

a) $H_0: \beta_1 = 0$

$H_1: \beta_1 \neq 0$

$\alpha = 0.05$

$f_0 = 53.50$

$f_{.05,1,18} = 4.414$

$f_0 > f_{\alpha,1,18}$

Therefore, reject H_0. P-value = 0.000009.

b) $se(\hat{\beta}_1) = 0.0256613$

$se(\hat{\beta}_0) = 2.13526$

c) $H_0: \beta_0 = 0$

$H_1: \beta_0 \neq 0$

$\alpha = 0.05$

$t_0 = -5.079$

$t_{.025,18} = 2.101$

$|t_0| > t_{\alpha/2,18}$

Therefore, reject H_0. P-value = 0.000078.

11-26. Refer to ANOVA of Exercise 11-11

a) $H_0: \beta_1 = 0$

$H_1: \beta_1 \neq 0$

$\alpha = 0.01$

$f_0 = 92.224$

$f_{.01,1,16} = 8.531$

$f_0 > f_{\alpha,1,16}$

Therefore, reject H_0.

b) P-value < 0.00001

c) $se(\hat{\beta}_1) = 2.14169$

$se(\hat{\beta}_0) = 1.93591$

d) $H_0: \beta_0 = 0$

$H_1: \beta_0 \neq 0$

$\alpha = 0.01$

$t_0 = 0.243$

$t_{.005,16} = 2.921$

$t_0 \not> t_{\alpha/2,16}$

Therefore, do not rejectH_0. Conclude, Yes, the intercept should be removed.

11-27. Refer to ANOVA of Exercise 11-12

a) $H_0: \beta_1 = 0$

$H_1: \beta_1 \neq 0$

$\alpha = 0.01$

$f_0 = 155.2$

$f_{.01,1,18} = 8.285$

$f_0 > f_{\alpha,1,18}$

Therefore, reject H_0. P-value < 0.00001.

b) $se(\hat{\beta}_1) = 45.3468$

$se(\hat{\beta}_0) = 2.96681$

c) $H_0: \beta_1 = -30$

$H_1: \beta_1 \neq -30$

$\alpha = 0.01$

$$t_0 = \frac{-36.9618 - (-30)}{2.96681} = -2.3466$$

$t_{.005,18} = 2.878$

$|t_0| \not> -t_{\alpha/2,18}$

Therefore, do not reject H_0. P-value = 0.0153(2) = 0.0306.

d) $H_0: \beta_0 = 0$

$H_1: \beta_0 \neq 0$

$\alpha = 0.01$

$t_0 = 57.8957$

$t_{.005,18} = 2.878$

$t_0 > t_{\alpha/2,18}$, therefore, reject H_0. P-value < 0.00001.

e) $H_0: \beta_0 = 2500$

$H_1: \beta_0 > 2500$

$\alpha = 0.01$

$$t_0 = \frac{2625.39 - 2500}{45.3468} = 2.7651$$

$t_{.01,18} = 2.552$

$t_0 > t_{\alpha,18}$, therefore reject H_0. P-value = 0.0064.

11-28. $t_0 = \frac{\hat{\beta}_1}{\sqrt{\hat{\sigma}^2 / S_{xx}}}$ After the transformation $\hat{\beta}_1^* = \frac{b}{a}\hat{\beta}_1$, $S_{xx}^* = a^2 S_{xx}$, $\bar{x}^* = a\bar{x}$, $\hat{\beta}_0^* = b\hat{\beta}_0$, and

$\hat{\sigma}^* = b\hat{\sigma}$. Therefore, $t_0^* = \frac{b\hat{\beta}_1 / a}{\sqrt{(b\hat{\sigma})^2 / a^2 S_{xx}}} = t_0$.

11-29. a) $\frac{\hat{\beta}}{\sqrt{\frac{\hat{\sigma}^2}{\sum x_i^2}}}$ has a t distribution with n-1 degree of freedom.

b) From Exercise 11-17, $\hat{\beta} = 21.031461, \hat{\sigma} = 3.611768,$ and $\sum x_i^2 = 14.7073$.

The t-statistic in part a. is 22.3314 and $H_0 : \beta_0 = 0$ is rejected at usual α values.

11-30. $d = \frac{|-0.01 - (-0.005)|}{2.4\sqrt{\frac{27}{3608611.96}}} = 0.76$, $S_{xx} = 3608611.96$.

Assume $\alpha = 0.05$, from Chart VI and interpolating between the curves for n = 20 and n = 30, $\beta \cong 0.05$.

<u>Sections 11-6 and 11-7</u>

11-31. $t_{\alpha/2,n-2} = t_{0.025,12} = 2.179$

a) 95% confidence interval on β_1.

$\hat{\beta}_1 \pm t_{\alpha/2,n-2} se(\hat{\beta}_1)$

$-2.3298 \pm t_{.025,12}(0.2696)$

$-2.3298 \pm 2.179(0.2696)$

$-2.9173. \le \beta_1 \le -1.7423.$

b) 95% confidence interval on β_0.

$\hat{\beta}_0 \pm t_{.025,12} se(\hat{\beta}_0)$

$48.0130 \pm 2.179(0.5959)$

$46.7145 \le \beta_0 \le 49.3115.$

c) 95% confidence interval on μ when $x_0 = 2.5$.

$\hat{\mu}_{Y|x_0} = 48.0130 - 2.3298(2.5) = 42.1885$

$\hat{\mu}_{Y|x_0} \pm t_{.025,12}\sqrt{\hat{\sigma}^2\left(\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}\right)}$

$42.1885 \pm (2.179)\sqrt{1.844\left(\frac{1}{14} + \frac{(2.5 - 3.0714)^2}{25.3486}\right)}$

$42.1885 \pm 2.179(0.3943)$

$41.3293 \le \hat{\mu}_{Y|x_0} \le 43.0477$

d) 95% on prediction interval when $x_0 = 2.5$.

$\hat{y}_0 \pm t_{.025,12}\sqrt{\hat{\sigma}^2\left(1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}\right)}$

$42.1885 \pm 2.179\sqrt{1.844\left(1 + \frac{1}{14} + \frac{(2.5 - 3.0714)^2}{25.348571}\right)}$

$42.1885 \pm 2.179(1.4056)$

$39.1257 \le y_0 \le 45.2513$

It is wider because it depends on both the error associated with the fitted model as well as that with the future observation.

11-32. $t_{\alpha/2,n-2} = t_{0.005,18} = 2.878$

a) $\hat{\beta}_1 \pm \left(t_{0.005,18}\right) se(\hat{\beta}_1)$.

$0.0041612 \pm (2.878)(0.000484)$

$0.0027682 \le \beta_1 \le 0.0055542$

b) $\hat{\beta}_0 \pm \left(t_{0.005,18}\right) se(\hat{\beta}_0)$.

$0.3299892 \pm (2.878)(0.04095)$

$0.212135 \le \beta_0 \le 0.447843$

c) 99% confidence interval on μ when $x_0 = 85^\circ F$.

$\hat{\mu}_{Y|x_0} = 0.683689$

$\hat{\mu}_{Y|x_0} \pm t_{.005,18}\sqrt{\hat{\sigma}^2\left(\frac{1}{n} + \frac{(x_0-\bar{x})^2}{S_{xx}}\right)}$

$0.683689 \pm (2.878)\sqrt{0.00796\left(\frac{1}{20} + \frac{(85-73.9)^2}{33991.6}\right)}$

0.683689 ± 0.0594607

$0.6242283 \le \hat{\mu}_{Y|x_0} \le 0.7431497$

d) 99% prediction interval when $x_0 = 90^\circ F$.

$\hat{y}_0 = 0.7044949$

$\hat{y}_0 \pm t_{.005,18}\sqrt{\hat{\sigma}^2\left(1 + \frac{1}{n} + \frac{(x_0-\bar{x})^2}{S_{xx}}\right)}$

$0.7044949 \pm 2.878\sqrt{0.00796\left(1 + \frac{1}{20} + \frac{(90-73.9)^2}{33991.6}\right)}$

0.7044949 ± 0.2640665

$0.4404284 \le y_0 \le 0.9685614$

Note for Problems 11-33 through 11-35: These computer printouts were obtained from Statgraphics. For Minitab users, the standard errors are obtained from the Regression subroutine.

11-33. 95 percent confidence intervals for coefficient estimates

	Estimate	Standard error	Lower Limit	Upper Limit
CONSTANT	21.7883	2.69623	16.2448	27.3318
Yards	-0.00703	0.00126	-0.00961	-0.00444

a) $-0.00961 \le \beta_1 \le -0.00444$.

b) $16.2448 \le \beta_0 \le 27.3318$.

c) $9.143 \pm (2.056)\sqrt{5.72585\left(\frac{1}{28} + \frac{(1800-2110.14)^2}{3608325.5}\right)}$

9.143 ± 1.2287

$7.9143 \le \hat{\mu}_{Y|x_0} \le 10.3717$

d) $9.143 \pm (2.056)\sqrt{5.72585\left(1 + \frac{1}{28} + \frac{(1800-2110.14)^2}{3608325.5}\right)}$

9.143 ± 5.0709

$4.0721 \le y_0 \le 14.2139$

11-34. 95 percent confidence intervals for coefficient estimates

	Estimate	Standard error	Lower Limit	Upper Limit
CONSTANT	13.3202	2.57172	7.98547	18.6549
Taxes	3.32437	0.39028	2.51479	4.13395

a) $2.51479 \le \beta_1 \le 4.13395$.

b) $7.98547 \le \beta_0 \le 18.6549$.

c) $38.253 \pm (2.074)\sqrt{8.76775(\frac{1}{24} + \frac{(7.5-6.40492)^2}{57.563139})}$

38.253 ± 1.5353

$36.7177 \le \hat{\mu}_{Y|x_0} \le 39.7883$

d) $38.253 \pm (2.074)\sqrt{8.76775(1 + \frac{1}{24} + \frac{(7.5-6.40492)^2}{57.563139})}$

38.253 ± 6.3302

$31.9228 \le y_0 \le 44.5832$

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) $500.124 \pm (2.228)\sqrt{3.774609(\frac{1}{12} + \frac{(55-46.5)^2}{3308.9994})}$

500.124 ± 1.4037586

$498.72024 \le \hat{\mu}_{Y|x_0} \le 501.52776$

d) $500.124 \pm (2.228)\sqrt{3.774609(1 + \frac{1}{12} + \frac{(55-46.5)^2}{3308.9994})}$

500.124 ± 4.5505644

$495.57344 \le y_0 \le 504.67456$

It is wider because the prediction interval includes error for both the fitted model and from that associated with the future observation.

11-36. a) $-0.07034 \le \beta_1 \le -0.00045$

b) $28.0417 \le \beta_0 \le 39.027$

c) $28.225 \pm (2.101)\sqrt{13.39232(\frac{1}{20} + \frac{(150-149.3)^2}{48436.256})}$

28.225 ± 1.7194236

$26.5406 \le \mu_{y|x_0} \le 29.9794$

d) $28.225 \pm (2.101)\sqrt{13.39232\left(1+\frac{1}{20}+\frac{(150-149.3)^2}{48436.256}\right)}$

28.225 ± 7.87863

$20.3814 \le y_0 \le 36.1386$

11-37. a) $0.03689 \le \beta_1 \le 0.10183$

b) $-47.0877 \le \beta_0 \le 14.0691$

c) $46.6041 \pm (3.106)\sqrt{7.324951\left(\frac{1}{13}+\frac{(910-939)^2}{67045.97}\right)}$

46.6041 ± 2.514401

$44.0897 \le \mu_{y|x_0} \le 49.1185)$

d) $46.6041 \pm (3.106)\sqrt{7.324951\left(1+\frac{1}{13}+\frac{(910-939)^2}{67045.97}\right)}$

46.6041 ± 8.779266

$37.8298 \le y_0 \le 55.3784$

11-38. a) $0.11756 \le \beta_1 \le 0.22541$

b) $-14.3002 \le \beta_0 \le -5.32598$

c) $4.76301 \pm (2.101)\sqrt{1.982231\left(\frac{1}{20}+\frac{(85-82.3)^2}{3010.2111}\right)}$

4.76301 ± 0.6772655

$4.0857 \le \mu_{y|x_0} \le 5.4403$

d) $4.76301 \pm (2.101)\sqrt{1.982231\left(1+\frac{1}{20}+\frac{(85-82.3)^2}{3010.2111}\right)}$

4.76301 ± 3.0345765

$1.7284 \le y_0 \le 7.7976$

11-39. a) $201.552 \le \beta_1 \le 266.590$

b) $-4.67015 \le \beta_0 \le -2.34696$

c) $128.814 \pm (2.365)\sqrt{398.2804\left(\frac{1}{9}+\frac{(30-24.5)^2}{1651.4214}\right)}$

128.814 ± 16.980124

$111.8339 \le \mu_{y|x_0} \le 145.7941$

11-40. a) $14.3107 \le \beta_1 \le 26.8239$

b) $-5.18501 \le \beta_0 \le 6.12594$

c) $21.038 \pm (2.921)\sqrt{13.8092\left(\frac{1}{18}+\frac{(1-0.806111)^2}{3.01062}\right)}$

21.038 ± 2.8314277

$18.2066 \le \mu_{y|x_0} \le 23.8694$

d) $21.038 \pm (2.921)\sqrt{13.8092\left(1+\frac{1}{18}+\frac{(1-0.806111)^2}{3.01062}\right)}$

21.038 ± 11.217861

$9.8201 \le y_0 \le 32.2559$

11-41. a) $-43.1964 \le \beta_1 \le -30.7272$

b) $2530.09 \le \beta_0 \le 2720.68$

c) $1886.154 \pm (2.101)\sqrt{9811.21\left(\frac{1}{20}+\frac{(20-13.3375)^2}{1114.6618}\right)}$

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

$1668.9013 \le y_0 \le 2103.4067$

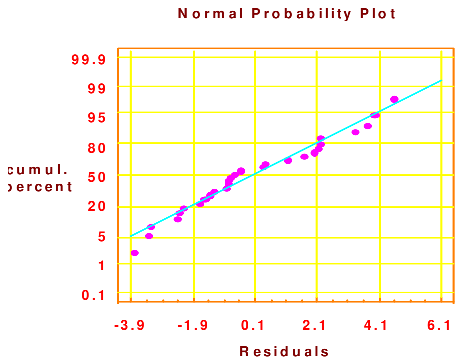
Section 11-7

11-42. Use the results of Exercise 11-4 to answer the following questions.

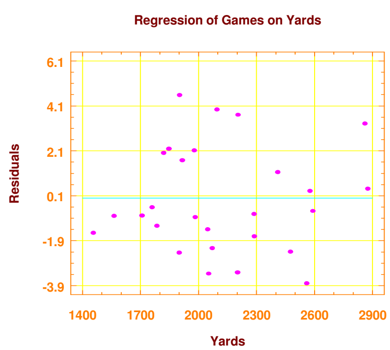
a) $R^2 = 0.544684$; The proportion of variability explained by the model.

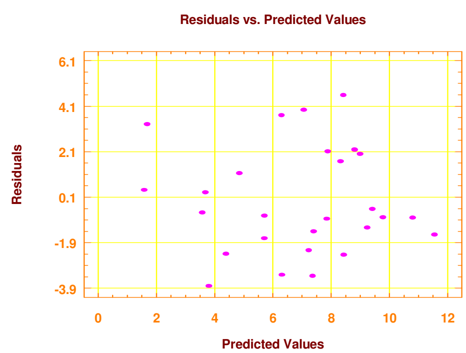
$$R^2_{Adj} = 1 - \frac{148.87197/26}{326.96429/27} = 1 - 0.473 = 0.527$$

b) Yes, normality seems to be satisfied since the data appear to fall along the straight line.



c) Since the residuals plots appear to be random, the plots do not include any serious model inadequacies.



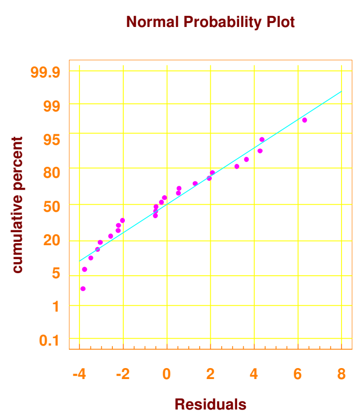


11-43. Use the Results of exercise 11-5 to answer the following questions.

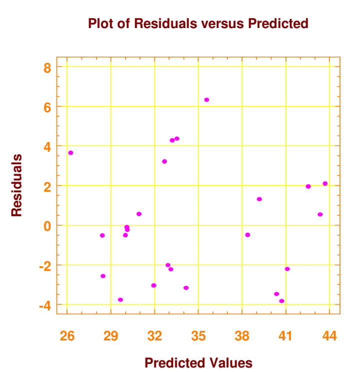
a)

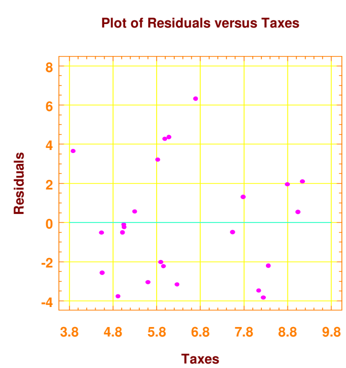
SalePrice	Taxes	Predicted	Residuals
25.9	4.9176	29.6681073	-3.76810726
29.5	5.0208	30.0111824	-0.51118237
27.9	4.5429	28.4224654	-0.52246536
25.9	4.5573	28.4703363	-2.57033630
29.9	5.0597	30.1405004	-0.24050041
29.9	3.8910	26.2553078	3.64469225
30.9	5.8980	32.9273208	-2.02732082
28.9	5.6039	31.9496232	-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) There are no serious departures from the assumption of constant variance. This is evident by the random pattern of the residuals.

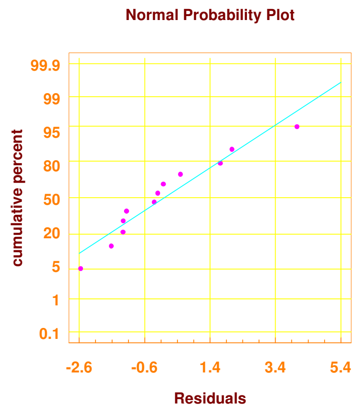




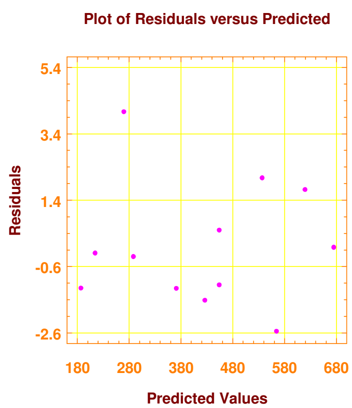
d) $R^2 \equiv 76.73\%$;

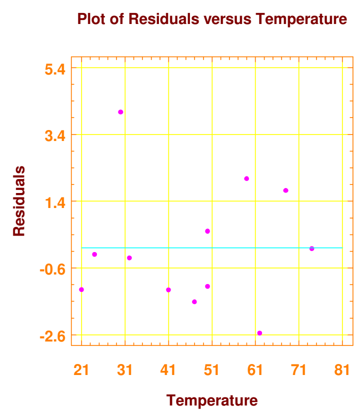
11-44. Use the results of Exercise 11-6 to answer the following questions

a) $R^2 = 99.986\%$; The proportion of variability explained by the model.
b) Yes, normality seems to be satisfied since the data appear to fall along the straight line.

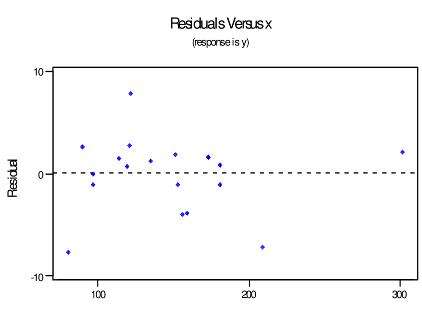


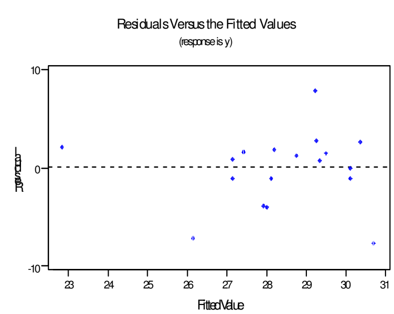
c) There might be lower variance at the middle settings of x. However, this data does not indicate a serious departure from the assumptions.



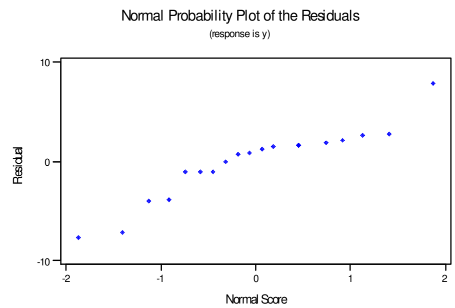


11-45. a) $R^2 = 20.1121\%$
b) These plots indicate presence of outliers, but no real problem with assumptions.

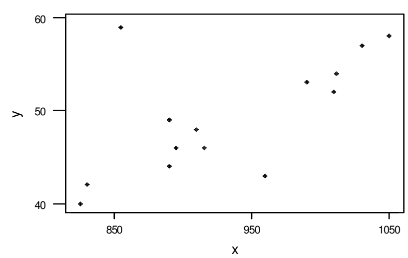




c) The normality assumption appears marginal.



11-46. a)



$$\hat{y} = 0.677559 + 0.0521753x$$

b) $H_0 : \beta_1 = 0$ $H_1 : \beta_1 \neq 0$ $\alpha = 0.05$

$f_0 = 7.9384$

$f_{.05,1,12} = 4.75$

$f_0 > f_{\alpha,1,12}$

Reject H_o.

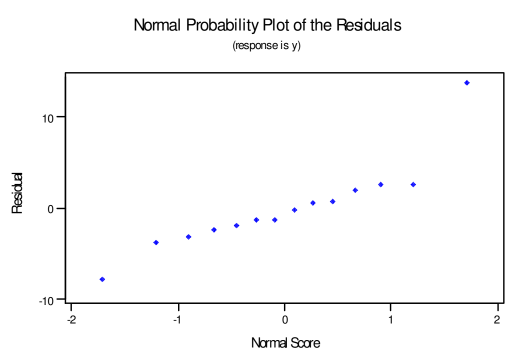
c) $\hat{\sigma}^2 = 25.23842$

d) $\hat{\sigma}_{orig}{}^2 = 7.324951$

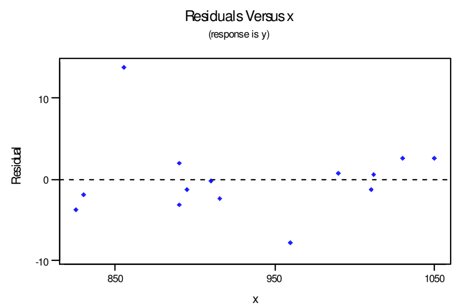
The new estimate is larger because the new point added additional variance not accounted for by the model.

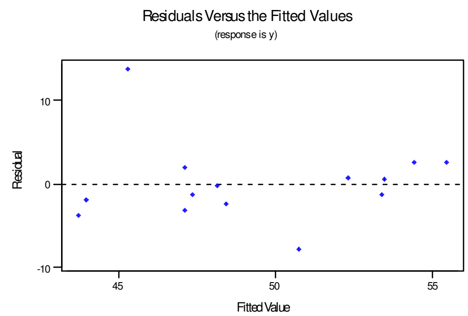
e) Yes, e_{14} is especially large compared to the other residuals.

f) The one added point is an outlier and the normality assumption is not as valid with the point included.



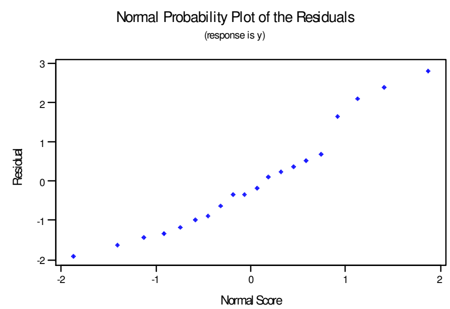
g) Constant variance assumption appears valid except for the added point.



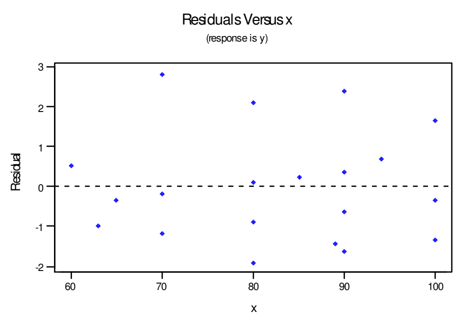


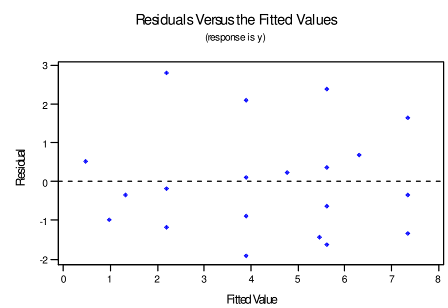
11-47. a) $R^2 = 71.27\%$

b) No major departure from normality assumptions.



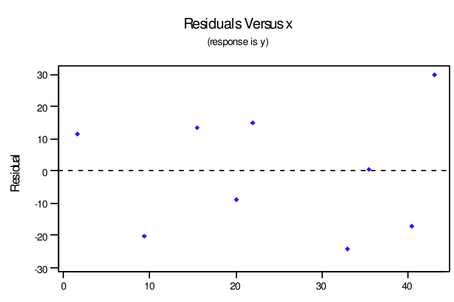
c) Assumption of constant variance appears reasonable.

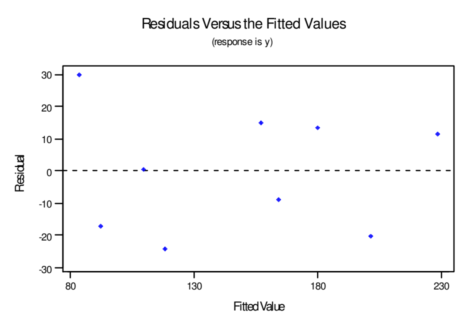




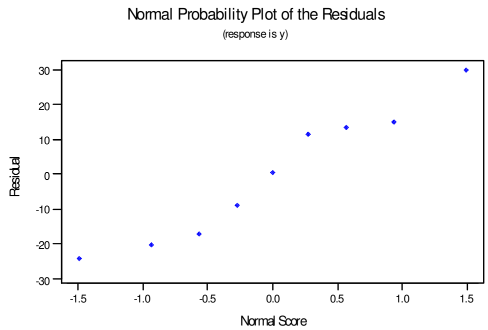
11-48. a) $R^2 = 0.879397$

b) No departures from constant variance are noted.



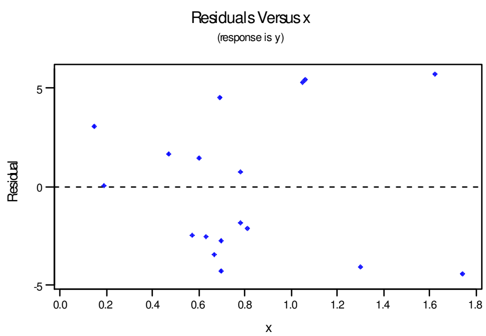


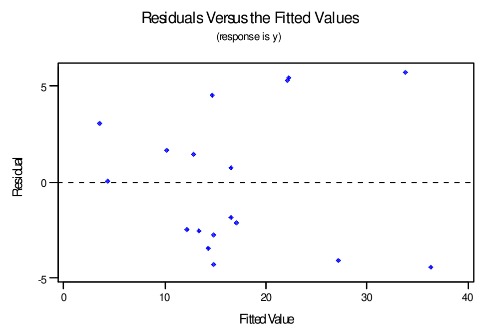
c) Normality assumption appears reasonable.



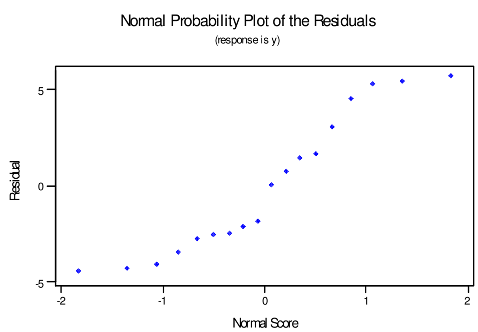
11-49. a) $R^2 = 85.22\%$

b) Assumptions appear reasonable, but there is a suggestion that variability increases slightly with $\hat{y}$.



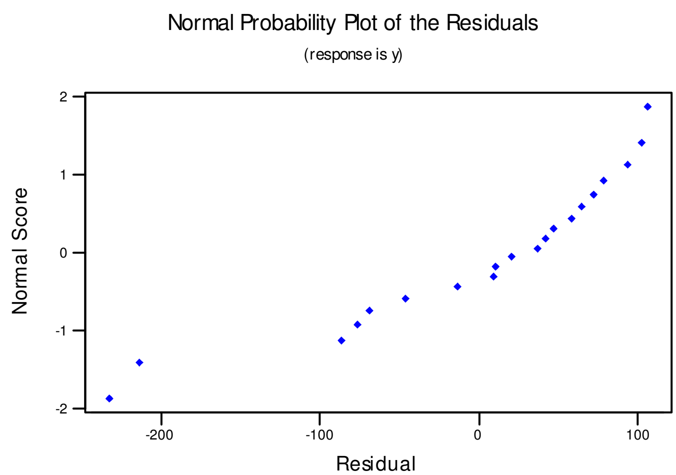


c) Normality assumption may be questionable. There is some " bending" away from a straight line in the tails of the normal probability plot.



11-50. a) $R^2 = 0.896081$ 89% of the variability is explained by the model.

b) Yes, the two points with residuals much larger in magnitude than the others.



c) $R^2_{new\ \ model} = 0.9573$

Larger, because the model is better able to account for the variability in the data with these two outlying data points removed.

d) $\hat{\sigma}^2_{old\ \ model} = 9811.21$

$\hat{\sigma}^2_{new\ \ model} = 4022.93$

Yes, reduced more than 50%, because the two removed points accounted for a large amount of the error.

11-51. Using $R^2 = 1 - \frac{SS_E}{S_{yy}}$, $F_0 = \dfrac{(n-2)(1-\frac{SS_E}{S_{yy}})}{\frac{SS_E}{S_{yy}}} = \dfrac{S_{yy} - SS_E}{\frac{SS_E}{n-2}} = \dfrac{S_{yy} - SS_E}{\hat{\sigma}^2}$

Also,

$$SS_E = \sum (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2$$
$$= \sum (y_i - \bar{y} - \hat{\beta}_1(x_i - \bar{x}))^2$$
$$= \sum (y_i - \bar{y}) + \hat{\beta}_1^2 \sum (x_i - \bar{x})^2 - 2\hat{\beta}_1 \sum (y_i - \bar{y})(x_i - \bar{x})$$
$$= \sum (y_i - \bar{y})^2 - \hat{\beta}_1^2 \sum (x_i - \bar{x})^2$$

$$S_{yy} - SS_E = \hat{\beta}_1^2 \sum (x_i - \bar{x})^2$$

Therefore, $F_0 = \dfrac{\hat{\beta}_1^2}{\hat{\sigma}^2 / S_{xx}} = t_0^2$

Because the square of a t random variable with n-2 degrees of freedom is an F random variable with 1 and n-2 degrees of freedom, the usually t-test that compares $|t_0|$ to $t_{\alpha/2,n-2}$ is equivalent to comparing $f_0 = t_0^2$ to

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

11-52. a) $f_0 = \frac{0.9(23)}{1-0.9} = 207$. Reject $H_0 : \beta_1 = 0$.

b) Because $f_{.05,1,23} = 4.28$, H_0 is rejected if $\frac{23R^2}{1-R^2} > 4.28$.

That is, H_0 is rejected if

$$23R^2 > 4.28(1-R^2)$$
$$27.28R^2 > 4.28$$
$$R^2 > 0.157$$

11-53. Yes, the larger residuals are easier to identify.

1.10269	-0.75866	-0.14376	0.66992
-2.49758	-2.25949	0.50867	0.46158
0.10242	0.61161	0.21046	-0.94548
0.87051	0.74766	-0.50425	0.97781
0.11467	0.38479	1.13530	-0.82398

11-54. For two random variables X_1 and X_2,

$V(X_1 + X_2) = V(X_1) + V(X_2) + 2Cov(X_1, X_2)$

Then,

$$\begin{aligned} V(Y_i - \hat{Y}_i) &= V(Y_i) + V(\hat{Y}_i) - 2Cov(Y_i, \hat{Y}_i) \\ &= \sigma^2 + V(\hat{\beta}_0 + \hat{\beta}_1 x_i) - 2\sigma^2\left[\tfrac{1}{n} + \tfrac{(x_i-\bar{x})^2}{S_{xx}}\right] \\ &= \sigma^2 + \sigma^2\left[\tfrac{1}{n} + \tfrac{(x_i-\bar{x})^2}{S_{xx}}\right] - 2\sigma^2\left[\tfrac{1}{n} + \tfrac{(x_i-\bar{x})^2}{S_{xx}}\right] \\ &= \sigma^2\left[1 - \left(\tfrac{1}{n} + \tfrac{(x_i-\bar{x})^2}{S_{xx}}\right)\right] \end{aligned}$$

a) Because e_i is divided by an estimate of its standard error (when σ^2 is estimated by $\hat{\sigma}^2$), r_i has approximate unit variance.

b) No, the term in brackets in the denominator is necessary.

c) If x_i is near $\bar{x}$ and n is reasonably large, r_i is approximately equal to the standardized residual.

d) If x_i is far from $\bar{x}$, the standard error of e_i is small. Consequently, extreme points are better fit by least squares regression than points near the middle range of x. Because the studentized residual at any point has variance of approximately one, the studentized residuals can be used to compare the fit of points to the regression line over the range of x.

Section 11-9

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) $\tanh(\text{arcta nh } 0.90334 - \frac{z_{.025}}{\sqrt{17}}) \leq \rho \leq \tanh(\text{arcta nh } 0.90334 + \frac{z_{.025}}{\sqrt{17}})$ where $z_{.025} = 1.96$.

$0.7677 \leq \rho \leq 0.9615$.

11-56. a) $\hat{y} = 69.1044 + 0.419415\ x$

b) $H_0 : \beta_1 = 0$

$H_1 : \beta_1 \neq 0$ $\alpha = 0.05$

$f_0 = 35.744$

$f_{.05,1,24} = 4.260$

$f_0 > f_{\alpha,1,24}$

Reject H_0 .

c) $r = 0.77349$

d) $H_0 : \rho = 0$

$H_1 : \rho \neq 0$ $\alpha = 0.05$

$t_0 = \frac{0.77349\ \sqrt{24}}{\sqrt{1-0.5983}} = 5.9787$

$t_{.025,24} = 2.064$

$t_0 > t_{\alpha/2,24}$

Reject H_0 .

e) $H_0 : \rho = 0.6$

$H_1 : \rho \neq 0.6$ $\alpha = 0.05$

$z_0 = (\text{arctanh } 0.77349 - \text{arctanh } 0.6)(23)^{1/2} = 1.6105$

$z_{.025} = 1.96$

$z_0 \ngtr z_{\alpha/2}$

Do not reject H_0 .

f) $\tanh(\text{arcta nh } 0.77349 - \frac{z_{.025}}{\sqrt{23}}) \leq \rho \leq \tanh(\text{arcta nh } 0.77349 + \frac{z_{.025}}{\sqrt{23}})$ where $z_{.025} = 1.96$.

$0.5513 \leq \rho \leq 0.8932$.

11-57. a) r = -0.738027

b) $H_0 : \rho = 0$

$H_1 : \rho \neq 0$ α = 0.05

$t_0 = \frac{-0.738027\sqrt{26}}{\sqrt{1-0.5447}} = -5.577$

$t_{.025,26} = 2.056$

$|t_0| > t_{\alpha/2,26}$

Reject H_0. P-value = (3.69E-6)(2) = 7.38E-6

c) $\tanh(\text{arctanh}\ -0.738 - \frac{z_{.025}}{\sqrt{25}}) \le \rho \le \tanh(\text{arctanh}\ -0.738 + \frac{z_{.025}}{\sqrt{25}})$

where $z_{.025} = 1.96$. $-0.871 \le \rho \le -0.504$.

d) $H_0 : \rho = -0.7$

$H_1 : \rho \neq -0.7$ α = 0.05

$z_0 = (\text{arctanh}\ -0.738 - \text{arctanh}\ -0.7)(25)^{1/2} = -0.394$

$z_{.025} = 1.96$

$|z_0| < z_{\alpha/2}$

Do not reject H_0. P-value = (0.3468)(2) = 0.6936

11-58 $R = \hat{\beta}_1 \left(\frac{S_{xx}}{S_{yy}} \right)^{1/2}$ and $1 - R^2 = \frac{SS_E}{S_{yy}}$.

Therefore, $T_0 = \frac{R\sqrt{n-2}}{\sqrt{1-R^2}} = \frac{\hat{\beta}_1 \left(\frac{S_{xx}}{S_{yy}} \right)^{1/2} \sqrt{n-2}}{\left(\frac{SS_E}{S_{yy}} \right)^{1/2}} = \frac{\hat{\beta}_1}{\sqrt{\frac{\hat{\sigma}^2}{S_{xx}}}}$ where $\hat{\sigma}^2 = \frac{SS_E}{n-2}$.

11-59 n = 50 r = 0.62

a) $H_0 : \rho = 0$

$H_1 : \rho \neq 0$ α = 0.01

$t_0 = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \frac{0.62\sqrt{48}}{\sqrt{1-(0.62)^2}} = 5.475$

$t_{.005,48} = 2.682$

$t_0 > t_{0.005,48}$

Reject H_0. *P*-value ≅ 0

b) $\tanh(\text{arctanh}\ 0.62 - \frac{z_{.005}}{\sqrt{47}}) \le \rho \le \tanh(\text{arctanh}\ 0.62 + \frac{z_{.005}}{\sqrt{47}})$

where $z_{.005} = 2.575$. $0.3358 \le \rho \le 0.8007$.

c) Yes.

11-60. n = 10000, r = 0.02

a) $H_0 : \rho = 0$

$H_1 : \rho \neq 0$ α = 0.05

$t_0 = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \frac{0.02\sqrt{10000}}{\sqrt{1-(0.02)^2}} = 2.0002$

$t_{.025,9998} = 1.96$

$t_0 > t_{\alpha/2,9998}$

Reject H_0. P-value = 2(0.02274) = 0.04548

b) Since the sample size is so large, the standard error is very small. Therefore, very small differences are found to be "statistically" significant. However, the practical significance is minimal since r = 0.02 is essentially zero.

11-61. a) r = 0.933203

b) $H_0 : \rho = 0$

$H_1 : \rho \neq 0$ $\alpha = 0.05$

$t_0 = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \frac{0.933203\ \sqrt{15}}{\sqrt{1-(0.8709)}} = 10.06$

$t_{.025,15} = 2.131$

$t_0 > t_{\alpha/2,15}$

Reject H_0.

c) $\hat{y} = 0.72538 + 0.498081x$

$H_0 : \beta_1 = 0$

$H_1 : \beta_1 \neq 0$ $\alpha = 0.05$

$f_0 = 101.16$

$f_{.05,1,15} = 4.543$

$f_0 >> f_{\alpha,1,15}$

Reject H_0. Conclude that the model is significant at $\alpha = 0.05$. This test and the one in part b are identical.

d) $H_0 : \beta_0 = 0$

$H_1 : \beta_0 \neq 0$ $\alpha = 0.05$

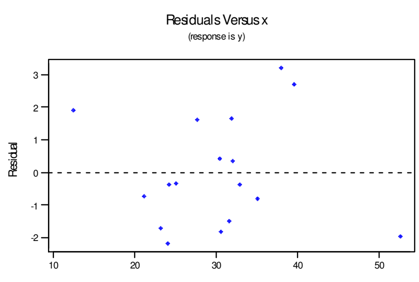
$t_0 = 0.468345$

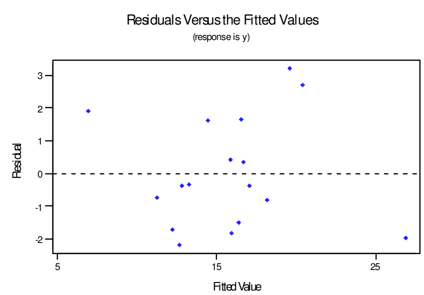
$t_{.025,15} = 2.131$

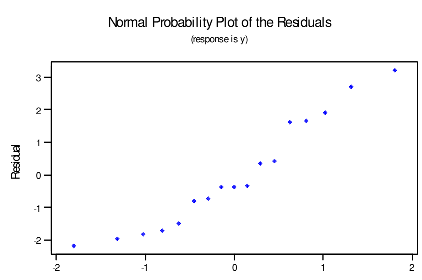
$t_0 \not> t_{\alpha/2,15}$

Do not reject H_0. We cannot conclude β_0 is different from zero.

e) No problems with model assumptions are noted.







11-62. n = 25 r = 0.83

a) $H_0 : \rho = 0$

$H_1 : \rho \neq 0$ $\alpha = 0.05$

$t_0 = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \frac{0.83\sqrt{23}}{\sqrt{1-(0.83)^2}} = 7.137$

$t_{.025,23} = 2.069$

$t_0 > t_{\alpha/2,23}$

Reject H_0. P-value = 0.

b) $\tanh(\text{arctanh } 0.83 - \frac{z_{.025}}{\sqrt{22}}) \le \rho \le \tanh(\text{arctanh } 0.83 + \frac{z_{.025}}{\sqrt{22}})$

where $z_{.025} = 1.96$. $0.6471 \le \rho \le 0.9226$.

c) $H_0 : \rho = 0.8$

$H_1 : \rho \neq 0.8$ $\alpha = 0.05$

$z_0 = (\text{arctanh } 0.83 - \text{arctanh } 0.8)(22)^{1/2} = 0.4199$

$z_{.025} = 1.96$

$z_0 \not> z_{\alpha/2}$

Do not reject H_0. P-value = (0.3373)(2) = 0.6746.

Supplemental Exercises

11-63. a) $\sum_{i=1}^{n}(y_i - \hat{y}_i) = \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} \hat{y}_i$ and $\sum y_i = n\hat{\beta}_0 + \hat{\beta}_1 \sum x_i$ from normal equation

Then,

$$(n\hat{\beta}_0 + \hat{\beta}_1 \sum_{i=1}^{n} x_i) - \sum \hat{y}_i$$

$$= n\hat{\beta}_0 + \hat{\beta}_1 \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} (\hat{\beta}_0 + \hat{\beta}_1 x_i)$$

$$= n\hat{\beta}_0 + \hat{\beta}_1 \sum_{i=1}^{n} x_i - n\hat{\beta}_0 - \hat{\beta}_1 \sum_{i=1}^{n} x_i = 0$$

b) $\sum_{i=1}^{n}(y_i - \hat{y}_i)x_i = \sum_{i=1}^{n} y_i x_i - \sum_{i=1}^{n} \hat{y}_i x_i$

and $\sum_{i=1}^{n} y_i x_i = \hat{\beta}_0 \sum_{i=1}^{n} x_i + \hat{\beta}_1 \sum_{i=1}^{n} x_i^2$ from normal equations. Then,

$$\hat{\beta}_0 \sum_{i=1}^{n} x_i + \hat{\beta}_1 \sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} (\hat{\beta}_0 + \hat{\beta}_1 x_i) x_i =$$

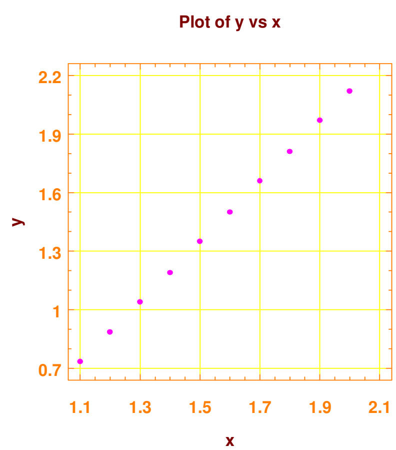
$$\hat{\beta}_0 \sum_{i=1}^{n} x_i + \hat{\beta}_1 \sum_{i=1}^{n} x_i^2 - \hat{\beta}_0 \sum_{i=1}^{n} x_i - \hat{\beta}_1 \sum_{i=1}^{n} x_i^2 = 0$$

c) $\frac{1}{n}\sum_{i=1}^{n}\hat{y}_i = \overline{y}$

$\sum \hat{y} = \sum (\hat{\beta}_0 + \hat{\beta}_1 x)$

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}\hat{y}_i &= \frac{1}{n}\sum(\hat{\beta}_0 + \hat{\beta}_1 x_i)\\ &= \frac{1}{n}(n\hat{\beta}_0 + \hat{\beta}_1 \sum x_i)\\ &= \frac{1}{n}(n(\overline{y} - \hat{\beta}_1\overline{x}) + \hat{\beta}_1 \sum x_i)\\ &= \frac{1}{n}(n\overline{y} - n\hat{\beta}_1\overline{x} + \hat{\beta}_1 \sum x_i)\\ &= \overline{y} - \hat{\beta}\overline{x} + \hat{\beta}_1\overline{x}\\ &= \overline{y}\end{aligned}$$

11-64. a)



Yes, a straight line relationship seems plausible.

b) Model fitting results for: y

Independent variable	coefficient	std. error	t-value	sig.level
CONSTANT	-0.966824	0.004845	-199.5413	0.0000
x	1.543758	0.003074	502.2588	0.0000

```
R-SQ. (ADJ.) = 1.0000  SE=      0.002792  MAE=      0.002063  DurbWat=  2.843
Previously:    0.0000           0.000000            0.000000              0.000
10 observations fitted, forecast(s) computed for 0 missing val. of dep. var.
```

$\hat{y} = -0.966824 + 1.54376x$

c) Analysis of Variance for the Full Regression

Source	Sum of Squares	DF	Mean Square	F-Ratio	P-value
Model	1.96613	1	1.96613	252264.	.0000
Error	0.0000623515	8	0.00000779394		
Total (Corr.)	1.96619	9			

```
R-squared = 0.999968                          Stnd. error of est. = 2.79176E-3
R-squared (Adj. for d.f.) = 0.999964          Durbin-Watson statistic = 2.84309
```

2) $H_0: \beta_1 = 0$

3) $H_1: \beta_1 \neq 0$

4) $\alpha = 0.05$

5) The test statistic is $f_0 = \dfrac{SS_R / k}{SS_E /(n-p)}$

6) Reject H_0 if $f_0 > f_{\alpha,1,8}$ where $f_{0.05,1,8} = 5.32$

7) Using the results from the ANOVA table

$$f_0 = \frac{1.96613/1}{0.0000623515/8} = 255263.9$$

8) Since 2552639 > 5.32 reject H_0 and conclude that the regression model is significant at $\alpha = 0.05$.
P-value < 0.000001

d) 95 percent confidence intervals for coefficient estimates

	Estimate	Standard error	Lower Limit	Upper Limit
CONSTANT	-0.96682	0.00485	-0.97800	-0.95565
x	1.54376	0.00307	1.53667	1.55085

$-0.97800 \leq \beta_0 \leq -0.95565$

e) 2) $H_0: \beta_0 = 0$

3) $H_1: \beta_0 \neq 0$

4) $\alpha = 0.05$

5) The test statistic is $t_0 = \dfrac{\hat{\beta}_0}{se(\hat{\beta}_0)}$

6) Reject H_0 if $t_0 < -t_{\alpha/2,n-2}$ where $-t_{0.025,8} = -2.306$ or $t_0 > t_{0.025,8} = 2.306$

7) Using the results from the table above

$$t_0 = \frac{-0.96682}{0.00485} = -199.34$$

8) Since $-199.34 < -2.306$ reject H_0 and conclude the intercept is significant at $\alpha = 0.05$.

11-65. a) $\hat{y} = 93.34 + 15.64x$

b) $H_0 : \beta_1 = 0$

$H_1 : \beta_1 \neq 0$ $\alpha = 0.05$

$f_0 = 12.872$

$f_{.05,1,14} = 4.60$

$f_0 > f_{0.05,1,14}$

Reject H_0. Conclude that $\beta_1 \neq 0$ at $\alpha = 0.05$.

c) $(7.961 \leq \beta_1 \leq 23.322)$

d) $(74.758 \leq \beta_0 \leq 111.923)$

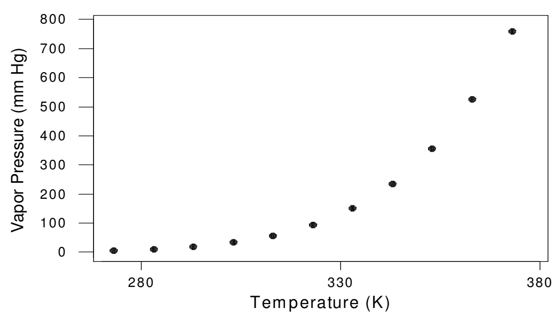
e) $\hat{y} = 93.34 + 15.64(2.5) = 132.44$

$$132.44 \pm 2.145\sqrt{136.27\left[\frac{1}{16} + \frac{(2.5-2.325)^2}{7.017}\right]}$$

$$132.44 \pm 6.47$$

$$125.97 \leq \hat{\mu}_{Y|x_0=2.5} \leq 138.91$$

11-66 a) There is curvature in the data.



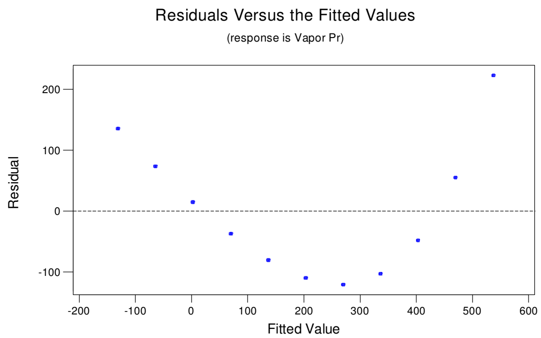
x

b) $y = -1956.3 + 6.686x$

c)

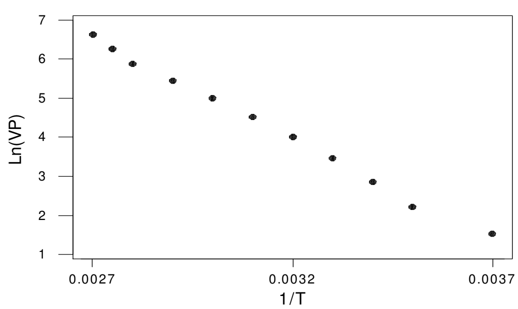
Source	DF	SS	MS	F	P
Regression	1	491662	491662	35.57	0.000
Residual Error	9	124403	13823		
Total	10	616065			

d)



There is a curve in the residuals.

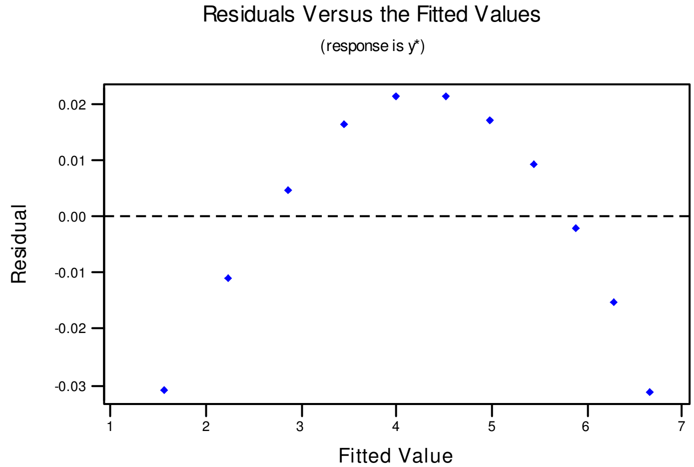
e) The data are linear after the transformation.



$\ln y = 20.6 - 5201\ (1/x)$

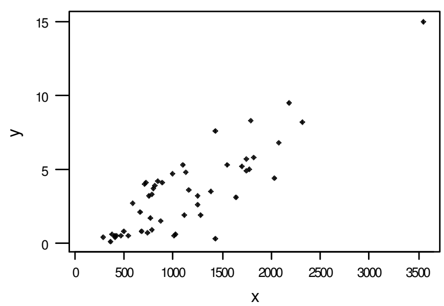
Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	28.511	28.511	66715.47	0.000
Residual Error	9	0.004	0.000		
Total	10	28.515			



There is still curvature in the data, but now the plot is convex instead of concave.

11-67. a)



b) $\hat{y} = -0.8819 + 0.00385x$

c) $H_0 : \beta_1 = 0$

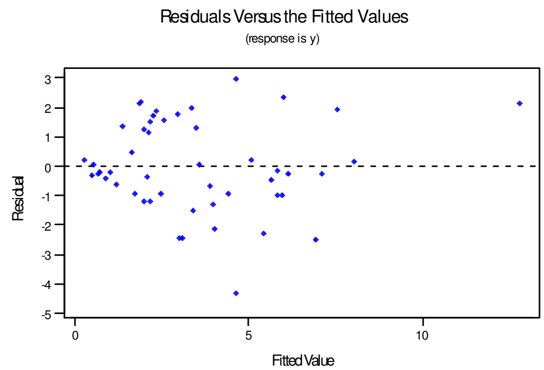
$H_1 : \beta_1 \neq 0$ $\alpha = 0.05$

$f_0 = 122.03$

$f_0 > f_{0.05,1,48}$

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.

11-68. $\hat{y}^* = 1.2232 + 0.5075x$ where $y^* = \frac{1}{y}$. No, model does not seem reasonable. The residual plots indicate a possible outlier.

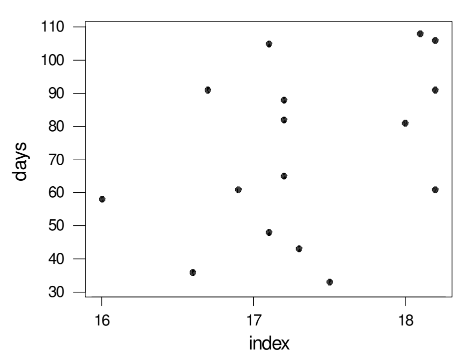
11-69. $\hat{y} = 0.7916x$

Even though y should be zero when x is zero, because the regressor variable does not normally assume values near zero, a model with an intercept fits this data better. Without an intercept, the MS_E is larger because there are fewer terms and the residuals plots are not satisfactory.

11-70. $\hat{y} = 4.5755 + 2.2047x$, r= 0.992, $R^2 = 98.40\%$

The model appears to be an excellent fit. Significance of regressor is strong and R^2 is large. Both regression coefficients are significant. No, the existence of a strong correlation does not imply a cause and effect relationship.

11-71 a)



b) The regression equation is

$$\hat{y} = -193 + 15.296x$$

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	1492.6	1492.6	2.64	0.127
Residual 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

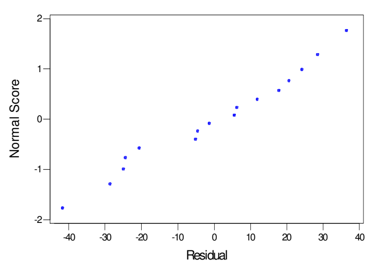
$$\hat{\beta}_1 \pm t_{\alpha/2,n-2} se(\hat{\beta}_1)$$

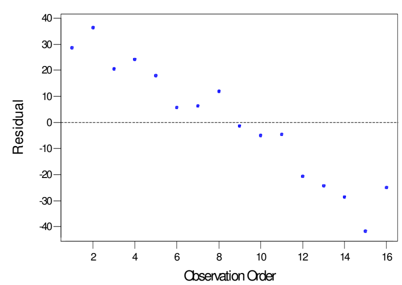
$$15.296 \pm t_{.025,12}(9.421)$$

$$15.296 \pm 2.145(9.421)$$

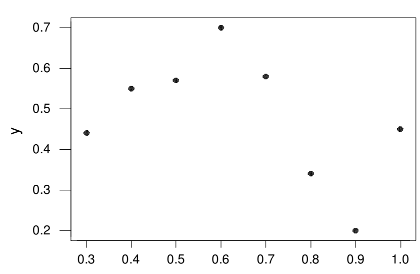
$$-4.912 \le \beta_1 \le 35.504$$

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.





11-72 a)



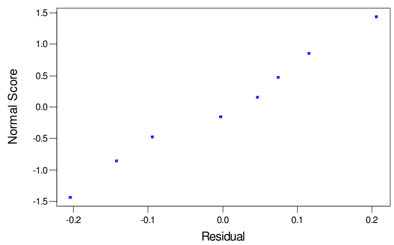
b) $\hat{y} = .6714 - 2964x$

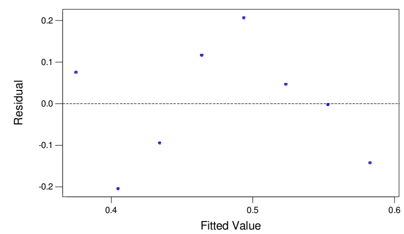
c) Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	0.03691	0.03691	1.64	0.248
Residual Error	6	0.13498	0.02250		
Total	7	0.17189			

$R^2 = 21.47\%$

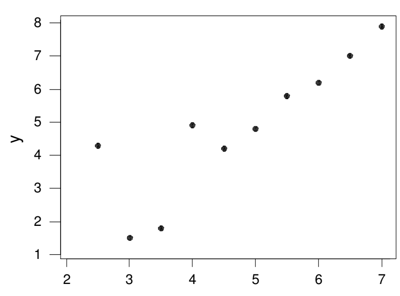
d) There appears to be curvature in the data. There is a dip in the middle of the normal probability plot and the plot of the residuals versus the fitted values shows curvature.





11-73 The correlation coefficient for the n pairs of data (x_i, z_i) will not be near unity. It will be near zero. The data for the pairs (x_i, z_i) where $z_i = y_i^2$ will not fall along the straight line $y_i = x_i$ which has a slope near unity and gives a correlation coefficient near unity. These data will fall on a line $y_i = \sqrt{x_i}$ that has a slope near zero and gives a much smaller correlation coefficient.

11-74 a)



b) $\hat{y} = -0.699 + 1.66x$

c)

Source	DF	SS	MS	F	P
Regression	1	28.044	28.044	22.75	0.001
Residual Error	8	9.860	1.233		
Total	9	37.904			

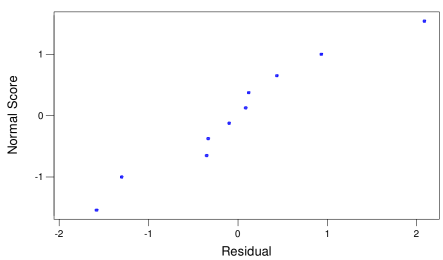
d) x=4.25 $\mu_{y|x_0} = 4.257$

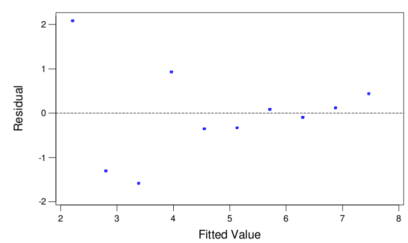
$$4.257 \pm 2.306\sqrt{1.2324\left(\frac{1}{10} + \frac{(4.25-4.75)^2}{20.625}\right)}$$

$$4.257 \pm 2.306(0.3717)$$

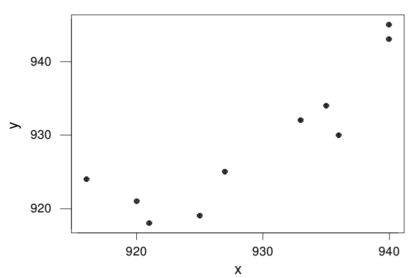
$$3.399 \le \mu_{y|x_o} \le 5.114$$

e) The normal probability plot of the residuals appears straight, but there are some large residuals in the lower fitted values. There may be some problems with the model.





11-75 a)



b) $\hat{y} = 33.3 + 0.9636x$

c)

Predictor	Coef	SE Coef	T	P
Constant	66.0	194.2	0.34	0.743
Therm	0.9299	0.2090	4.45	0.002

S = 5.435   R-Sq = 71.2%   R-Sq(adj) = 67.6%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	584.62	584.62	19.79	0.002
Residual Error	8	236.28	29.53		
Total	9	820.90			

Reject the hull hypothesis and conclude that the model is significant. 77.3% of the variability is explained by the model.

d) $H_0 : \beta_1 = 1$

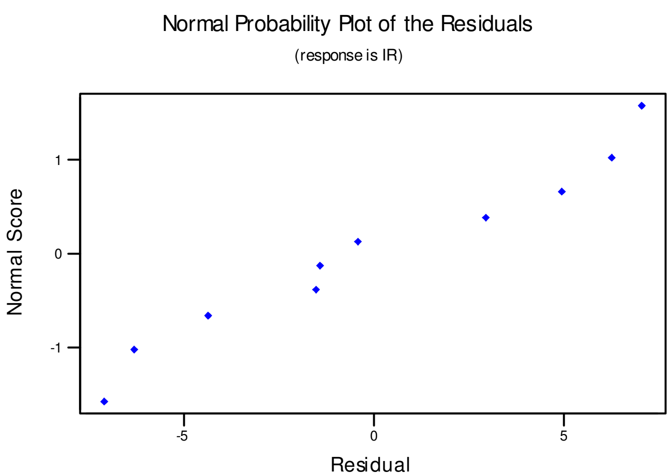
$H_1 : \beta_1 \neq 1$ α=.05

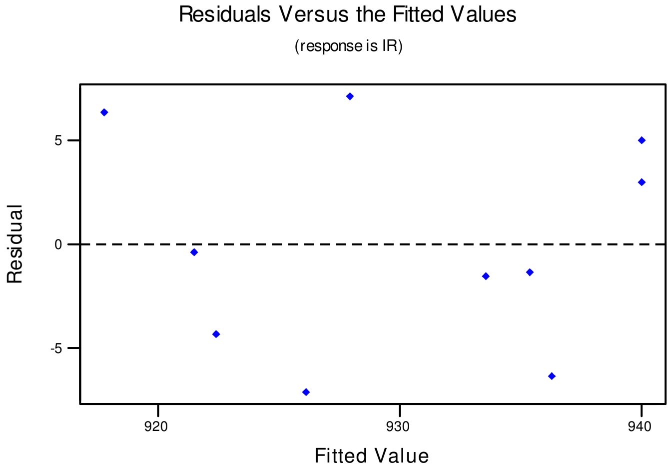
$$t_0 = \frac{\hat{\beta}_1 - 1}{se(\hat{\beta}_1)} = \frac{0.9299 - 1}{0.2090} = -0.3354$$

$$t_{a/2,n-2} = t_{.025,8} = 2.306$$

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.





Mind-Expanding Exercises

11-76. a) $\hat{\beta}_1 = \frac{S_{xY}}{S_{xx}}, \quad \hat{\beta}_0 = \overline{Y} - \hat{\beta}_1 \overline{x}$

$$Cov(\hat{\beta}_0, \hat{\beta}_1) = Cov(\overline{Y}, \hat{\beta}_1) - \overline{x}Cov(\hat{\beta}_1, \hat{\beta}_1)$$

$$Cov(\overline{Y}, \hat{\beta}_1) = \frac{Cov(\overline{Y}, S_{xY})}{S_{xx}} = \frac{Cov(\sum Y_i, \sum Y_i(x_i - \overline{x}))}{nS_{xx}} = \frac{\sum (x_i - \overline{x})\sigma^2}{nS_{xx}} = 0.$$ Therefore,

$$Cov(\hat{\beta}_1, \hat{\beta}_1) = V(\hat{\beta}_1) = \frac{\sigma^2}{S_{xx}}$$

$$Cov(\hat{\beta}_0, \hat{\beta}_1) = \frac{-\overline{x}\sigma^2}{S_{xx}}$$

b) The requested result is shown in part a.

11-77. a) $MS_E = \frac{\sum (Y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2}{n-2} = \frac{\sum e_i^2}{n-2}$

$$E(e_i) = E(Y_i) - E(\hat{\beta}_0) - E(\hat{\beta}_1)x_i = 0$$

$$V(e_i) = \sigma^2[1 - (\tfrac{1}{n} + \tfrac{(x_i - \overline{x})^2}{S_{xx}})]$$ Therefore,

$$E(MS_E) = \frac{\sum E(e_i^2)}{n-2} = \frac{\sum V(e_i)}{n-2}$$

$$= \frac{\sum \sigma^2[1 - (\frac{1}{n} + \frac{(x_i - \overline{x})^2}{S_{xx}})]}{n-2}$$

$$= \frac{\sigma^2[n-1-1]}{n-2} = \sigma^2$$

b) Using the fact that $SS_R = MS_R$, we obtain

$$E(MS_R) = E(\hat{\beta}_1^2 S_{xx}) = S_{xx}\left\{V(\hat{\beta}_1) + [E(\hat{\beta}_1)]^2\right\}$$
$$= S_{xx}\left(\frac{\sigma^2}{S_{xx}} + \beta_1^2\right) = \sigma^2 + \beta_1^2 S_{xx}$$

11-78. $\hat{\beta}_1 = \frac{S_{x_1Y}}{S_{x_1x_1}}$

$$E(\hat{\beta}_1) = \frac{E\left[\sum_{i=1}^{n} Y_i(x_{1i} - \bar{x}_1)\right]}{S_{x_1x_1}} = \frac{\sum_{i=1}^{n}(\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i})(x_{1i} - \bar{x}_1)}{S_{x_1x_1}}$$
$$= \frac{\beta_1 S_{x_1x_1} + \beta_2 \sum_{i=1}^{n} x_{2i}(x_{1i} - \bar{x}_1)}{S_{x_1x_1}} = \beta_1 + \frac{\beta_2 S_{x_1x_2}}{S_{x_1x_1}}$$

No, $\hat{\beta}_1$ is no longer unbiased.

11-79. $V(\hat{\beta}_1) = \frac{\sigma^2}{S_{xx}}$. To minimize $V(\hat{\beta}_1)$, S_{xx} should be maximized. Because $S_{xx} = \sum_{i=1}^{n}(x_i - \bar{x})^2$, S_{xx} is maximized by choosing approximately half of the observations at each end of the range of x.
From a practical perspective, this allocation assumes the linear model between Y and x holds throughout the range of x and observing Y at only two x values prohibits verifying the linearity assumption. It is often preferable to obtain some observations at intermediate values of x.

11-80. One might minimize a weighted some of squares $\sum_{i=1}^{n} w_i(y_i - \beta_0 - \beta_1 x_i)^2$ in which a Y_i with small variance (w_i large) receives greater weight in the sum of squares.

$$\frac{\partial}{\beta_0}\sum_{i=1}^{n} w_i(y_i - \beta_0 - \beta_1 x_i)^2 = -2\sum_{i=1}^{n} w_i(y_i - \beta_0 - \beta_1 x_i)$$
$$\frac{\partial}{\beta_1}\sum_{i=1}^{n} w_i(y_i - \beta_0 - \beta_1 x_i)^2 = -2\sum_{i=1}^{n} w_i(y_i - \beta_0 - \beta_1 x_i)x_i$$

Setting these derivatives to zero yields

$$\hat{\beta}_0 \sum w_i + \hat{\beta}_1 \sum w_i x_i = \sum w_i y_i$$

$$\hat{\beta}_0 \sum w_i x_i + \hat{\beta}_1 \sum w_i x_i^2 = \sum w_i x_i y_i$$

as requested.

and

$$\hat{\beta}_1 = \frac{\left(\sum w_i x_i y_i\right)\left(\sum w_i\right) - \sum w_i y_i}{\left(\sum w_i\right)\left(\sum w_i x_i^2\right) - \left(\sum w_i x_i\right)^2}$$

$$\hat{\beta}_0 = \frac{\sum w_i y_i}{\sum w_i} - \frac{\sum w_i x_i}{\sum w_i} \hat{\beta}_1 \quad .$$

11-81. $$\hat{y} = \overline{y} + r \frac{s_y}{s_x}(x - \overline{x})$$

$$= \overline{y} + \frac{S_{xy}\sqrt{\sum (y_i - \overline{y})^2}\,(x - \overline{x})}{\sqrt{S_{xx}S_{yy}}\sqrt{\sum (x_i - \overline{x})^2}}$$

$$= \overline{y} + \frac{S_{xy}}{S_{xx}}(x - \overline{x})$$

$$= \overline{y} + \hat{\beta}_1 x - \hat{\beta}_1 \overline{x} = \hat{\beta}_0 + \hat{\beta}_1 x$$

11-82. a) $$\frac{\partial}{\beta_1} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2 = -2\sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i) x_i$$

Upon setting the derivative to zero, we obtain

$$\beta_0 \sum x_i + \beta_1 \sum x_i^2 = \sum x_i y_i$$

Therefore,

$$\hat{\beta}_1 = \frac{\sum x_i y_i - \beta_0 \sum x_i}{\sum x_i^2} = \frac{\sum x_i (y_i - \beta_0)}{\sum x_i^2}$$

b) $$V(\hat{\beta}_1) = V\left(\frac{\sum x_i \ (Y_i - \beta_0)}{\sum x_i^2}\right) = \frac{\sum x_i^2 \sigma^2}{[\sum x_i^2]^2} = \frac{\sigma^2}{\sum x_i^2}$$

c) $$\hat{\beta}_1 \pm t_{\alpha/2, n-1}\sqrt{\frac{\hat{\sigma}^2}{\sum x_i^2}}$$

This confidence interval is shorter because $\sum x_i^2 \geq \sum (x_i - \overline{x})^2$. Also, the t value based on n-1 degrees of freedom is slightly smaller than the corresponding t value based on n-2 degrees of freedom.