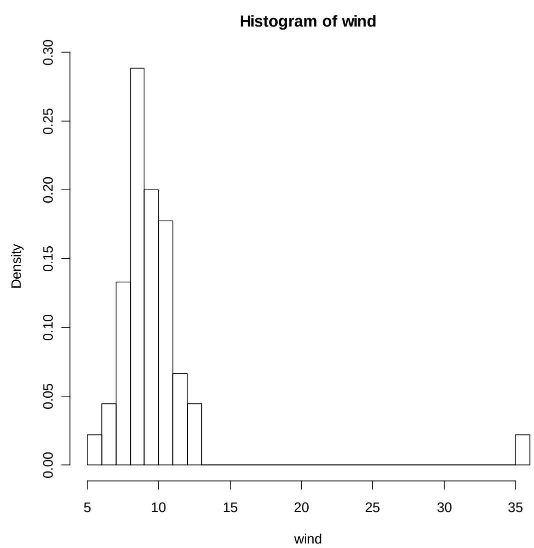
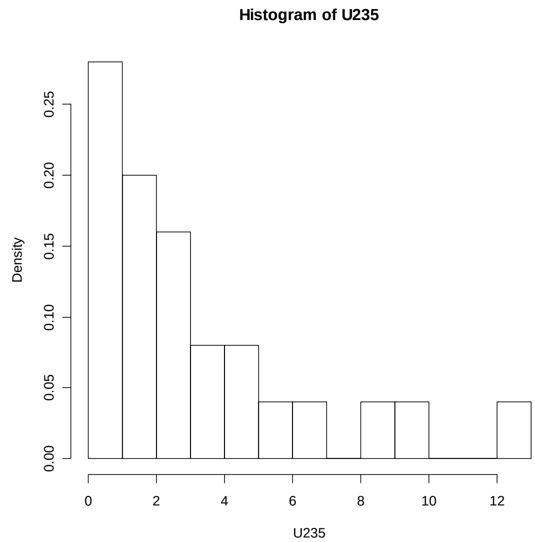
Chapter 1: What is Statistics?

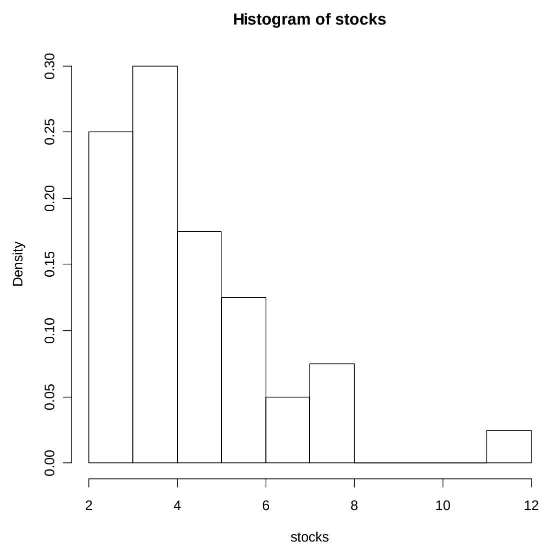
1.1 **a.** Population: all generation X age US citizens (specifically, assign a '1' to those who want to start their own business and a '0' to those who do not, so that the population is the set of 1's and 0's). Objective: to estimate the proportion of generation X age US citizens who want to start their own business.
b. Population: all healthy adults in the US. Objective: to estimate the true mean body temperature
c. Population: single family dwelling units in the city. Objective: to estimate the true mean water consumption
d. Population: all tires manufactured by the company for the specific year. Objective: to estimate the proportion of tires with unsafe tread.
e. Population: all adult residents of the particular state. Objective: to estimate the proportion who favor a unicameral legislature.
f. Population: times until recurrence for all people who have had a particular disease. Objective: to estimate the true average time until recurrence.
g. Population: lifetime measurements for all resistors of this type. Objective: to estimate the true mean lifetime (in hours).



1.2 **a.** This histogram is above.
b. Yes, it is quite windy there.
c. 11/45, or approx. 24.4%
d. it is not especially windy in the overall sample.



1.3 The histogram is above.

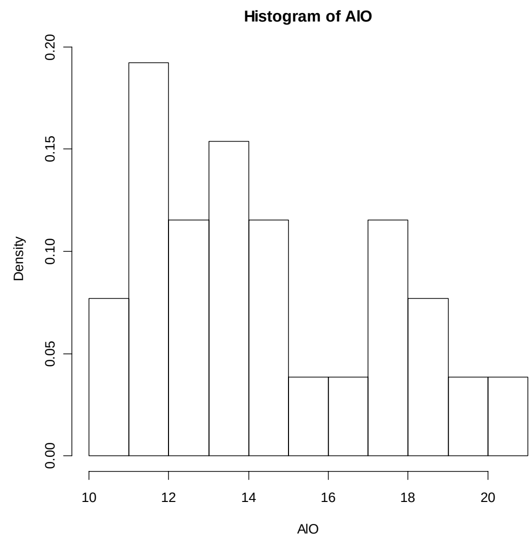


1.4 **a.** The histogram is above.
b. 18/40 = 45%
c. 29/40 = 72.5%

1.5 **a.** The categories with the largest grouping of students are 2.45 to 2.65 and 2.65 to 2.85. (both have 7 students).
b. 7/30
c. 7/30 + 3/30 + 3/30 + 3/30 = 16/30

1.6 **a.** The modal category is 2 (quarts of milk). About 36% (9 people) of the 25 are in this category.
b. .2 + .12 + .04 = .36
c. Note that 8% purchased 0 while 4% purchased 5. Thus, 1 – .08 – .04 = .88 purchased between 1 and 4 quarts.

1.7 **a.** There is a possibility of bimodality in the distribution.
b. There is a dip in heights at 68 inches.
c. If all of the students are roughly the same age, the bimodality could be a result of the men/women distributions.



1.8 **a.** The histogram is above.
b. The data appears to be bimodal. Llanederyn and Caldicot have lower sample values than the other two.

1.9 **a.** Note that 9.7 = 12 – 2.3 and 14.3 = 12 + 2.3. So, (9.7, 14.3) should contain approximately 68% of the values.
b. Note that 7.4 = 12 – 2(2.3) and 16.6 = 12 + 2(2.3). So, (7.4, 16.6) should contain approximately 95% of the values.
c. From parts (a) and (b) above, 95% - 68% = 27% lie in both (14.3. 16.6) and (7.4, 9.7). By symmetry, 13.5% should lie in (14.3, 16.6) so that 68% + 13.5% = 81.5% are in (9.7, 16.6)
d. Since 5.1 and 18.9 represent three standard deviations away from the mean, the proportion outside of these limits is approximately 0.

1.10 **a.** 14 – 17 = -3.
b. Since 68% lie within one standard deviation of the mean, 32% should lie outside. By symmetry, 16% should lie below one standard deviation from the mean.
c. If normally distributed, approximately 16% of people would spend less than –3 hours on the internet. Since this doesn't make sense, the population is not normal.

1.11 **a.** $\sum_{i=1}^{n} c = c + c + \ldots + c = nc.$

b. $\sum_{i=1}^{n} cy_i = c(y_1 + \ldots + y_n) = c\sum_{i=1}^{n} y_i$

c. $\sum_{i=1}^{n}(x_i + y_i) = x_1 + y_1 + x_2 + y_2 + \ldots + x_n + y_n = (x_1 + x_2 + \ldots + x_n) + (y_1 + y_2 + \ldots + y_n)$

Using the above, the numerator of s^2 is $\sum_{i=1}^{n}(y_i - \bar{y})^2 = \sum_{i=1}^{n}(y_i^{\,2} - 2y_i\bar{y} + \bar{y}^2) = \sum_{i=1}^{n} y_i^{\,2} - 2\bar{y}\sum_{i=1}^{n} y_i + n\bar{y}^2$ Since $n\bar{y} = \sum_{i=1}^{n} y_i$, we have $\sum_{i=1}^{n}(y_i - \bar{y})^2 = \sum_{i=1}^{n} y_i^{\,2} - n\bar{y}^2$. Let $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ to get the result.

1.12 Using the data, $\sum_{i=1}^{6} y_i = 14$ and $\sum_{i=1}^{6} y_i^{\,2} = 40$. So, $s^2 = (40 - 14^2/6)/5 = 1.47$. So, $s = 1.21$.

1.13 **a.** With $\sum_{i=1}^{45} y_i = 440.6$ and $\sum_{i=1}^{45} y_i^{\,2} = 5067.38$, we have that $\bar{y} = 9.79$ and $s = 4.14$.

b.

k	interval	frequency	Exp. frequency
1	5.65, 13.93	44	30.6
2	1.51, 18.07	44	42.75
3	-2.63, 22.21	44	45

1.14 **a.** With $\sum_{i=1}^{25} y_i = 80.63$ and $\sum_{i=1}^{25} y_i^{\,2} = 500.7459$, we have that $\bar{y} = 3.23$ and $s = 3.17$.

b.

k	interval	frequency	Exp. frequency
1	0.063, 6.397	21	17
2	-3.104, 9.564	23	23.75
3	-6.271, 12.731	25	25

1.15 **a.** With $\sum_{i=1}^{40} y_i = 175.48$ and $\sum_{i=1}^{40} y_i^{\,2} = 906.4118$, we have that $\bar{y} = 4.39$ and $s = 1.87$.

b.

k	interval	frequency	Exp. frequency
1	2.52, 6.26	35	27.2
2	0.65, 8.13	39	38
3	-1.22, 10	39	40

1.16 **a.** Without the extreme value, $\bar{y} = 4.19$ and $s = 1.44$.
b. These counts compare more favorably:

k	interval	frequency	Exp. frequency
1	2.75, 5.63	25	26.52
2	1.31, 7.07	36	37.05
3	-0.13, 8.51	39	39

1.17 For Ex. 1.2, range/4 = 7.35, while s = 4.14. For Ex. 1.3, range/4 = 3.04, while = s = 3.17. For Ex. 1.4, range/4 = 2.32, while s = 1.87.

1.18 The approximation is (800–200)/4 = 150.

1.19 One standard deviation below the mean is 34 – 53 = –19. The empirical rule suggests that 16% of all measurements should lie one standard deviation below the mean. Since chloroform measurements cannot be negative, this population cannot be normally distributed.

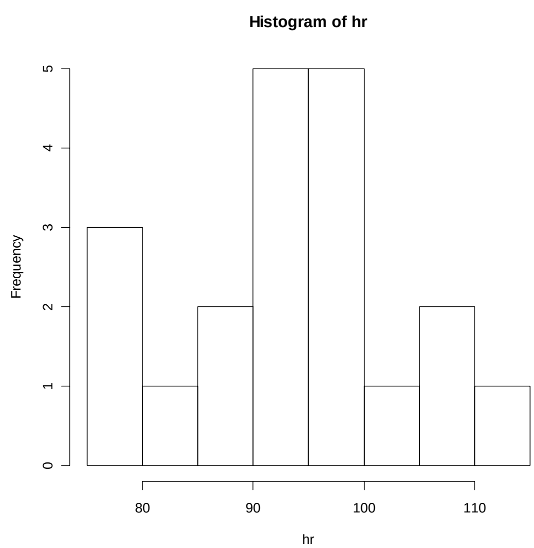
1.20 Since approximately 68% will fall between \$390 (\$420 – \$30) to \$450 (\$420 + \$30), the proportion above \$450 is approximately 16%.

1.21 (Similar to exercise 1.20) Having a gain of more than 20 pounds represents all measurements greater than one standard deviation below the mean. By the empirical rule, the proportion above this value is approximately 84%, so the manufacturer is probably correct.

1.22 (See exercise 1.11) $\sum_{i=1}^{n}(y_i - \overline{y}) = \sum_{i=1}^{n} y_i - n\overline{y} = \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} y_i = 0$.

1.23 **a.** (Similar to exercise 1.20) 95 sec = 1 standard deviation above 75 sec, so this percentage is 16% by the empirical rule.
b. (35 sec., 115 sec) represents an interval of 2 standard deviations about the mean, so approximately 95%
c. 2 minutes = 120 sec = 2.5 standard deviations above the mean. This is unlikely.

1.24 **a.** (112-78)/4 = 8.5



b. The histogram is above.

c. With $\sum_{i=1}^{20} y_i$ = 1874.0 and $\sum_{i=1}^{20} y_i^2$ = 117,328.0, we have that $\overline{y}$ = 93.7 and s = 9.55.

d.

k	interval	frequency	Exp. frequency
1	84.1, 103.2	13	13.6
2	74.6, 112.8	20	19
3	65.0, 122.4	20	20

1.25 **a.** (716-8)/4 = 177
b. The figure is omitted.
c. With $\sum_{i=1}^{88} y_i = 18{,}550$ and $\sum_{i=1}^{88} y_i^2 = 6{,}198{,}356$, we have that $\bar{y} = 210.8$ and $s = 162.17$.

d.

k	interval	frequency	Exp. frequency
1	48.6, 373	63	59.84
2	-113.5, 535.1	82	83.6
3	-275.7, 697.3	87	88

1.26 For Ex. 1.12, 3/1.21 = 2.48. For Ex. 1.24, 34/9.55 = 3.56. For Ex. 1.25, 708/162.17 = 4.37. The ratio increases as the sample size increases.

1.27 (64, 80) is one standard deviation about the mean, so 68% of 340 or approx. 231 scores. (56, 88) is two standard deviations about the mean, so 95% of 340 or 323 scores.

1.28 (Similar to 1.23) 13 mg/L is one standard deviation below the mean, so 16%.

1.29 If the empirical rule is assumed, approximately 95% of all bearing should lie in (2.98, 3.02) – this interval represents two standard deviations about the mean. So, approximately 5% will lie outside of this interval.

1.30 If $\mu = 0$ and $\sigma = 1.2$, we expect 34% to be between 0 and 0 + 1.2 = 1.2. Also, approximately 95%/2 = 47.5% will lie between 0 and 2.4. So, 47.5% – 34% = 13.5% should lie between 1.2 and 2.4.

1.31 Assuming normality, approximately 95% will lie between 40 and 80 (the standard deviation is 10). The percent below 40 is approximately 2.5% which is relatively unlikely.

1.32 For a sample of size n, let n' denote the number of measurements that fall outside the interval $\bar{y} \pm ks$, so that $(n - n')/n$ is the fraction that falls inside the interval. To show this fraction is greater than or equal to $1 - 1/k^2$, note that

$$(n-1)s^2 = \sum_{i \in A} (y_i - \bar{y})^2 + \sum_{i \in b} (y_i - \bar{y})^2 \text{, (both sums must be positive)}$$

where $A = \{i\colon |y_i - \bar{y}| \geq ks\}$ and $B = \{i\colon |y_i - \bar{y}| < ks\}$. We have that $\sum_{i \in A} (y_i - \bar{y})^2 \geq \sum_{i \in A} k^2 s^2 = n'k^2s^2$, since if i is in A, $|y_i - \bar{y}| \geq ks$ and there are n' elements in A. Thus, we have that $s^2 \geq k^2s^2n'/(n-1)$, or $1 \geq k^2n'/(n-1) \geq k^2n'/n$. Thus, $1/k^2 \geq n'/n$ or $(n - n')/n \geq 1 - 1/k^2$.

1.33 With k =2, at least 1 – 1/4 = 75% should lie within 2 standard deviations of the mean. The interval is (0.5, 10.5).

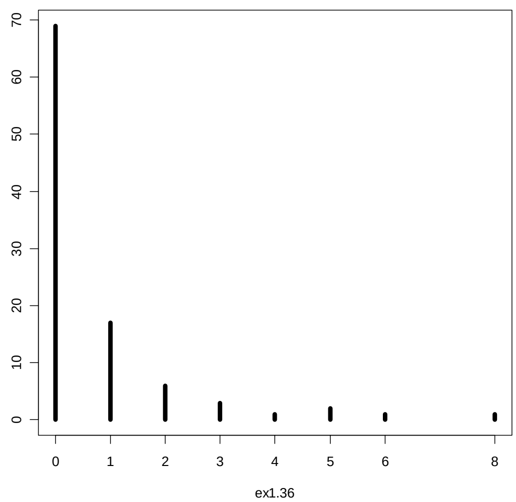
1.34 The point 13 is 13 – 5.5 = 7.5 units above the mean, or 7.5/2.5 = 3 standard deviations above the mean. By Tchebysheff's theorem, at least $1 - 1/3^2 = 8/9$ will lie within 3 standard deviations of the mean. Thus, at most 1/9 of the values will exceed 13.

1.35 **a.** (172 – 108)/4 =16

b. With $\sum_{i=1}^{15} y_i = 2041$ and $\sum_{i=1}^{15} y_i^2 = 281{,}807$ we have that $\bar{y} = 136.1$ and $s = 17.1$

c. $a = 136.1 - 2(17.1) = 101.9$, $b = 136.1 + 2(17.1) = 170.3$.

d. There are 14 observations contained in this interval, and 14/15 = 93.3%. 75% is a lower bound.



1.36 **a.** The histogram is above.

b. With $\sum_{i=1}^{100} y_i = 66$ and $\sum_{i=1}^{100} y_i^2 = 234$ we have that $\bar{y} = 0.66$ and $s = 1.39$.

c. Within two standard deviations: 95, within three standard deviations: 96. The calculations agree with Tchebysheff's theorem.

1.37 Since the lead readings must be non negative, 0 (the smallest possible value) is only 0.33 standard deviations from the mean. This indicates that the distribution is skewed.

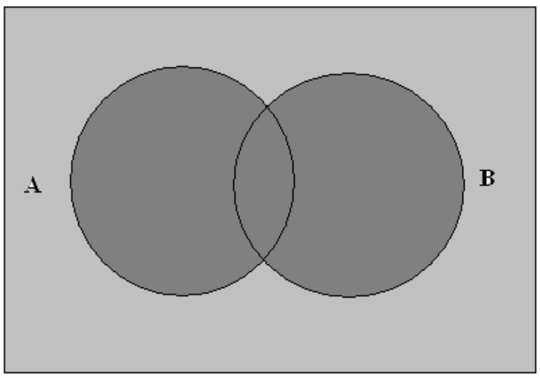
1.38 By Tchebysheff's theorem, at least 3/4 = 75% lie between (0, 140), at least 8/9 lie between (0, 193), and at least 15/16 lie between (0, 246). The lower bounds are all truncated a 0 since the measurement cannot be negative.

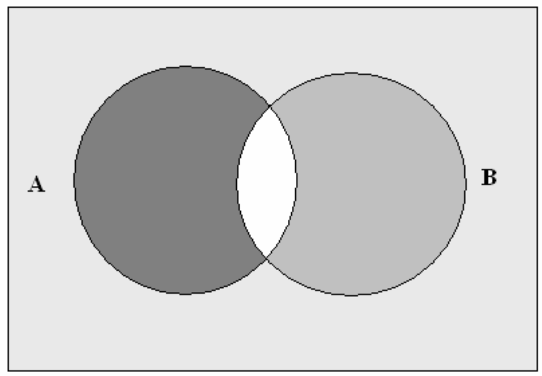
Chapter 2: Probability

2.1 A = {FF}, B = {MM}, C = {MF, FM, MM}. Then, $A \cap B = \emptyset$, $B \cap C$ = {MM}, $C \cap \overline{B}$ = {MF, FM}, $A \cup B$ ={FF,MM}, $A \cup C = S$, $B \cup C = C$.

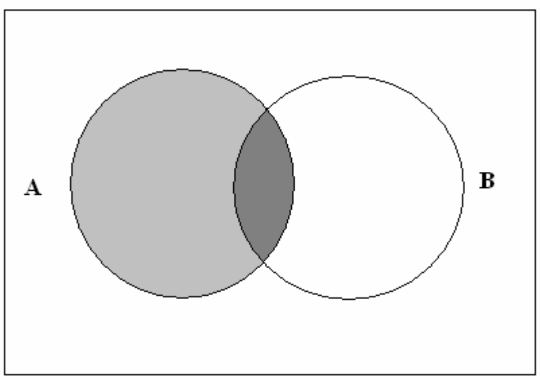
2.2 **a.** $A \cap B$ **b.** $A \cup B$ **c.** $\overline{A \cup B}$ **d.** $(A \cap \overline{B}) \cup (\overline{A} \cap B)$

2.3

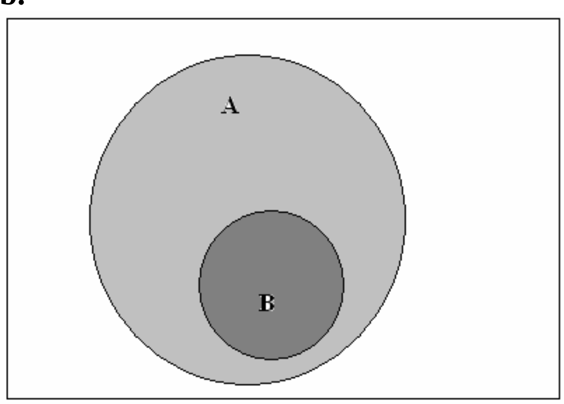




2.4 **a.**



b.



2.5 **a.** $(A\cap B)\cup(A\cap\overline{B}) = A\cap(B\cup\overline{B}) = A\cap S = A$.
b. $B\cup(A\cap\overline{B}) = (B\cap A)\cup(B\cap\overline{B}) = (B\cap A) = A$.
c. $(A\cap B)\cap(A\cap\overline{B}) = A\cap(B\cap\overline{B}) = \emptyset$. The result follows from part a.
d. $B\cap(A\cap\overline{B}) = A\cap(B\cap\overline{B}) = \emptyset$. The result follows from part b.

2.6 A = {(1,2), (2,2), (3,2), (4,2), (5,2), (6,2), (1,4), (2,4), (3,4), (4,4), (5,4), (6,4), (1,6), (2,6), (3,6), (4,6), (5,6), (6,6)}
$\overline{C}$ = {(2,2), (2,4), (2,6), (4,2), (4,4), (4,6), (6,2), (6,4), (6,6)}
$A\cap B$ = {(2,2), (4,2), (6,2), (2,4), (4,4), (6,4), (2,6), (4,6), (6,6)}
$A\cap\overline{B}$ = {(1,2), (3,2), (5,2), (1,4), (3,4), (5,4), (1,6), (3,6), (5,6)}
$\overline{A}\cup B$ = everything but {(1,2), (1,4), (1,6), (3,2), (3,4), (3,6), (5,2), (5,4), (5,6)}
$\overline{A}\cap C = \overline{A}$

2.7 A = {two males} = {M_1, M_2), (M_1,M_3), (M_2,M_3)}
B = {at least one female} = {(M_1,W_1), (M_2,W_1), (M_3,W_1), (M_1,W_2), (M_2,W_2), (M_3,W_2), {W_1,W_2)}
$\overline{B}$ = {no females} = A $\quad A\cup B = S \quad A\cap B = \emptyset \quad A\cap\overline{B} = A$

2.8 **a.** 36 + 6 = 42 **b.** 33 **c.** 18

2.9 S = {A+, B+, AB+, O+, A-, B-, AB-, O-}

2.10 **a.** S = {A, B, AB, O}
b. P({A}) = 0.41, P({B}) = 0.10, P({AB}) = 0.04, P({O}) = 0.45.
c. P({A} or {B}) = P({A}) + P({B}) = 0.51, since the events are mutually exclusive.

2.11 **a.** Since $P(S) = P(E_1) + \ldots + P(E_5) = 1$, $1 = .15 + .15 + .40 + 3P(E_5)$. So, $P(E_5) = .10$ and $P(E_4) = .20$.
b. Obviously, $P(E_3) + P(E_4) + P(E_5) = .6$. Thus, they are all equal to .2

2.12 **a.** Let L = {left tern}, R = {right turn}, C = {continues straight}.
b. P(vehicle turns) = P(L) + P(R) = 1/3 + 1/3 = 2/3.

2.13 **a.** Denote the events as very likely (VL), somewhat likely (SL), unlikely (U), other (O).
b. Not equally likely: P(VL) = .24, P(SL) = .24, P(U) = .40, P(O) = .12.
c. P(at least SL) = P(SL) + P(VL) = .48.

2.14 **a.** P(needs glasses) = .44 + .14 = .48
b. P(needs glasses but doesn't use them) = .14
c. P(uses glasses) = .44 + .02 = .46

2.15 **a.** Since the events are M.E., $P(S) = P(E_1) + \ldots + P(E_4) = 1$. So, $P(E_2) = 1 - .01 - .09 - .81 = .09$.
b. P(at least one hit) = $P(E_1) + P(E_2) + P(E_3) = .19$.

2.16 **a.** 1/3 **b.** 1/3 + 1/15 = 6/15 **c.** 1/3 + 1/16 = 19/48 **d.** 49/240

2.17 Let B = bushing defect, SH = shaft defect.
a. P(B) = .06 + .02 = .08
b. P(B or SH) = .06 + .08 + .02 = .16
c. P(exactly one defect) = .06 + .08 = .14
d. P(neither defect) = 1 – P(B or SH) = 1 – .16 = .84

2.18 **a.** S = {HH, TH, HT, TT}
b. if the coin is fair, all events have probability .25.
c. A = {HT, TH}, B = {HT, TH, HH}
d. $P(A) = .5$, $P(B) = .75$, $P(A \cap B) = P(A) = .5$, $P(A \cup B) = P(B) = .75$, $P(\overline{A} \cup B) = 1$.

2.19 **a.** $(V_1, V_1), (V_1, V_2), (V_1, V_3), (V_2, V_1), (V_2, V_2), (V_2, V_3), (V_3, V_1), (V_3, V_2), (V_3, V_3)$
b. if equally likely, all have probability of 1/9.
c. A = {same vendor gets both} = $\{(V_1, V_1), (V_2, V_2), (V_3, V_3)\}$
B = {at least one V2} = $\{(V_1, V_2), (V_2, V_1), (V_2, V_2), (V_2, V_3), (V_3, V_2)\}$
So, $P(A) = 1/3$, $P(B) = 5/9$, $P(A \cup B) = 7/9$, $P(A \cap B) = 1/9$.

2.20 **a.** $P(G) = P(D_1) = P(D_2) = 1/3$.
b. i. The probability of selecting the good prize is 1/3.
ii. She will get the other dud.
iii. She will get the good prize.
iv. Her probability of winning is now 2/3.
v. The best strategy is to switch.

2.21 $P(A) = P((A \cap B) \cup (A \cap \overline{B})) = P(A \cap B) + P(A \cap \overline{B})$ since these are M.E. by Ex. 2.5.

2.22 $P(A) = P(B \cup (A \cap \overline{B})) = P(B) + P(A \cap \overline{B})$ since these are M.E. by Ex. 2.5.

2.23 All elements in B are in A, so that when B occurs, A must also occur. However, it is possible for A to occur and B *not* to occur.

2.24 From the relation in Ex. 2.22, $P(A \cap \overline{B}) \geq 0$, so $P(B) \leq P(A)$.

2.25 Unless exactly 1/2 of all cars in the lot are Volkswagens, the claim is not true.

2.26 **a.** Let N_1, N_2 denote the empty cans and W_1, W_2 denote the cans filled with water. Thus, $S = \{N_1N_2, N_1W_2, N_2W_2, N_1W_1, N_2W_1, W_1W_2\}$
b. If this a merely a guess, the events are equally likely. So, $P(W_1W_2) = 1/6$.

2.27 **a.** S = {CC, CR, CL, RC, RR, RL, LC, LR, LL}
b. 5/9
c. 5/9

2.28 **a.** Denote the four candidates as A_1, A_2, A_3, and M. Since order is not important, the outcomes are {A_1A_2, A_1A_3, A_1M, A_2A_3, A_2M, A_3M}.
b. assuming equally likely outcomes, all have probability 1/6.
c. P(minority hired) = $P(A_1M) + P(A_2M) + P(A_3M) = .5$

2.29 **a.** The experiment consists of randomly selecting two jurors from a group of two women and four men.
b. Denoting the women as w1, w2 and the men as m1, m2, m3, m4, the sample space is

w1,m1	w2,m1	m1,m2	m2,m3	m3,m4
w1,m2	w2,m2	m1,m3	m2,m4	
w1,m3	w2,m3	m1,m4		
w1,m4	w2,m4			w1,w2

c. P(w1,w2) = 1/15

2.30 **a.** Let w1 denote the first wine, w2 the second, and w3 the third. Each sample point is an ordered triple indicating the ranking.
b. triples: (w1,w2,w3), (w1,w3,w2), (w2,w1,w3), (w2,w3,w1), (w3,w1,w2), (w3,w2,w1)
c. For each wine, there are 4 ordered triples where it is not last. So, the probability is 2/3.

2.31 **a.** There are four "good" systems and two "defactive" systems. If two out of the six systems are chosen randomly, there are 15 possible unique pairs. Denoting the systems as g1, g2, g3, g4, d1, d2, the sample space is given by S = {g1g2, g1g3, g1g4, g1d1, g1d2, g2g3, g2g4, g2d1, g2d2, g3g4, g3d1, g3d2, g4g1, g4d1, d1d2}. Thus:

P(at least one defective) = 9/15 P(both defective) = P(d1d2) = 1/15

b. If four are defective, P(at least one defective) = 14/15. P(both defective) = 6/15.

2.32 **a.** Let "1" represent a customer seeking style 1, and "2" represent a customer seeking style 2. The sample space consists of the following 16 four-tuples:

1111, 1112, 1121, 1211, 2111, 1122, 1212, 2112, 1221, 2121,
2211, 2221, 2212, 2122, 1222, 2222

b. If the styles are equally in demand, the ordering should be equally likely. So, the probability is 1/16.
c. $P(A) = P(1111) + P(2222) = 2/16$.

2.33 **a.** Define the events: G = family income is greater than \$43,318, N otherwise. The points are

E1: GGGG	E2: GGGN	E3: GGNG	E4: GNGG
E5: NGGG	E6: GGNN	E7: GNGN	E8: NGGN
E9: GNNG	E10: NGNG	E11: NNGG	E12: GNNN
E13: NGNN	E14: NNGN	E15: NNNG	E16: NNNN

b. A = {E1, E2, …, E11} B = {E6, E7, …, E11} C = {E2, E3, E4, E5}
c. If P(E) = P(N) = .5, each element in the sample space has probability 1/16. Thus,

$P(A) = 11/16$, $P(B) = 6/16$, and $P(C) = 4/16$.

2.34 **a.** Three patients enter the hospital and randomly choose stations 1, 2, or 3 for service. Then, the sample space S contains the following 27 three-tuples:
111, 112, 113, 121, 122, 123, 131, 132, 133, 211, 212, 213, 221, 222, 223, 231, 232, 233, 311, 312, 313, 321, 322, 323, 331, 332, 333
b. A = {123, 132, 213, 231, 312, 321}
c. If the stations are selected at random, each sample point is equally likely. $P(A) = 6/27$.

2.35 The total number of flights is 6(7) = 42.

2.36 There are 3! = 6 orderings.

2.37 **a.** There are 6! = 720 possible itineraries.
b. In the 720 orderings, exactly 360 have Denver before San Francisco and 360 have San Francisco before Denver. So, the probability is .5.

2.38 By the *mn* rule, 4(3)(4)(5) = 240.

2.39 **a.** By the *mn* rule, there are 6(6) = 36 possible roles.
b. Define the event $A = \{(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)\}$. Then, $P(A) = 6/36$.

2.40 **a.** By the *mn* rule, the dealer must stock 5(4)(2) = 40 autos.
b. To have each of these in every one of the eight colors, he must stock 8*40 = 320 autos.

2.41 If the first digit cannot be zero, there are 9 possible values. For the remaining six, there are 10 possible values. Thus, the total number is $9(10)(10)(10)(10)(10)(10) = 9*10^6$.

2.42 There are three *different* positions to fill using ten engineers. Then, there are $P_3^{10} = 10!/3!$ = 720 different ways to fill the positions.

2.43 $\binom{9}{3}\binom{6}{5}\binom{1}{1} = 504$ ways.

2.44 **a.** The number of ways the taxi needing repair can be sent to airport C is $\binom{8}{5}\binom{5}{5} = 56$.
So, the probability is 56/504 = 1/9.
b. $3\binom{6}{2}\binom{4}{4} = 45$, so the probability that every airport receives one of the taxis requiring repair is 45/504.

2.45 $\binom{17}{2\ 7\ 10} = 408{,}408$.

2.46 There are $\binom{10}{2}$ ways to chose two teams for the first game, $\binom{8}{2}$ for second, etc. So, there are $\binom{10}{2}\binom{8}{2}\binom{6}{2}\binom{4}{2}\binom{2}{2} = \frac{10!}{2^5}$ = 113,400 ways to assign the ten teams to five games.

2.47 There are $\binom{2n}{2}$ ways to chose two teams for the first game, $\binom{2n-2}{2}$ for second, etc. So, following Ex. 2.46, there are $\frac{2n!}{2^n}$ ways to assign $2n$ teams to n games.

2.48 Same answer: $\binom{8}{5} = \binom{8}{3}$ = 56.

2.49 **a.** $\binom{130}{2}$ = 8385.

b. There are 26*26 = 676 two-letter codes and 26(26)(26) = 17,576 three-letter codes. Thus, 18,252 total major codes.
c. 8385 + 130 = 8515 required.
d. Yes.

2.50 Two numbers, 4 and 6, are possible for each of the three digits. So, there are 2(2)(2) = 8 potential winning three-digit numbers.

2.51 There are $\binom{50}{3}$ = 19,600 ways to choose the 3 winners. Each of these is equally likely.

a. There are $\binom{4}{3}$ = 4 ways for the organizers to win all of the prizes. The probability is 4/19600.

b. There are $\binom{4}{2}\binom{46}{1}$ = 276 ways the organizers can win two prizes and one of the other 46 people to win the third prize. So, the probability is 276/19600.

c. $\binom{4}{1}\binom{46}{2}$ = 4140. The probability is 4140/19600.

d. $\binom{46}{3}$ = 15,180. The probability is 15180/19600.

2.52 The *mn* rule is used. The total number of experiments is 3(3)(2) = 18.

2.53 **a.** In choosing three of the five firms, order is important. So $P_3^5 = 60$ sample points.

b. If F_3 is awarded a contract, there are $P_2^4 = 12$ ways the other contracts can be assigned. Since there are 3 possible contracts, there are 3(12) = 36 total number of ways to award F_3 a contract. So, the probability is 36/60 = 0.6.

2.54 There are $\binom{8}{4} = 70$ ways to chose four students from eight. There are $\binom{3}{2}\binom{5}{2} = 30$ ways to chose exactly 2 (of the 3) undergraduates and 2 (of the 5) graduates. If each sample point is equally likely, the probability is 30/70 = 0.7.

2.55 **a.** $\binom{90}{10}$ **b.** $\binom{20}{4}\binom{70}{6}\Big/\binom{90}{10} = 0.111$

2.56 The student can solve all of the problems if the teacher selects 5 of the 6 problems that the student can do. The probability is $\binom{6}{5}\Big/\binom{10}{5} = 0.0238$.

2.57 There are $\binom{52}{2} = 1326$ ways to draw two cards from the deck. The probability is 4*12/1326 = 0.0362.

2.58 There are $\binom{52}{5} = 2{,}598{,}960$ ways to draw five cards from the deck.

a. There are $\binom{4}{3}\binom{4}{2} = 24$ ways to draw three Aces and two Kings. So, the probability is 24/2598960.

b. There are 13(12) = 156 types of "full house" hands. From part a. above there are 24 different ways each type of full house hand can be made. So, the probability is 156*24/2598960 = 0.00144.

2.59 There are $\binom{52}{5} = 2{,}598{,}960$ ways to draw five cards from the deck.

a. $\binom{4}{1}\binom{4}{1}\binom{4}{1}\binom{4}{1}\binom{4}{1} = 4^5 = 1024$. So, the probability is 1024/2598960 = 0.000394.

b. There are 9 different types of "straight" hands. So, the probability is $9(4^5)/2598960 =$ 0.00355. Note that this also includes "straight flush" and "royal straight flush" hands.

2.60 **a.** $\dfrac{365(364)(363)\cdots(365-n+1)}{365^n}$ **b.** With $n = 23$, $1-\dfrac{365(364)\cdots(343)}{365^{23}} = 0.507$.

2.61 **a.** $\frac{364(364)(364)\cdots(364)}{365^n} = \frac{364^n}{365^n}$. **b.** With n = 253, $1-\left(\frac{364}{365}\right)^{253}$ = 0.5005.

2.62 The number of ways to divide the 9 motors into 3 groups of size 3 is $\binom{9!}{3!\ 3!\ 3!}$ = 1680. If both motors from a particular supplier are assigned to the first line, there are only 7 motors to be assigned: one to line 1 and three to lines 2 and 3. This can be done $\binom{7!}{1!\ 3!\ 3!}$ = 140 ways. Thus, 140/1680 = 0.0833.

2.63 There are $\binom{8}{5}$ = 56 sample points in the experiment, and only one of which results in choosing five women. So, the probability is 1/56.

2.64 $6!\left(\frac{1}{6}\right)^6$ = 5/324.

2.65 $5!\left(\frac{2}{6}\right)^6\left(\frac{1}{6}\right)^4$ = 5/162.

2.66 **a.** After assigning an ethnic group member to each type of job, there are 16 laborers remaining for the other jobs. Let n_a be the number of ways that one ethnic group can be assigned to each type of job. Then:

$n_a = \binom{4}{1\ 1\ 1\ 1}\binom{16}{5\ 3\ 4\ 4}$. The probability is n_a/N = 0.1238.

b. It doesn't matter how the ethnic group members are assigned to jobs type 1, 2, and 3. Let n_a be the number of ways that no ethnic member gets assigned to a type 4 job. Then:

$n_a = \binom{4}{0}\binom{16}{5}$. The probability is $\binom{4}{0}\binom{16}{5}\Big/\binom{20}{5}$ = 0.2817.

2.67 As shown in Example 2.13, N = 10^7.

a. Let A be the event that all orders go to different vendors. Then, A contains n_a = 10(9)(8)…(4) = 604,800 sample points. Thus, $P(A)$ = 604,800/10^7 = 0.0605.

b. The 2 orders assigned to Distributor I can be chosen from the 7 in $\binom{7}{2}$ = 21 ways.

The 3 orders assigned to Distributor II can be chosen from the remaining 5 in $\binom{5}{3}$ = 10 ways. The final 2 orders can be assigned to the remaining 8 distributors in 8^2 ways. Thus, there are 21(10)(8^2) = 13,440 possibilities so the probability is 13440/10^7 = 0.001344.

c. Let A be the event that Distributors I, II, and III get exactly 2, 3, and 1 order(s) respectively. Then, there is one remaining unassigned order. Thus, A contains $\binom{7}{2}\binom{5}{3}\binom{2}{1}7 = 2940$ sample points and $P(A) = 2940/10^7 = 0.00029$.

2.68 **a.** $\binom{n}{n} = \frac{n!}{n!(n-n)!} = 1$. There is only one way to choose all of the items.

b. $\binom{n}{0} = \frac{n!}{0!(n-0)!} = 1$. There is only one way to chose none of the items.

c. $\binom{n}{r} = \frac{n!}{r!(n-r)!} = \frac{n!}{(n-r)!(n-(n-r))!} = \binom{n}{n-r}$. There are the same number of ways to choose r out of n objects as there are to choose $n - r$ out of n objects.

d. $2^n = (1+1)^n = \sum_{i=1}^{n}\binom{n}{i}1^{n-i}1^i = \sum_{i=1}^{n}\binom{n}{i}$.

2.69 $\binom{n}{k}+\binom{n}{k-1} = \frac{n!}{k!(n-k)!}+\frac{n!}{(k-1)!(n-k+1)!} = \frac{n!(n-k+1)}{k!(n-k+1)!}+\frac{n!k}{k!(n-k+1)!} = \frac{(n+1)!}{k!(n+1-k)!}$

2.70 From Theorem 2.3, let $y_1 = y_2 = \ldots = y_k = 1$.

2.71 **a.** $P(A|B) = .1/.3 = 1/3$.
b. $P(B|A) = .1/.5 = 1/5$.
c. $P(A|\, A \cup B\,) = .5/(.5+.3-.1) = 5/7$
d. $P(A|A\cap B) = 1$, since A has occurred.
e. $P(A\cap B|\, A \cup B\,) = .1(.5+.3-.1) = 1/7$.

2.72 Note that $P(A) = 0.6$ and $P(A|M) = .24/.4 = 0.6$. So, A and M are independent. Similarly, $P(\overline{A}\,|\,F\,) = .24/.6 = 0.4 = P(\overline{A}\,)$, so $\overline{A}$ and F are independent.

2.73 **a.** P(at least one R) = P(Red) 3/4. **b.** P(at least one r) = 3/4.
c. P(one r | Red) = .5/.75 = 2/3.

2.74 **a.** $P(A) = 0.61$, $P(D) = .30$. $P(A\cap D) = .20$. Dependent.
b. $P(B) = 0.30$, $P(D) = .30$. $P(B\cap D) = 0.09$. Independent.
c. $P(C) = 0.09$, $P(D) = .30$. $P(C\cap D) = 0.01$. Dependent.

2.75 **a.** Given the first two cards drawn are spades, there are 11 spades left in the deck. Thus, the probability is $\dfrac{\binom{11}{3}}{\binom{50}{3}} = 0.0084$. Note: this is also equal to $P(S_3S_4S_5|S_1S_2)$.

b. Given the first three cards drawn are spades, there are 10 spades left in the deck. Thus, the probability is $\dfrac{\binom{10}{2}}{\binom{49}{2}} = 0.0383$. Note: this is also equal to $P(S_4S_5|S_1S_2S_3)$.

c. Given the first four cards drawn are spades, there are 9 spades left in the deck. Thus, the probability is $\dfrac{\binom{9}{1}}{\binom{48}{1}} = 0.1875$. Note: this is also equal to $P(S_5|S_1S_2S_3S_4)$

2.76 Define the events: U: job is unsatisfactory A: plumber A does the job
a. $P(U|A) = P(A\cap U)/P(A) = P(A|U)P(U)/P(A) = .5*.1/.4 = 0.125$
b. From part a. above, $1 - P(U|A) = 0.875$.

2.77 **a.** 0.40 **b.** 0.37 **c.** 0.10 **d.** $0.40 + 0.37 - 0.10 = 0.67$
e. $1 - 0.4 = 0.6$ **f.** $1 - 0.67 = 0.33$ **g.** $1 - 0.10 = 0.90$
h. $.1/.37 = 0.27$ **i.** $1/.4 = 0.25$

2.78 **1.** Assume $P(A|B) = P(A)$. Then:
$P(A\cap B) = P(A|B)P(B) = P(A)P(B)$. $P(B|A) = P(B\cap A)/P(A) = P(A)P(B)/P(A) = P(B)$.
2. Assume $P(B|A) = P(B)$. Then:
$P(A\cap B) = P(B|A)P(A) = P(B)P(A)$. $P(A|B) = P(A\cap B)/P(B) = P(A)P(B)/P(B) = P(A)$.
3. Assume $P(A\cap B) = P(B)P(A)$. The results follow from above.

2.79 If A and B are M.E., $P(A\cap B) = 0$. But, $P(A)P(B) > 0$. So they are not independent.

2.80 If $A \subset B$, $P(A\cap B) = P(A) \neq P(A)P(B)$, unless $B = S$ (in which case $P(B) = 1$).

2.81 Given $P(A) < P(A|B) = P(A\cap B)/P(B) = P(B|A)P(A)/P(B)$, solve for $P(B|A)$ in the inequality.

2.82 $P(B|A) = P(B\cap A)/P(A) = P(A)/P(A) = 1$
$P(A|B) = P(A\cap B)/P(B) = P(A)/P(B)$.

2.83 $P(A|A\cup B) = P(A)/P(A\cup B) = \dfrac{P(A)}{P(A)+P(B)}$, since A and B are M.E. events.

2.84 Note that if $P(A_2\cap A_3) = 0$, then $P(A_1\cap A_2\cap A_3)$ also equals 0. The result follows from Theorem 2.6.

2.85 $P(A|\overline{B}) = P(A\cap\overline{B})/P(\overline{B}) = \dfrac{P(\overline{B}|A)P(A)}{P(\overline{B})} = \dfrac{[1-P(B|A)]P(A)}{P(\overline{B})} = \dfrac{[1-P(B)]P(A)}{P(\overline{B})} =$

$\dfrac{P(\overline{B})P(A)}{P(\overline{B})} = P(A)$. So, A and $\overline{B}$ are independent.

$P(\overline{B}|\overline{A}) = P(\overline{B}\cap\overline{A})/P(\overline{A}) = \dfrac{P(\overline{A}|\overline{B})P(\overline{B})}{P(\overline{A})} = \dfrac{[1-P(A|\overline{B})]P(\overline{B})}{P(\overline{A})}$. From the above,

A and $\overline{B}$ are independent. So $P(\overline{B}|\overline{A}) = \dfrac{[1-P(A)]P(\overline{B})}{P(\overline{A})} = \dfrac{P(\overline{A})P(\overline{B})}{P(\overline{A})} = P(\overline{B})$. So,

$\overline{A}$ and $\overline{B}$ are independent

2.86 **a.** No. It follows from $P(A\cup B) = P(A) + P(B) - P(A\cap B) \le 1$.
b. $P(A\cap B) \ge 0.5$
c. No.
d. $P(A\cap B) \le 0.70$.

2.87 **a.** $P(A) + P(B) - 1$.
b. the smaller of $P(A)$ and $P(B)$.

2.88 **a.** Yes.
b. 0, since they could be disjoint.
c. No, since $P(A\cap B)$ cannot be larger than either of $P(A)$ or $P(B)$.
d. $0.3 = P(A)$.

2.89 **a.** 0, since they could be disjoint.
b. the smaller of $P(A)$ and $P(B)$.

2.90 **a.** $(1/50)(1/50) = 0.0004$.
b. P(at least one injury) = 1 – P(no injuries in 50 jumps) = $1 = (49/50)^{50} = 0.636$. Your friend is not correct.

2.91 If A and B are M.E., $P(A\cup B) = P(A) + P(B)$. This value is greater than 1 if $P(A) = 0.4$ and $P(B) = 0.7$. So they cannot be M.E. It is possible if $P(A) = 0.4$ and $P(B) = 0.3$.

2.92 **a.** The three tests are independent. So, the probability in question is $(.05)^3 = 0.000125$.
b. P(at least one mistake) = 1 – P(no mistakes) = $1 - (.95)^3 = 0.143$.

2.93 Let H denote a hit and let M denote a miss. Then, she wins the game in three trials with the events HHH, HHM, and MHH. If she begins with her right hand, the probability she wins the game, assuming independence, is (.7)(.4)(.7) + (.7)(.4)(.3) + (.3)(.4)(.7) = 0.364.

2.94 Define the events *A*: device *A* detects smoke *B*: device *B* detects smoke
a. $P(A \cup B)$ = .95 + .90 - .88 = 0.97.
b. *P*(smoke is undetected) = 1 - $P(A \cup B)$ = 1 – 0.97 = 0.03.

2.95 Part a is found using the Addition Rule. Parts b and c use DeMorgan's Laws.
a. 0.2 + 0.3 – 0.4 = 0.1
b. 1 – 0.1 = 0.9
c. 1 – 0.4 = 0.6
d. $P(\overline{A} \mid B) = \dfrac{P(\overline{A} \cap B)}{P(B)} = \dfrac{P(B) - P(A \cap B)}{P(B)} = \dfrac{.3 - .1}{.3}$ = 2/3.

2.96 Using the results of Ex. 2.95:
a. 0.5 + 0.2 – (0.5)(0.2) = 0.6.
b. 1 – 0.6 = 0.4.
c. 1 – 0.1 = 0.9.

2.97 **a.** *P*(current flows) = 1 – *P*(all three relays are open) = 1 – $(.1)^3$ = 0.999.
b. Let *A* be the event that current flows and *B* be the event that relay 1 closed properly. Then, $P(B|A) = P(B \cap A)/P(A) = P(B)/P(A)$ = .9/.999 = 0.9009. Note that $B \subset A$.

2.98 Series system: *P*(both relays are closed) = (.9)(.9) = 0.81
Parallel system: *P*(at least one relay is closed) = .9 + .9 – .81 = 0.99.

2.99 Given that $P(\overline{A \cup B})$ = *a*, *P*(*B*) = *b*, and that *A* and *B* are independent. Thus $P(A \cup B)$ = 1 – *a* and $P(B \cap A) = bP(A)$. Thus, $P(A \cup B) = P(A) + b - bP(A) = 1 - a$. Solve for *P*(*A*).

2.100 $P(A \cup B \mid C) = \dfrac{P((A \cup B) \cap C)}{P(C)} = \dfrac{P((A \cap C) \cup (B \cap C))}{P(C)} =$

$$\frac{P(A \cap C) + P(B \cap C) - P(A \cap B \cap C)}{P(C)} = P(A|C) + P(B|C) + P(A \cap B|C).$$

2.101 Let A be the event the item gets past the first inspector and B the event it gets past the second inspector. Then, *P*(*A*) = 0.1 and *P*(*B*|*A*) = 0.5. Then $P(A \cap B)$ = .1(.5) = 0.05.

2.102 Define the events: I: disease I us contracted II: disease II is contracted. Then, *P*(I) = 0.1, *P*(II) = 0.15, and $P(\text{I} \cap \text{II})$ = 0.03.
a. $P(\text{I} \cup \text{II})$ = .1 + .15 – .03 = 0.22
b. $P(\text{I} \cap \text{II} \mid \text{I} \cup \text{II})$ = .03/.22 = 3/22.

2.103 Assume that the two state lotteries are independent.
a. P(666 in CT|666 in PA) = P(666 in CT) = 0.001
b. P(666 in CT∩666 in PA) = P(666 in CT)P(666 in PA) = .001(1/8) = 0.000125.

2.104 By DeMorgan's Law, $P(A\cap B) = 1 - P(\overline{A\cap B}) = 1 - P(\overline{A}\cup\overline{B})$. Since $P(\overline{A}\cup\overline{B}) \le P(\overline{A}) + P(\overline{B})$, $P(A\cap B) \ge 1 - P(\overline{A}) - P(\overline{B})$.

2.105 P(landing safely on both jumps) ≥ – 0.05 – 0.05 = 0.90.

2.106 Note that it must be also true that $P(\overline{A}) = P(\overline{B})$. Using the result in Ex. 2.104,

$P(A\cap B) \ge 1 - 2P(\overline{A}) \ge 0.98$, so $P(A) \ge 0.99$.

2.107 (Answers vary) Consider flipping a coin twice. Define the events:
A: observe at least one tail B: observe two heads or two tails C: observe two heads

2.108 Let U and V be two events. Then, by Ex. 2.104, $P(U\cap V) \ge 1 - P(\overline{U}) - P(\overline{V})$. Let U = A∩B and V = C. Then, $P(A\cap B\cap C) \ge 1 - P(\overline{A\cap B}) - P(\overline{C})$. Apply Ex. 2.104 to $P(\overline{A\cap B})$ to obtain the result.

2.109 This is similar to Ex. 2.106. Apply Ex. 2.108: $0.95 \le 1 - P(\overline{A}) - P(\overline{B}) - P(\overline{C}) \le P(A\cap B\cap C)$. Since the events have the same probability, $0.95 \le 1 - 3P(\overline{A})$. Thus, $P(A) \ge 0.9833$.

2.110 Define the events:
I: item is from line I II: item is from line II N: item is not defective
Then, $P(N) = P(N\cap(I\cup II)) = P(N\cap I) + P(N\cap II)$ = .92(.4) + .90(.6) = 0.908.

2.111 Define the following events:
A: buyer sees the magazine ad
B: buyer sees the TV ad
C: buyer purchases the product
The following are known: $P(A)$ = .02, $P(B)$ = .20, $P(A\cap B)$ = .01. Thus $P(A\cap B)$ = .21. Also, P(buyer sees no ad) = $P(\overline{A}\cap\overline{B}) = 1 - P(A\cup B) = 1 - 0.21 = 0.79$. Finally, it is known that $P(C\,|\,A\cup B) = 0.1$ and $P(C\,|\,\overline{A}\cap\overline{B}) = 1/3$. So, we can find $P(C)$ as $P(C) = P(C\cap(A\cup B)) + P(C\cap(\overline{A}\cap\overline{B}))$ = (1/3)(.21) + (.1)(.79) = 0.149.

2.112 **a.** P(aircraft undetected) = P(all three fail to detect) = (.02)(.02)(.02) = $(.02)^3$.
b. P(all three detect aircraft) = $(.98)^3$.

2.113 By independence, (.98)(.98)(.98)(.02) = $(.98)^3(.02)$.

2.114 Let T = {detects truth} and L = {detects lie}. The sample space is TT, TL, LT, LL. Since one suspect is guilty, assume the guilty suspect is questioned first:
a. $P(LL) = .95(.10) = 0.095$ **b.** $P(LT) = ..95(.9) = 0.885$
b. $P(TL) = .05(.10) = 0.005$ **d.** $1 - (.05)(.90) = 0.955$

2.115 By independence, $(.75)(.75)(.75)(.75) = (.75)^4$.

2.116 By the complement rule, P(system works) = 1 – P(system fails) = $1 - (.01)^3$.

2.117 **a.** From the description of the problem, there is a 50% chance a car will be rejected. To find the probability that three out of four will be rejected (i.e. the drivers chose team 2), note that there are $\binom{4}{3} = 4$ ways that three of the four cars are evaluated by team 2. Each one has probability (.5)(.5)(.5)(.5) of occurring, so the probability is $4(.5)^4 = 0.25$.

b. The probability that all four pass (i.e. all four are evaluated by team 1) is $(.5)^4 = 1/16$.

2.118 If the victim is to be saved, a proper donor must be found within eight minutes. The patient will be saved if the proper donor is found on the 1st, 2nd, 3rd, or 4th try. But, if the donor is found on the 2nd try, that implies he/she wasn't found on the 1st try. So, the probability of saving the patient is found by, letting A = {correct donor is found}:

$$P(\text{save}) = P(A) + P(\overline{A}A) + P(\overline{A}\overline{A}A) + P(\overline{A}\overline{A}\overline{A}A).$$

By independence, this is $.4 + .6(.4) + (.6)^2(.4) + (.6)^3(.4) = 0.8704$

2.119 **a.** Define the events: A: obtain a sum of 3 B: do not obtain a sum of 3 or 7
Since there are 36 possible rolls, $P(A) = 2/36$ and $P(B) = 28/36$. Obtaining a sum of 3 before a sum of 7 can happen on the 1st roll, the 2nd roll, the 3rd roll, etc. Using the events above, we can write these as A, BA, BBA, $BBBA$, etc. The probability of obtaining a sum of 3 before a sum of 7 is given by $P(A) + P(B)P(A) + [P(B)]^2P(A) + [P(B)]^3P(A) + \ldots$. (Here, we are using the fact that the rolls are independent.) This is an infinite sum, and it follows as a geometric series. Thus, $2/36 + (28/36)(2/36) + (28/36)^2(2/26) + \ldots = 1/4$.

b. Similar to part a. Define C: obtain a sum of 4 D: do not obtain a sum of 4 or 7
Then, $P(C) = 3/36$ and $P(D) = 27/36$. The probability of obtaining a 4 before a 7 is 1/3.

2.120 Denote the events G: good refrigerator D: defective refrigerator
a. If the last defective refrigerator is found on the 4th test, this means the first defective refrigerator was found on the 1st, 2nd, or 3rd test. So, the possibilities are DGGD, GDGD, and GGDD. So, the probability is $\left(\frac{2}{6}\right)\left(\frac{4}{5}\right)\left(\frac{3}{4}\right)\frac{1}{3}$. The probabilities associated with the other two events are identical to the first. So, the desired probability is $3\left(\frac{2}{6}\right)\left(\frac{4}{5}\right)\left(\frac{3}{4}\right)\frac{1}{3} = \frac{1}{5}$.

b. Here, the second defective refrigerator must be found on the 2nd, 3rd, or 4th test.
Define: A_1: second defective found on 2nd test
A_2: second defective found on 3rd test
A_3: second defective found on 4th test

Clearly, $P(A_1) = \left(\frac{2}{6}\right)\left(\frac{1}{5}\right) = \frac{1}{15}$. Also, $P(A_3) = \frac{1}{5}$ from part a. Note that $A_2 = \{DGD, GDD\}$. Thus, $P(A_2) = 2\left(\frac{2}{6}\right)\left(\frac{4}{5}\right)\left(\frac{1}{4}\right) = \frac{2}{15}$. So, $P(A_1) + P(A_2) + P(A_3) = 2/5$.

c. Define: B_1: second defective found on 3rd test
B_2: second defective found on 4th test

Clearly, $P(B_1) = 1/4$ and $P(B_2) = (3/4)(1/3) = 1/4$. So, $P(B_1) + P(B_2) = 1/2$.

2.121 **a.** $1/n$
b. $\frac{n-1}{n}\cdot\frac{1}{n-1} = 1/n$. $\frac{n-1}{n}\cdot\frac{n-2}{n-1}\cdot\frac{1}{n-2} = 1/n$.
c. P(gain access) = P(first try) + P(second try) + P(third try) = 3/7.

2.122 Applet exercise (answers vary).

2.123 Applet exercise (answers vary).

2.124 Define the events for the voter: D: democrat R: republican F: favors issue

$$P(D \mid F) = \frac{P(F \mid D)P(D)}{P(F \mid D)P(D) + P(F \mid R)P(R)} = \frac{.7(.6)}{.7(.6)+.3(.4)} = 7/9$$

2.125 Define the events for the person: D: has the disease H: test indicates the disease
Thus, $P(H|D) = .9$, $P(\overline{H} \mid \overline{D}) = .9$, $P(D) = .01$, and $P(\overline{D}) = .99$. Thus,

$$P(D \mid H) = \frac{P(H \mid D)P(D)}{P(H \mid D)P(D) + P(H \mid \overline{D})P(\overline{D})} = 1/12.$$

2.126 **a.** (.95*.01)/(.95*.01 + .1*.99) = 0.08756.
b. .99*.01/(.99*.01 + .1*.99) = 1/11.
c. Only a small percentage of the population has the disease.
d. If the specificity is .99, the positive predictive value is .5.
e. The sensitivity and specificity should be as close to 1 as possible.

2.127 **a.** .9*.4/(.9*.4 + .1*.6) = 0.857.
b. A larger proportion of the population has the disease, so the numerator and denominator values are closer.
c. No; if the sensitivity is 1, the largest value for the positive predictive value is .8696.
d. Yes, by increasing the specificity.
e. The specificity is more important with tests used for rare diseases.

2.128 **a.** Let $P(A \mid B) = P(A \mid \overline{B}) = p$. By the Law of Total Probability,

$$P(A) = P(A \mid B)P(B) + P(A \mid \overline{B})P(B) = p\big(P(B) + P(\overline{B})\big) = p.$$

Thus, A and B are independent.

b. $P(A) = P(A \mid C)P(C) + P(A \mid \overline{C})P(\overline{C}) > P(B \mid C)P(C) + P(B \mid \overline{C})P(\overline{C}) = P(B)$.

2.129 Define the events: P: positive response M: male respondent F: female respondent
$P(P|F) = .7$, $P(P|M) = .4$, $P(M) = .25$. Using Bayes' rule,

$$P(M \mid \overline{P}) = \frac{P(\overline{P} \mid M)P(M)}{P(\overline{P} \mid M)P(M) + P(\overline{P} \mid F)P(F)} = \frac{.6(.25)}{.6(.25)+.3(.75)} = 0.4.$$

2.130 Define the events: C: contract lung cancer S: worked in a shipyard
Thus, $P(S|C) = .22$, $P(S \mid \overline{C}) = .14$, and $P(C) = .0004$. Using Bayes' rule,

$$P(C \mid S) = \frac{P(S \mid C)P(C)}{P(S \mid C)P(C) + P(S \mid \overline{C})P(\overline{C})} = \frac{.22(.0004)}{.22(.0004)+.14(.9996)} = 0.0006.$$

2.131 The additive law of probability gives that $P(A\Delta B) = P(A \cap \overline{B}) + P(\overline{A} \cap B)$. Also, A and B can be written as the union of two disjoint sets: $A = (A \cap \overline{B}) \cup (A \cap B)$ and $B = (\overline{A} \cap B) \cup (A \cap B)$. Thus, $P(A \cap \overline{B}) = P(A) - P(A \cap B)$ and $P(\overline{A} \cap B) = P(B) - P(A \cap B)$. Thus, $P(A\Delta B) = P(A) + P(B) - 2P(A \cap B)$.

2.132 For i = 1, 2, 3, let F_i represent the event that the plane is found in region i and N_i be the complement. Also R_i is the event the plane is in region i. Then $P(F_i|R_i) = 1 - \alpha_i$ and $P(R_i) = 1/3$ for all i. Then,

a. $$P(R_1 \mid N_1) = \frac{P(N_1 \mid R_1)P(R_1)}{P(N_1 \mid R_1)P(R_1) + P(N_1 \mid R_2)P(R_2) + P(N_1 \mid R_3)P(R_3)} = \frac{\alpha_1 \frac{1}{3}}{\alpha_1 \frac{1}{3} + \frac{1}{3} + \frac{1}{3}}$$
$$= \frac{\alpha_1}{\alpha_1 + 2}.$$

b. Similarly, $P(R_2 \mid N_1) = \dfrac{1}{\alpha_1 + 2}$ and c. $P(R_3 \mid N_1) = \dfrac{1}{\alpha_1 + 2}$.

2.133 Define the events: G: student guesses C: student is correct

$$P(\overline{G} \mid C) = \frac{P(C \mid \overline{G})P(\overline{G})}{P(C \mid \overline{G})P(\overline{G}) + P(C \mid G)P(G)} = \frac{1(.8)}{1(.8)+.25(.2)} = 0.9412.$$

2.134 Define F as "failure to learn. Then, $P(F|A) = .2$, $P(F|B) = .1$, $P(A) = .7$, $P(B) = .3$. By Bayes' rule, $P(A|F) = 14/17$.

2.135 Let M = major airline, P = private airline, C = commercial airline, B = travel for business
a. $P(B) = P(B|M)P(M) + P(B|P)P(P) + P(B|C)P(C) = .6(.5) + .3(.6) + .1(.9) = 0.57$.
b. $P(B \cap P) = P(B|P)P(P) = .3(.6) = 0.18$.
c. $P(P|B) = P(B \cap P)/P(B) = .18/.57 = 0.3158$.
d. $P(B|C) = 0.90$.

2.136 Let A = woman's name is selected from list 1, B = woman's name is selected from list 2.
Thus, $P(A) = 5/7$, $P(\overline{B} \mid A) = 2/3$, $P(\overline{B} \mid \overline{A}) = 7/9$.

$$P(A \mid \overline{B}) = \frac{P(\overline{B} \mid A)P(A)}{P(\overline{B} \mid A)P(A) + P(\overline{B} \mid \overline{A})P(\overline{A})} = \frac{\frac{2}{3}\left(\frac{5}{7}\right)}{\frac{2}{3}\left(\frac{5}{7}\right) + \frac{7}{9}\left(\frac{2}{7}\right)} = \frac{30}{44}.$$

2.137 Let A = {both balls are white}, and for i = 1, 2, … 5

A_i = both balls selected from bowl i are white. Then $\bigcup A_i = A$.

B_i = bowl i is selected. Then, $P(B_i) = .2$ for all i.

a. $P(A) = \sum P(A_i \mid B_i)P(B_i) = \frac{1}{5}\left[0 + \frac{2}{5}\left(\frac{1}{4}\right) + \frac{3}{5}\left(\frac{2}{4}\right) + \frac{4}{5}\left(\frac{3}{4}\right) + 1\right]$ = 2/5.

b. Using Bayes' rule, $P(B_3|A) = \dfrac{\frac{3}{50}}{\frac{2}{50}}$ = 3/20.

2.138 Define the events:

A: the player wins
B_i: a sum of i on first toss
C_k: obtain a sum of k before obtaining a 7

Now, $P(A) = \sum_{i=1}^{12} P(A \cap B_i)$. We have that $P(A \cap B_2) = P(A \cap B_3) = P(A \cap B_{12})$ = 0.

Also, $P(A \cap B_7) = P(B_7) = \frac{6}{36}$, $P(A \cap B_{11}) = P(B_{11}) = \frac{2}{36}$.

Now, $P(A \cap B_4) = P(C_4 \cap B_7) = P(C_4)P(B_7) = \frac{1}{3}\left(\frac{3}{36}\right) = \frac{3}{36}$ (using independence Ex. 119).

Similarly, $P(C_5) = P(C_9) = \frac{4}{10}$, $P(C_6) = P(C_8) = \frac{5}{11}$, and $P(C_{10}) = \frac{3}{9}$.

Thus, $P(A \cap B_5) = P(A \cap B_9) = \frac{2}{45}$, $P(A \cap B_6) = P(A \cap B_8) = \frac{25}{396}$, $P(A \cap B_{10}) = \frac{1}{36}$.

Putting all of this together, $P(A)$ = 0.493.

2.139 From Ex. 1.112, $P(Y = 0) = (.02)^3$ and $P(Y = 3) = (.98)^3$. The event $Y = 1$ are the events *FDF*, *DFF*, and *FFD*, each having probability $(.02)^2(.98)$. So, $P(Y = 1) = 3(.02)^2(.98)$. Similarly, $P(Y = 2) = 3(.02)(.98)^2$.

2.140 The total number of ways to select 3 from 6 refrigerators is $\binom{6}{3}$ = 20. The total number of ways to select y defectives and 3 – y nondefectives is $\binom{2}{y}\binom{4}{3-y}$, y = 0, 1, 2. So,

$P(Y = 0) = \dfrac{\binom{2}{0}\binom{4}{3}}{20}$ = 4/20, $P(Y = 1)$ = 4/20, and $P(Y = 2)$ = 12/20.

2.141 The events $Y = 2$, $Y = 3$, and $Y = 4$ were found in Ex. 2.120 to have probabilities 1/15, 2/15, and 3/15 (respectively). The event $Y = 5$ can occur in four ways:

DGGGD *GDGGD* *GGDGD* *GGGDD*

Each of these possibilities has probability 1/15, so that P(Y = 5) = 4/15. By the complement rule, P(Y = 6) = 5/15.

2.142 Each position has probability 1/4, so every ordering of two positions (from two spins) has probability 1/16. The values for Y are 2, 3. P(Y = 2) = $\binom{4}{2}\frac{1}{16}$ = 3/4. So, P(Y = 3) = 1/4.

2.143 Since $P(B) = P(B \cap A) + P(B \cap \overline{A})$, 1 = $\dfrac{P(B \cap A)}{P(B)} + \dfrac{P(B \cap \overline{A})}{P(B)} = P(A \mid B) + P(\overline{A} \mid B)$.

2.144 **a.** S = {16 possibilities of drawing 0 to 4 of the sample points}
b. $\binom{4}{0}+\binom{4}{1}+\binom{4}{2}+\binom{4}{3}+\binom{4}{4} = 1+4+6+4+1 = 16 = 2^4$.
c. $A \cup B = \{E_1, E_2, E_3, E_4\}$, $A \cap B = \{E_2\}$, $\overline{A} \cap \overline{B} = \emptyset$, $\overline{A} \cup B = \{E_2, E_4\}$.

2.145 All 18 orderings are possible, so the total number of orderings is 18!

2.146 There are $\binom{52}{5}$ ways to draw 5 cards from the deck. For each suit, there are $\binom{13}{5}$ ways to select 5 cards. Since there are 4 suits, the probability is $4\binom{13}{5}\Big/\binom{52}{5}$ = 0.00248.

2.147 The gambler will have a full house if he is dealt {two kings} or {an ace and a king} (there are 47 cards remaining in the deck, two of which are aces and three are kings).
The probabilities of these two events are $\binom{3}{2}\Big/\binom{47}{2}$ and $\binom{3}{1}\binom{2}{1}\Big/\binom{47}{2}$, respectively.
So, the probability of a full house is $\binom{3}{2}\Big/\binom{47}{2} + \binom{3}{1}\binom{2}{1}\Big/\binom{47}{2}$ = 0.0083.

2.148 Note that $\binom{12}{4} = 495$. P(each supplier has at least one component tested) is given by

$$\frac{\binom{3}{2}\binom{4}{1}\binom{5}{1}+\binom{3}{1}\binom{4}{2}\binom{5}{1}+\binom{3}{1}\binom{4}{1}\binom{5}{2}}{495} = 270/475 = 0.545.$$

2.149 Let A be the event that the person has symptom A and define B similarly. Then
a. $P(\overline{A \cup B}) = P(\overline{A} \cap \overline{B}) = 0.4$
b. $P(A \cup B) = 1 - P(\overline{A \cup B}) = 0.6$.
c. $P(A \cap B \mid B) = P(A \cap B)/P(B) = .1/.4 = 0.25$

2.150 $P(Y = 0) = 0.4$, $P(Y = 1) = 0.2 + 0.3 = 0.5$, $P(Y = 2) = 0.1$.

2.151 The probability that team A wins in 5 games is $p^4(1 - p)$ and the probability that team B wins in 5 games is $p(1 - p)^4$. Since there are 4 ways that each team can win in 5 games, the probability is $4[p^4(1 - p) + p(1 - p)^4]$.

2.152 Let R denote the event that the specimen turns red and N denote the event that the specimen contains nitrates.

a. $P(R) = P(R \mid N)P(N) + P(R \mid \overline{N})P(\overline{N})$ = .95(.3) + .1(.7) = 0.355.

b. Using Bayes' rule, $P(N|R)$ = .95(.3)/.355 = 0.803.

2.153 Using Bayes' rule,

$$P(I_1 \mid H) = \frac{P(H \mid I_1)P(I_1)}{P(H \mid I_1)P(I_1) + P(H \mid I_2)P(I_2) + P(H \mid I_3)P(I_3)} = 0.313.$$

2.154 Let Y = the number of pairs chosen. Then, the possible values are 0, 1, and 2.

a. There are $\binom{10}{4}$ = 210 ways to choose 4 socks from 10 and there are $\binom{5}{4}2^4$ = 80 ways to pick 4 non-matching socks. So, $P(Y = 0)$ = 80/210.

b. Generalizing the above, the probability is $\binom{n}{2r}2^{2r} \Big/ \binom{2n}{2r}$.

2.155 **a.** $P(A)$ = .25 + .1 + .05 + .1 = .5
b. $P(A \cap B)$ = .1 + .05 = 0.15.
c. 0.10
d. Using the result from Ex. 2.80, $\frac{.25 + .25 - .15}{.4}$ = 0.875.

2.156 **a.**
i. 1 – 5686/97900 = 0.942
ii. (97900 – 43354)/97900 = 0.557
ii. 10560/14113 = 0.748
iv. (646+375+568)/11533 = 0.138

b. If the US population in 2002 was known, this could be used to divide into the total number of deaths in 2002 to give a probability.

2.157 Let D denote death due to lung cancer and S denote being a smoker. Thus:
$P(D) = P(D \mid S)P(S) + P(D \mid \overline{S})P(\overline{S}) = 10P(D \mid \overline{S})(.2) + P(D \mid \overline{S})(.8)$ = 0.006. Thus, $P(D \mid S) = 0.021$.

2.158 Let W denote the even that the first ball is white and B denote the event that the second ball is black. Then:

$$P(W \mid B) = \frac{P(B \mid W)P(W)}{P(B \mid W)P(W) + P(B \mid \overline{W})P(\overline{W})} = \frac{\frac{b}{w+b+n}\left(\frac{w}{w+b}\right)}{\frac{b}{w+b+n}\left(\frac{w}{w+b}\right) + \frac{b+n}{w+b+n}\left(\frac{b}{w+b}\right)} = \frac{w}{w+b+n}$$

2.159 Note that $S = S \cup \emptyset$, and S and $\emptyset$ are disjoint. So, 1 = $P(S) = P(S) + P(\emptyset)$ and therefore $P(\emptyset) = 0$.

2.160 There are 10 nondefective and 2 defective tubes that have been drawn from the machine, and number of distinct arrangements is $\binom{12}{2} = 66$.

a. The probability of observing the specific arrangement is 1/66.

b. There are two such arrangements that consist of "runs." In addition to what was given in part (a), the other is *DDNNNNNNNNNN*. Thus, the probability of two runs is 2/66 = 1/33.

2.161 We must find $P(R \le 3) = P(R = 3) + P(R = 2)$, since the minimum value for R is 2. Id the two D's occurs on consecutive trials (but not in positions 1 and 2 or 11 and 12), there are 9 such arrangements. The only other case is a defective in position 1 and 12, so that (combining with Ex. 2.160 with $R = 2$), there are 12 possibilities. So, $P(R \le 3) = 12/66$.

2.162 There are 9! ways for the attendant to park the cars. There are 3! ways to park the expensive cars together and there are 7 ways the expensive cars can be next to each other in the 9 spaces. So, the probability is 7(3!)/9! = 1/12.

2.163 Let A be the event that current flows in design A and let B be defined similarly. Design A will function if (1 or 2) & (3 or 4) operate. Design B will function if (1 & 3) or (2 & 4) operate. Denote the event R_i = {relay i operates properly}, i = 1, 2, 3, 4. So, using independence and the addition rule,

$$P(A) = (R_1 \cup R_2) \cap (R_3 \cup R_4) = (.9 + .9 - .9^2)(.9 + .9 - .9^2) = 0.9801.$$

$$P(B) = (R_1 \cap R_3) \cup (R_2 \cap R_4) = .9^2 + .9^2 - (.9^2)^2 = .9639.$$

So, design A has the higher probability.

2.164 Using the notation from Ex. 2.163, $P(R_1 \cap R_4 \mid A) = P(R_1 \cap R_4 \cap A)/P(A)$.
Note that $R_1 \cap R_4 \cap A = R_1 \cap R_4$, since the event $R_1 \cap R_4$ represents a path for the current to flow. The probability of this above event is $.9^2 = .81$, and the conditional probability is in question is .81/.9801 = 0.8264.

2.165 Using the notation from Ex. 2.163, $P(R_1 \cap R_4 \mid B) = P(R_1 \cap R_4 \cap B)/P(B)$.
$R_1 \cap R_4 \cap B = (R_1 \cap R_4) \cap (R_1 \cap R_3) \cup (R_2 \cap R_4) = (R_1 \cap R_4 \cap R_3) \cup (R_2 \cap R_4)$. The probability of the above event is $.9^3 + .9^2 - .9^4 = 0.8829$. So, the conditional probability in question is .8829/.9639 = 0.916.

2.166 There are $\binom{8}{4}$ = 70 ways to choose the tires. If the best tire the customer has is ranked #3, the other three tires are from ranks 4, 5, 6, 7, 8. There are $\binom{5}{3}$ = 10 ways to select three tires from these five, so that the probability is 10/70 = 1/7.

2.167 If $Y = 1$, the customer chose the best tire. There are $\binom{7}{3}$ = 35 ways to choose the remaining tires, so $P(Y = 1) = 35/70 = .5$.
If $Y = 2$, the customer chose the second best tire. There are $\binom{6}{3}$ = 20 ways to choose the remaining tires, so $P(Y = 2) = 20/70 = 2/7$. Using the same logic, $P(Y = 4) = 4/70$ and so $P(Y = 5) = 1/70$.

2.168 **a.** The two other tires picked by the customer must have ranks 4, 5, or 6. So, there are $\binom{3}{2}$ = 3 ways to do this. So, the probability is 3/70.

b. There are four ways the range can be 4: #1 to #5, #2 to #6, #3 to #7, and #4 to #8. Each has probability 3/70 (as found in part **a**). So, $P(R = 4) = 12/70$.

c. Similar to parts a and b, $P(R = 3) = 5/70$, $P(R = 5) = 18/70$, $P(R = 6) = 20/70$, and $P(R = 7) = 15/70$.

2.169 **a.** For each beer drinker, there are 4! = 24 ways to rank the beers. So there are 24^3 = 13,824 total sample points.

b. In order to achieve a combined score of 4 our less, the given beer may receive at most one score of two and the rest being one. Consider brand A. If a beer drinker assigns a one to A there are still 3! = 6 ways to rank the other brands. So, there are 6^3 ways for brand A to be assigned all ones. Similarly, brand A can be assigned two ones and one two in $3(3!)^3$ ways. Thus, some beer may earn a total rank less than or equal to four in $4[6^3 + 3(3!)^3] = 3456$ ways. So, the probability is 3456/13824 = 0.25.

2.170 There are $\binom{7}{3}$ = 35 ways to select three names from seven. If the first name on the list is included, the other two names can be picked $\binom{6}{2}$ = 15 ways. So, the probability is 15/35 = 3/7.

2.171 It is stated that the probability that Skylab will hit someone is (unconditionally) 1/150, without regard to where that person lives. If one wants to know the probability condition on living in a certain area, it is not possible to determine.

2.172 Only $P(A \mid B + P(\overline{A} \mid B) = 1$ is <u>true</u> for any events *A* and *B*.

2.173 Define the events: *D*: item is defective *C*: item goes through inspection
Thus $P(D) = .1$, $P(C|D) = .6$, and $P(C \mid \overline{D}) = .2$. Thus,
$$P(D \mid C) = \frac{P(C \mid D)P(D)}{P(C \mid D)P(D) + P(C \mid \overline{D})P(\overline{D})} = .25.$$

2.174 Let A = athlete disqualified previously B = athlete disqualified next term
Then, we know $P(B \mid \overline{A}) = .15, P(B \mid A) = .5, P(A) = .3$. To find $P(B)$, use the law of total probability: $P(B) = .3(.5) + .7(.15) = 0.255$.

2.175 Note that $P(A) = P(B) = P(C) = .5$. But, $P(A \cap B \cap C) = P(HH) = .25 \neq (.5)^3$. So, they are not mutually independent.

2.176 **a.** $P[(A \cup B) \cap C)] = P[(A \cap C) \cup (B \cap C)] = P(A \cap C) + P(B \cap C) - P(A \cap B \cap C)$
$= P(A)P(C) + P(B)P(C) - P(A)P(B)P(C) = [P(A) + P(B) - P(A)P(B)]P(C)$
$= P(A \cap B)P(C)$.
b. Similar to part a above.

2.177 **a.** P(no injury in 50 jumps) = $(49/50)^{50} = 0.364$.
b. P(at least one injury in 50 jumps) = 1 – P(no injury in 50 jumps) = 0.636.
c. P(no injury in n jumps) = $(49/50)^n \geq 0.60$, so n is at most 25.

2.178 Define the events: *E*: person is exposed to the flu *F*: person gets the flu
Consider two employees, one of who is inoculated and one not. The probability of interest is the probability that at least one contracts the flu. Consider the complement:
P(at least one gets the flu) = 1 – P(neither employee gets the flu).
For the inoculated employee: $P(\overline{F}) = P(\overline{F} \cap E) + P(\overline{F} \cap \overline{E}) = .8(.6) + 1(.4) = 0.88$.
For the non-inoculated employee: $P(\overline{F}) = P(\overline{F} \cap E) + P(\overline{F} \cap \overline{E}) = .1(.6) + 1(.4) = 0.46$.
So, P(at least one gets the flu) = 1 – .88(.46) = 0.5952

2.179 **a.** The gamblers break even if each win three times and lose three times. Considering the possible sequences of "wins" and "losses", there are $\binom{6}{3} = 20$ possible orderings. Since each has probability $\left(\tfrac{1}{2}\right)^6$, the probability of breaking even is $20\left(\tfrac{1}{2}\right)^6 = 0.3125$.

b. In order for this event to occur, the gambler Jones must have \$11 at trial 9 and must win on trial 10. So, in the nine remaining trials, seven "wins" and two "losses" must be placed. So, there are $\binom{9}{2} = 36$ ways to do this. However, this includes cases where Jones would win before the 10th trial. Now, Jones can only win the game on an even trial (since he must gain \$6). Included in the 36 possibilities, there are three ways Jones could win on trial 6: *WWWWWWWLL*, *WWWWWWLLW*, *WWWWWWLWL*, and there are six ways Jones could win on trial 8: *LWWWWWWWL*, *WLWWWWWWL*, *WWLWWWWWL*, *WWWLWWWWL*, *WWWWLWWWL*, *WWWWWLWWL*. So, these nine cases must be removed from the 36. So, the probability is $27\left(\frac{1}{2}\right)^{10}$.

2.180 **a.** If the patrolman starts in the center of the 16x16 square grid, there are 4^8 possible paths to take. Only four of these will result in reaching the boundary. Since all possible paths are equally likely, the probability is $4/4^8 = 1/4^7$.

b. Assume the patrolman begins by walking north. There are nine possible paths that will bring him back to the starting point: *NNSS*, *NSNS*, *NSSN*, *NESW*, *NWSE*, *NWES*, *NEWS*, *NSEW*, *NSWE*. By symmetry, there are nine possible paths for each of north, south, east, and west as the starting direction. Thus, there are 36 paths in total that result in returning to the starting point. So, the probability is $36/4^8 = 9/4^7$.

2.181 We will represent the n balls as 0's and create the N boxes by placing bars (|) between the 0's. For example if there are 6 balls and 4 boxes, the arrangement

0|00||000

represents one ball in box 1, two balls in box 2, no balls in box 3, and three balls in box 4. Note that six 0's were need but only 3 bars. In general, n 0's and $N - 1$ bars are needed to represent each possible placement of n balls in N boxes. Thus, there are $\binom{N+n-1}{N-1}$ ways to arrange the 0's and bars. Now, if no two bars are placed next to each other, no box will be empty. So, the $N - 1$ bars must be placed in the $n - 1$ spaces between the 0's. The total number of ways to do this is $\binom{n-1}{N-1}$, so that the probability is as given in the problem.

Chapter 3: Discrete Random Variables and Their Probability Distributions

3.1 $P(Y = 0) = P(\text{no impurities}) = .2$, $P(Y = 1) = P(\text{exactly one impurity}) = .7$, $P(Y = 2) = .1$.

3.2 We know that $P(HH) = P(TT) = P(HT) = P(TH) = 0.25$. So, $P(Y = -1) = .5$, $P(Y = 1) = .25 = P(Y = 2)$.

3.3 $p(2) = P(DD) = 1/6$, $p(3) = P(DGD) + P(GDD) = 2(2/4)(2/3)(1/2) = 2/6$, $p(4) = P(GGDD) + P(DGGD) + P(GDGD) = 3(2/4)(1/3)(2/2) = 1/2$.

3.4 Define the events: A: value 1 fails B: valve 2 fails C: valve 3 fails

$P(Y = 2) = P(\overline{A} \cap \overline{B} \cap \overline{C}) = .8^3 = 0.512$

$P(Y = 0) = P(A \cap (B \cup C)) = P(A)P(B \cup C) = .2(.2 + .2 - .2^2) = 0.072$.

Thus, $P(Y = 1) = 1 - .512 - .072 = 0.416$.

3.5 There are $3! = 6$ possible ways to assign the words to the pictures. Of these, one is a perfect match, three have one match, and two have zero matches. Thus,

$p(0) = 2/6$, $p(1) = 3/6$, $p(3) = 1/6$.

3.6 There are $\binom{5}{2} = 10$ sample points, and all are equally likely: (1,2), (1,3), (1,4), (1,5), (2,3), (2,4), (2,5), (3,4), (3,5), (4,5).

a. $p(2) = .1$, $p(3) = .2$, $p(4) = .3$, $p(5) = .4$.

b. $p(3) = .1$, $p(4) = .1$, $p(5) = .2$, $p(6) = .2$, $p(7) = .2$, $p(8) = .1$, $p(9) = .1$.

3.7 There are $3^3 = 27$ ways to place the three balls into the three bowls. Let Y = # of empty bowls. Then:

$p(0) = P(\text{no bowls are empty}) = \frac{3!}{27} = \frac{6}{27}$

$p(2) = P(\text{2 bowls are empty}) = \frac{3}{27}$

$p(1) = P(\text{1 bowl is empty}) = 1 - \frac{6}{27} - \frac{3}{27} = \frac{18}{27}$.

3.8 Note that the number of cells cannot be odd.

$p(0) = P$(no cells in the next generation) $= P$(the first cell dies or the first cell splits and both die) $= .1 + .9(.1)(.1) = 0.109$

$p(4) = P$(four cells in the next generation) = P(the first cell splits and both created cells split) $= .9(.9)(.9) = 0.729$.

$p(2) = 1 – .109 – .729 = 0.162$.

3.9 The random variable Y takes on vales 0, 1, 2, and 3.

a. Let E denote an error on a single entry and let N denote no error. There are 8 sample points: EEE, EEN, ENE, NEE, ENN, NEN, NNE, NNN. With $P(E) = .05$ and $P(N) = .95$ and assuming independence:

$P(Y = 3) = (.05)^3 = 0.000125$ $P(Y = 2) = 3(.05)2(.95) = 0.007125$

$P(Y = 1) = 3(.05)^2(.95) = 0.135375$ $P(Y = 0) = (.95)^3 = 0.857375$.

b. The graph is omitted.
c. $P(Y > 1) = P(Y = 2) + P(Y = 3) = 0.00725$.

3.10 Denote R as the event a rental occurs on a given day and N denotes no rental. Thus, the sequence of interest is RR, RNR, $RNNR$, $RNNNR$, … . Consider the position immediately following the first R: it is filled by an R with probability .2 and by an N with probability .8. Thus, $P(Y = 0) = .2$, $P(Y = 1) = .8(.2) = .16$, $P(Y = 2) = .128$, … . In general,

$$P(Y = y) = .2(.8)^y,\ y = 0, 1, 2, \dots .$$

3.11 There is a 1/3 chance a person has O^+ blood and 2/3 they do not. Similarly, there is a 1/15 chance a person has O^- blood and 14/15 chance they do not. Assuming the donors are randomly selected, if X = # of O^+ blood donors and Y = # of O^- blood donors, the probability distributions are

	0	1	2	3
$p(x)$	$(2/3)^3 = 8/27$	$3(2/3)^2(1/3) = 12/27$	$3(2/3)(1/3)^2 = 6/27$	$(1/3)^3 = 1/27$
$p(y)$	2744/3375	196/3375	14/3375	1/3375

Note that $Z = X + Y$ = # will type O blood. The probability a donor will have type O blood is $1/3 + 1/15 = 6/15 = 2/5$. The probability distribution for Z is

	0	1	2	3
$p(z)$	$(2/5)^3 = 27/125$	$3(2/5)^2(3/5) = 54/27$	$3(2/5)(3/5)^2 = 36/125$	$(3/5)^3 = 27/125$

3.12 $E(Y) = 1(.4) + 2(.3) + 3(.2) + 4(.1) = 2.0$
$E(1/Y) = 1(.4) + 1/2(.3) + 1/3(.2) + 1/4(.1) = 0.6417$
$E(Y^2 - 1) = E(Y^2) - 1 = [1(.4) + 2^2(.3) + 3^2(.2) + 4^2(.1)] - 1 = 5 - 1 = 4$.
$V(Y) = E(Y^2) = [E(Y)]^2 = 5 - 22 = 1$.

3.13 $E(Y) = -1(1/2) + 1(1/4) + 2(1/4) = 1/4$
$E(Y^2) = (-1)^2(1/2) + 1^2(1/4) + 2^2(1/4) = 7/4$
$V(Y) = 7/4 - (1/4)^2 = 27/16$.
Let C = cost of play, then the net winnings is $Y - C$. If $E(Y - C) = 0$, $C = 1/4$.

3.14 **a.** $\mu = E(Y) = 3(.03) + 4(.05) + 5(.07) + \dots + 13(.01) = 7.9$
b. $\sigma^2 = V(Y) = E(Y^2) - [E(Y)]^2 = 3^2(.03) + 4^2(.05) + 5^2(.07) + \dots + 13^2(.01) - 7.9^2 = 67.14 - 62.41 = 4.73$. So, $\sigma = 2.17$.
c. $(\mu - 2\sigma, \mu + 2\sigma) = (3.56, 12.24)$. So, $P(3.56 < Y < 12.24) = P(4 \le Y \le 12) = .05 + .07 + .10 + .14 + .20 + .18 + .12 + .07 + .03 = 0.96$.

3.15 **a**. $p(0) = P(Y = 0) = (.48)^3 = .1106$, $p(1) = P(Y = 1) = 3(.48)^2(.52) = .3594$, $p(2) = P(Y = 2) = 3(.48)(.52)^2 = .3894$, $p(3) = P(Y = 3) = (.52)^3 = .1406$.
b. The graph is omitted.
c. $P(Y = 1) = .3594$.

d. $\mu = E(Y) = 0(.1106) + 1(.3594) + 2(.3894) + 3(.1406) = 1.56$,
$\sigma^2 = V(Y) = E(Y^2) - [E(Y)]^2 = 0^2(.1106) + 1^2(.3594) + 2^2(.3894) + 3^2(.1406) - 1.56^2 = 3.1824 - 2.4336 = .7488$. So, $\sigma = 0.8653$.
e. $(\mu - 2\sigma, \mu + 2\sigma) = (-.1706, 3.2906)$. So, $P(-.1706 < Y < 3.2906) = P(0 \le Y \le 3) = 1$.

3.16 As shown in Ex. 2.121, $P(Y = y) = 1/n$ for $y = 1, 2, \ldots, n$. Thus, $E(Y) = \frac{1}{n}\sum_{y=1}^{n} y = \frac{n+1}{2}$.

$$E(Y^2) = \frac{1}{n}\sum_{y=1}^{n} y^2 = \frac{(n+1)(2n+1)}{6}. \text{ So, } V(Y) = \frac{(n+1)(2n+1)}{6} - \left(\frac{n+1}{2}\right)^2 = \frac{(n+1)(n-1)}{12} = \frac{n^2-1}{12}.$$

3.17 $\mu = E(Y) = 0(6/27) + 1(18/27) + 2(3/27) = 24/27 = .889$
$\sigma^2 = V(Y) = E(Y^2) - [E(Y)]^2 = 0^2(6/27) + 1^2(18/27) + 2^2(3/27) - (24/27)^2 = 30/27 - 576/729 = .321$. So, $\sigma = 0.567$
For $(\mu - 2\sigma, \mu + 2\sigma) = (-.245, 2.023)$. So, $P(-.245 < Y < 2.023) = P(0 \le Y \le 2) = 1$.

3.18 $\mu = E(Y) = 0(.109) + 2(.162) + 4(.729) = 3.24$.

3.19 Let P be a random variable that represents the company's profit. Then, $P = C - 15$ with probability 98/100 and $P = C - 15 - 1000$ with probability 2/100. Then,
$E(P) = (C - 15)(98/100) + (C - 15 - 1000)(2/100) = 50$. Thus, $C = \$85$.

3.20 With probability .3 the volume is 8(10)(30) = 2400. With probability .7 the volume is 8*10*40 = 3200. Then, the mean is .3(2400) + .7(3200) = 2960.

3.21 Note that $E(N) = E(8\pi R^2) = 8\pi E(R^2)$. So, $E(R^2) = 21^2(.05) + 22^2(.20) + \ldots + 26^2(.05) = 549.1$. Therefore $E(N) = 8\pi(549.1) = 13{,}800.388$.

3.22 Note that $p(y) = P(Y = y) = 1/6$ for $y = 1, 2, \ldots, 6$. This is similar to Ex. 3.16 with n = 6. So, $E(Y) = 3.5$ and $V(Y) = 2.9167$.

3.23 Define G to be the gain to a person in drawing one card. The possible values for G are \$15, \$5, or \$–4 with probabilities 3/13, 2/13, and 9/13 respectively. So,
$E(G) = 15(3/13) + 5(2/13) - 4(9/13) = 4/13$ (roughly \$.31).

3.24 The probability distribution for Y = number of bottles with serious flaws is:

$p(y)$	0	1	2
y	.81	.18	.01

Thus, $E(Y) = 0(.81) + 1(.18) + 2(.01) = 0.20$ and $V(Y) = 0^2(.81) + 1^2(.18) + 2^2(.01) - (.20)^2 = 0.18$.

3.25 Let X_1 = # of contracts assigned to firm 1; X_2 = # of contracts assigned to firm 2. The sample space for the experiment is {(I,I), (I,II), (I,III), (II,I), (II,II), (II,III), (III,I), (III,II), (III,III)}, each with probability 1/9. So, the probability distributions for X_1 and X_2 are:

x_1	0	1	2
$p(x_1)$	4/9	4/9	1/9

x_2	0	1	2
$p(x_2)$	4/9	4/9	1/9

Thus, $E(X_1) = E(X_2) = 2/3$. The expected profit for the owner of both firms is given by $90000(2/3 + 2/3) = \$120,000$.

3.26 The random variable Y = daily sales can have values \$0, \$50,000 and \$100,000.
If $Y = 0$, either the salesperson contacted only one customer and failed to make a sale or the salesperson contacted two customers and failed to make both sales. Thus $P(Y = 0) = 1/3(9/10) + 2/3(9/10)(9/10) = 252/300$.
If $Y = 2$, the salesperson contacted to customers and made both sales. So, $P(Y = 2) = 2/3(1/10)(1/10) = 2/300$.
Therefore, $P(Y = 1) = 1 – 252/300 – 2/300 = 46/300$.
Then, $E(Y) = 0(252/300) + 50000(46/300) + 100000(2/300) = 25000/3$ (or \$8333.33).
$V(Y)$ =380,561,111 and σ = \$19,507.98.

3.27 Let Y = the payout on an individual policy. Then, $P(Y = 85,000) = .001$, $P(Y = 42,500) = .01$, and $P(Y = 0) = .989$. Let C represent the premium the insurance company charges. Then, the company's net gain/loss is given by $C – Y$. If $E(C – Y) = 0$, $E(Y) = C$. Thus, $E(Y) = 85000(.001) + 42500(.01) + 0(.989) = 510 = C$.

3.28 Using the probability distribution found in Ex. 3.3, $E(Y) = 2(1/6) + 3(2/6) + 4(3/6) = 20/6$. The cost for testing and repairing is given by $2Y + 4$. So, $E(2Y + 4) = 2(20/6) + 4 = 64/6$.

3.29 $$\sum_{k=1}^{\infty} P(Y \ge k) = \sum_{k=1}^{\infty}\sum_{j=k}^{\infty} P(Y = k) = \sum_{k=1}^{\infty}\sum_{j=k}^{\infty} p(j) = \sum_{j=1}^{\infty}\sum_{k=1}^{j} p(j) = \sum_{j=1}^{\infty} j \cdot p(j) = \sum_{y=1}^{\infty} y \cdot p(y) = E(Y).$$

3.30 **a.** The mean of X will be larger than the mean of Y.
b. $E(X) = E(Y + 1) = E(Y) + 1 = \mu + 1$.
c. The variances of X and Y will be the same (the addition of 1 doesn't affect variability).
d. $V(X) = E[(X – E(X))^2] = E[(Y + 1 – \mu – 1)^2] = E[(Y – \mu)^2] = \sigma^2$.

3.31 **a.** The mean of W will be larger than the mean of Y if $\mu > 0$. If $\mu < 0$, the mean of W will be smaller than μ. If $\mu = 0$, the mean of W will equal μ.
b. $E(W) = E(2Y) = 2E(Y) = 2\mu$.
c. The variance of W will be larger than σ^2, since the spread of values of W has increased.
d. $V(X) = E[(X – E(X))^2] = E[(2Y – 2\mu)^2] = 4E[(Y – \mu)^2] = 4\sigma^2$.

3.32 **a.** The mean of W will be smaller than the mean of Y if $\mu > 0$. If $\mu < 0$, the mean of W will be larger than μ. If $\mu = 0$, the mean of W will equal μ.
b. $E(W) = E(Y/10) = (.1)E(Y) = (.1)\mu$.
c. The variance of W will be smaller than σ^2, since the spread of values of W has decreased.
d. $V(X) = E[(X – E(X))^2] = E[(.1Y – .1\mu)^2] = (.01)E[(Y – \mu)^2] = (.01)\sigma^2$.

3.33 **a.** $E(aY+b) = E(aY) + E(b) = aE(Y) + b = a\mu + b.$
b. $V(aY+b) = E[(aY+b-a\mu-b)^2] = E[(aY-a\mu)^2] = a^2E[(Y-\mu)^2] = a^2\sigma^2.$

3.34 The mean cost is $E(10Y) = 10E(Y) = 10[0(.1) + 1(.5) + 2(.4)] = \$13$. Since $V(Y)$ .41, $V(10Y) = 100V(Y) = 100(.41) = 41$.

3.35 With $B = SS \cup FS$, $P(B) = P(SS) + P(FS) = \frac{2000}{5000}\left(\frac{1999}{4999}\right) + \frac{3000}{5000}\left(\frac{2000}{4999}\right) = 0.4$
$P(B|\text{first trial success}) = \frac{1999}{4999} = 0.3999$, which is not very different from the above.

3.36 **a.** The random variable Y does not have a binomial distribution. The days are not independent.
b. This is not a binomial experiment. The number of trials is not fixed.

3.37 **a.** Not a binomial random variable.
b. Not a binomial random variable.
c. Binomial with $n = 100$, p = proportion of high school students who scored above 1026.
d. Not a binomial random variable (not discrete).
e. Not binomial, since the sample was not selected among all female HS grads.

3.38 Note that Y is binomial with $n = 4$, $p = 1/3 = P$(judge chooses formula B).
a. $p(y) = \binom{4}{y}\left(\frac{1}{3}\right)^y\left(\frac{2}{3}\right)^{4-y}$, $y = 0, 1, 2, 3, 4$.
b. $P(Y \ge 3) = p(3) + p(4) = 8/81 + 1/81 = 9/81 = 1/9$.
c. $E(Y) = 4(1/3) = 4/3$.
d. $V(Y) = 4(1/3)(2/3) = 8/9$

3.39 Let Y = # of components failing in less than 1000 hours. Then, Y is binomial with $n = 4$ and $p = .2$.
a. $P(Y = 2) = \binom{4}{2}.2^2(.8)^2 = 0.1536.$
b. The system will operate if 0, 1, or 2 components fail in less than 1000 hours. So, P(system operates) = .4096 + .4096 + .1536 = .9728.

3.40 Let Y = # that recover from stomach disease. Then, Y is binomial with $n = 20$ and $p = .8$. To find these probabilities, Table 1 in Appendix III will be used.
a. $P(Y \ge 10) = 1 - P(Y \le 9) = 1 - .001 = .999$.
b. $P(14 \le Y \le 18) = P(Y \le 18) - P(Y \le 13) - .931 - .087 = .844$
c. $P(Y \le 16) = .589$.

3.41 Let Y = # of correct answers. Then, Y is binomial with $n = 15$ and $p = .2$. Using Table 1 in Appendix III, $P(Y \ge 10) = 1 - P(Y \le 9) = 1 - 1.000 = 0.000$ (to three decimal places).

3.42 **a.** If one answer can be eliminated on every problem, then, Y is binomial with $n = 15$ and $p = .25$. Then, $P(Y \geq 10) = 1 - P(Y \leq 9) = 1 - 1.000 = 0.000$ (to three decimal places).

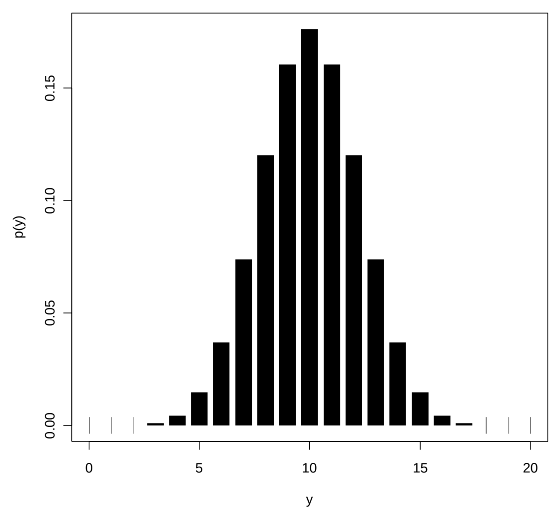
b. If two answers can be (correctly) eliminated on every problem, then, Y is binomial with $n = 15$ and $p = 1/3$. Then, $P(Y \geq 10) = 1 - P(Y \leq 9) = 0.0085$.

3.43 Let Y = # of qualifying subscribers. Then, Y is binomial with $n = 5$ and $p = .7$.
a. $P(Y = 5) = .7^5 = .1681$
b. $P(Y \geq 4) = P(Y = 4) + P(Y = 5) = 5(.7^4)(.3) + .7^5 = .3601 + .1681 = 0.5282$.

3.44 Let Y = # of successful operations. Then Y is binomial with $n = 5$.
a. With $p = .8$, $P(Y = 5) = .8^5 = 0.328$.
b. With $p = .6$, $P(Y = 4) = 5(.6^4)(.4) = 0.259$.
c. With $p = .3$, $P(Y < 2) = P(Y = 1) + P(Y = 0) = 0.528$.

3.45 Note that Y is binomial with $n = 3$ and $p = .8$. The alarm will function if $Y = 1$, 2, or 3. Thus, $P(Y \geq 1) = 1 - P(Y = 0) = 1 - .008 = 0.992$.

3.46 When $p = .5$, the distribution is symmetric. When $p < .5$, the distribution is skewed to the left. When $p > .5$, the distribution is skewed to the right.



3.47 The graph is above.

3.48 **a.** Let Y = # of sets that detect the missile. Then, Y has a binomial distribution with $n = 5$ and $p = .9$. Then, $P(Y = 4) = 5(.9)^4(.1) = 0.32805$ and
$P(Y \geq 1) = 1 - P(Y = 0) = 1 - 5(.9)^4(.1) = 0.32805$.
b. With n radar sets, the probability of at least one diction is $1 - (.1)^n$. If $1 - (.1)^n = .999$, $n = 3$.

3.49 Let Y = # of housewives preferring brand A. Thus, Y is binomial with $n = 15$ and $p = .5$.
a. Using the Appendix, $P(Y \geq 10) = 1 - P(Y \leq 9) = 1 - .849 = 0.151$.
b. $P(10 \text{ or more prefer } A \text{ or } B) = P(6 \leq Y \leq 9) = 0.302$.

3.50 The only way team A can win in exactly 5 games is to win 3 in the first 4 games and then win the 5th game. Let Y = # of games team A wins in the first 4 games. Thus, Y has a binomial distribution with $n = 4$. Thus, the desired probability is given by

$$P(\text{Team A wins in 5 games}) = P(Y = 3)P(\text{Team A wins game 5})$$
$$= \binom{4}{3} p^3(1-p)p = 4p^4(1-p).$$

3.51 **a.** P(at least one 6 in four rolls) = 1 – P(no 6's in four rolls) = 1 – $(5/6)^4$ = 0.51775.

b. Note that in a single toss of two dice, P(double 6) = 1/36. Then:
P(at least one double 6 in twenty–four rolls) = 1 – P(no double 6's in twenty–four rolls) =
= 1 – $(35/36)^{24}$ = 0.4914.

3.52 Let Y = # that are tasters. Then, Y is binomial with $n = 20$ and $p = .7$.
a. $P(Y \geq 17) = 1 - P(Y \leq 16) = 0.107$.
b. $P(Y < 15) = P(Y \leq 14) = 0.584$.

3.53 There is a 25% chance the offspring of the parents will develop the disease. Then, Y = # of offspring that develop the disease is binomial with $n = 3$ and $p = .25$.
a. $P(Y = 3) = (.25)^3 = 0.015625$.
b. $P(Y = 1) = 3(.25)(.75)^2 = 0.421875$
c. Since the pregnancies are mutually independent, the probability is simply 25%.

3.54 **a.** and **b.** follow from simple substitution
c. the classifications of "success" and "failure" are arbitrary.

3.55 $$E\{Y(Y-1)Y-2)\} = \sum_{y=0}^{n} \frac{y(y-1)(y-2)n!}{y!(n-y)!} p^y(1-p)^{n-y} = \sum_{y=3}^{n} \frac{n(n-1)(n-2)(n-3)!}{(y-3)!(n-3-(y-3))!} p^y(1-p)^{n-y}$$

$$= n(n-1)(n-2)p^3 \sum_{z=0}^{n-3} \binom{n-3}{z} p^z(1-p)^{n-3-z} = n(n-1)(n-2)p^3 \ .$$

Equating this to $E(Y^3) - 3E(Y^2) + 2E(Y)$, it is found that
$E(Y^3) = 3n(n-1)p^2 - n(n-1)(n-2)p^3 + np.$

3.56 Using expression for the mean and variance of Y = # of successful explorations, a binomial random variable with $n = 10$ and $p = .1$, $E(Y) = 10(.1) = 1$, and $V(Y) = 10(.1)(.9) = 0.9$.

3.57 If Y = # of successful explorations, then $10 - Y$ is the number of unsuccessful explorations. Hence, the cost C is given by $C = 20{,}000 + 30{,}000Y + 15{,}000(10 - Y)$. Therefore, $E(C) = 20{,}000 + 30{,}000(1) + 15{,}000(10 - 1) = \$185{,}000$.

3.58 If Y is binomial with $n = 4$ and $p = .1$, $E(Y) = .4$ and $V(Y) = .36$. Thus, $E(Y^2) = .36 + (.4)^2 = 0.52$. Therefore, $E(C) = 3(.52) + (.36) + 2 = 3.96$.

3.59 If Y = # of defective motors, then Y is binomial with n = 10 and p = .08. Then, $E(Y)$ = .8. The seller's expected next gain is \$1000 – \$200$E(Y)$ = \$840.

3.60 Let Y = # of fish that survive. Then, Y is binomial with n = 20 and p = .8.
a. $P(Y = 14) = .109$.
b. $P(Y \geq 10) = .999$.
c. $P(Y \leq 16) = .589$.
d. $\mu = 20(.8) = 16$, $\sigma^2 = 20(.8)(.2) = 3.2$.

3.61 Let Y = # with Rh^+ blood. Then, Y is binomial with n = 5 and p = .8
a. $1 - P(Y = 5) = .672$.
b. $P(Y \leq 4) = .672$.
c. We need n for which $P(Y \geq 5) = 1 - P(Y \leq 4) > .9$. The smallest n is 8.

3.62 **a.** Assume independence of the three inspection events.
b. Let Y = # of plane with wing cracks that are detected. Then, Y is binomial with n = 3 and $p = .9(.8)(.5) = .36$. Then, $P(Y \geq 1) = 1 - P(Y = 0) = 0.737856$.

3.63 **a.** Found by pulling in the formula for $p(y)$ and $p(y - 1)$ and simplifying.

b. Note that $P(Y < 3) = P(Y \leq 2) = P(Y = 2) + P(Y = 1) + P(Y = 0)$. Now, $P(Y = 0) = (.96)^{90} = .0254$. Then, $P(Y = 1) = \frac{(90-1+1).04}{1(.96)}(.0254) = .0952$ and $P(Y = 2) = \frac{(90-2+1).04}{2(.96)}(.0952) = .1765$. Thus, $P(Y < 3) = .0254 + .0952 + .1765 = 0.2971$

c. $\dfrac{(n-y+1)}{yq} > 1$ is equivalent to $(n+1)p - yp > yq$ is equivalent to $(n+1)p > y$. The others are similar.

d. Since for $y \leq (n + 1)p$, then $p(y) \geq p(y - 1) > p(y - 2) > \ldots$. Also, for $y \geq (n + 1)p$, then $p(y) \geq p(y + 1) > p(y + 2) > \ldots$. It is clear that p(y) is maximized when y is a close to $(n + 1)p$ as possible.

3.64 To maximize the probability distribution as a function of p, consider taking the natural log (since ln() is a strictly increasing function, it will not change the maximum). By taking the first derivative of $\ln[p(y_0)]$ and setting it equal to 0, the maximum is found to be y_0/n.

3.65 **a.** $E(Y/n) = E(Y)/n = np/n = p$.
b. $V(Y/n) = V(Y)/n^2 = npq/n^2 = pq/n$. This quantity goes to zero as n goes to infinity.

3.66 **a.** $\sum_{y=1}^{\infty} q^{y-1} p = p \sum_{x=0}^{\infty} q^x = p \dfrac{1}{1-q} = 1$ (infinite sum of a geometric series)

b. $\dfrac{q^{y-1}p}{q^{y-2}p} = q$. The event $Y = 1$ has the highest probability for all p, $0 < p < 1$.

3.67 $(.7)^4(.3) = 0.07203$.

3.68 $1/(.30) = 3.33$.

3.69 Y is geometric with $p = 1 – .41 = .59$. Thus, $p(y) = (.41)^{y-1}(.59)$, $y = 1, 2, \ldots$.

3.70 Let Y = # of holes drilled until a productive well is found.
a. $P(Y = 3) = (.8)^2(.2) = .128$
b. $P(Y > 10) = P(\text{first 10 are not productive}) = (.8)^{10} = .107$.

3.71 **a.** $P(Y > a) = \sum_{y=a+1}^{\infty} q^{y-1}p = q^a \sum_{x=1}^{\infty} q^{x-1}p = q^a$.

b. From part a, $P(Y > a+b \mid Y > a) = \dfrac{P(Y > a+b, Y > a)}{P(Y > a)} = \dfrac{P(Y > a+b)}{P(Y > a)} = \dfrac{q^{a+b}}{q^a} = q^b$.

c. The results in the past are not relevant to a future outcome (independent trials).

3.72 Let Y = # of tosses until the first head. $P(Y \geq 12 \mid Y > 10) = P(Y > 11 \mid Y > 10) = 1/2$.

3.73 Let Y = # of accounts audited until the first with substantial errors is found.
a. $P(Y = 3) = .1^2(.9) = .009$.
b. $P(Y \geq 3) = P(Y > 2) = .1^2 = .01$.

3.74 $\mu = 1/.9 = 1.1$, $\sigma = \sqrt{\frac{1-.9}{.9^2}} = .35$

3.75 Let Y = # of one second intervals until the first arrival, so that $p = .1$
a. $P(Y = 3) = (.9)2(.1) = .081$.
b. $P(Y \geq 3) = P(Y > 2) = .9^2 = .81$.

3.76 $P(Y > y_0) = (.7)^{y_0} \geq .1$. Thus, $y_0 \leq \frac{\ln(.1)}{\ln(.7)} = 6.46$, so $y_0 \leq= 6$.

3.77 $P(Y = 1, 3, 5, \ldots) = P(Y = 1) + P(Y + 3) + P(Y = 5) + \ldots = p + q^2p + q^4p + \ldots =$
$p[1 + q^2 + q^4 + \ldots] = p\dfrac{1}{1-q^2}$. (Sum an infinite geometric series in $\left(q^2\right)^k$.)

3.78 **a.** $(.4)^4(.6) = .01536$.
b. $(.4)^4 = .0256$.

3.79 Let Y = # of people questioned before a "yes" answer is given. Then, Y has a geometric distribution with $p = P(\text{yes}) = P(\text{smoker and "yes"}) + P(\text{nonsmoker and "yes"}) = .3(.2) + 0 = .06$. Thus, $p(y) = .06(.94)^{y-1}$. $y = 1, 2, \ldots$.

3.80 Let Y = # of tosses until the first 6 appears, so Y has a geometric distribution. Using the result from Ex. 3.77,

$$P(B \text{ tosses first } 6) = P(Y = 2, 4, 6, \ldots) = 1 - P(Y = 1, 3, 5, \ldots) = 1 - p\frac{1}{1-q^2}.$$

Since $p = 1/6$, $P(B$ tosses first 6$)$ = 5/11. Then,

$$\text{P(Y = 4 | B tosses the first 6)} \; \frac{P(Y=4)}{5/11} = \frac{(5/6)^2(1/6)}{5/11} = 275/1296.$$

3.81 With $p = 1/2$, then $\mu = 1/(1/2) = 2$.

3.82 With $p = .2$, then $\mu = 1/(.2) = 5$. The 5th attempt is the expected first successful well.

3.83 Let Y = # of trials until the correct password is picked. Then, Y has a geometric distribution with $p = 1/n$. $P(Y = 6) = \frac{1}{n}\left(\frac{n-1}{n}\right)^5$.

3.84 $E(Y) = n$, $V(Y) = (1-\frac{1}{n})n^2 = n(n-1)$.

3.85 Note that $\frac{d^2}{dq^2}q^y = y(y-1)q^{y-2}$. Thus, $\frac{d^2}{dq^2}\sum_{y=2}^{\infty} q^y = \sum_{y=2}^{\infty} y(y-1)q^{y-2}$. Thus,

$$E[Y(Y-1)] = \sum_{y=1}^{\infty} y(y-1)q^{y-1} = pq\sum_{y=1}^{\infty} y(y-1)q^{y-2} = pq\frac{d^2}{dq^2}\sum_{y=2}^{\infty} q^y = pq\frac{d^2}{dq^2}\left\{\frac{1}{1-q} - 1 - q\right\} =$$

$\frac{2pq}{(1-q)^3} = \frac{2q}{p^2}$. Use this with $V(Y) = E[Y(Y-1)] + E(Y) - [E(Y)]^2$.

3.86 $P(Y = y_0) = q^{y_0-1}p$. Like Ex. 3.64, maximize this probability by first taking the natural log.

3.87 $E(1/Y) = \sum_{y=1}^{\infty}\frac{1}{y}(1-p)^{y-1}p = \frac{p}{1-p}\sum_{y=1}^{\infty}\frac{(1-p)^y}{y} = -\frac{p\ln(p)}{1-p}$.

3.88 $P(Y^* = y) = P(Y = y+1) = q^{y+1-1}p = q^y p$, $y = 0, 1, 2, \ldots$.

3.89 $E(Y^*) = E(Y) - 1 = \frac{1}{p} - 1$. $V(Y^*) = V(Y - 1) = V(Y)$.

3.90 Let Y = # of employees tested until three positives are found. Then, Y is negative binomial with $r = 3$ and $p = .4$. $P(Y = 10) = \binom{9}{2}.4^3(.6)^7 = .06$.

3.91 The total cost is given by $20Y$. So, $E(20Y) = 20E(Y) = 20\frac{3}{.4}$ = \$50. Similarly, $V(20Y) = 400V(Y) = 4500$.

3.92 Let Y = # of trials until this first non–defective engine is found. Then, Y is geometric with $p = .9$. $P(Y = 2) = .9(.1) = .09$.

3.93 From Ex. 3.92:

a. $P(Y = 5) = \binom{4}{2}(.9)^3(.1)^2 = .04374$.

b. $P(Y \leq 5) = P(Y = 3) + P(Y = 4) + P(Y = 5) = .729 + .2187 + .04374 = .99144$.

3.94 **a.** $\mu = 1/(.9) = 1.11$, $\sigma^2 = (.1)/(.9)2 = .1234$.
b. $\mu = 3/(.9) = 3.33$, $\sigma^2 = 3(.1)/(.9)2 = .3704$.

3.95 From Ex. 3.92 (and the memory–less property of the geometric distribution),
$P(Y \geq 4 \mid Y > 2) = P(Y > 3 \mid Y > 2) = P(Y > 1) = 1 - P(Y = 0) = .1$.

3.96 **a.** Let Y = # of attempts until you complete your call. Thus, Y is geometric with $p = .4$. Thus, $P(Y = 1) = .4$, $P(Y = 2) = (.6).4 = .24$, $P(Y = 3) = (.6)^2.4 = .144$.
b. Let Y = # of attempts until both calls are completed. Thus, Y is negative binomial with $r = 2$ and $p = .4$. Thus, $P(Y = 4) = 3(.4)^2(.6)^2 = .1728$.

3.97 **a.** Geometric probability calculation: $(.8)^2(.2) = .128$.

b. Negative binomial probability calculation: $\binom{6}{2}(.2)^3(.8)^4 = .049$.

c. The trials are independent and the probability of success is the same from trial to trial.

d. $\mu = 3/.2 = 15$, $\sigma^2 = 3(.8)/(.04) = 60$.

3.98 **a.** $\dfrac{p(y)}{p(y-1)} = \dfrac{\frac{(y-1)!}{(r-1)!(y-r)!}p^r q^{y-r}}{\frac{(y-2)!}{(r-1)!(y-1-r)!}p^r q^{y-1-r}} = \dfrac{y-1}{y-r}q$

b. If $\dfrac{y-1}{y-r}q > 1$, then $yq - q > y - r$ or equivalently $\dfrac{r-q}{1-q} > y$. The 2nd result is similar.

c. If $r = 7$, $p = .5 = q$, then $\dfrac{r-q}{1-q} = \dfrac{7-.5}{1-.5} = 13 > y$.

3.99 Define a random variable $X = y$ trials before the before the first success, $y = r - 1, r, r + 1, \ldots$. Then, $X = Y - 1$, where Y has the negative binomial distribution with parameters r and p. Thus, $p(x) = \frac{y!}{(r-1)!(y-r+1)!}p^r q^{y+1-r}$, $y = r - 1, r, r + 1, \ldots$.

3.100 **a.** $P(Y^* = y) = P(Y = y + r) = \binom{y+r-1}{r-1}p^r q^{y+r-r} = \binom{y+r-1}{r-1}p^r q^y$, $y = 0, 1, 2, \ldots$.

b. $E(Y^*) = E(Y) - r = r/p - r = r/q$, $V(Y^*) = V(Y - r) = V(Y)$.

3.101 **a.** Note that $P(Y = 11) = \binom{10}{4} p^5 (1-p)^6$. Like Ex. 3.64 and 3.86, maximize this probability by first taking the natural log. The maximum is 5/11.
b. In general, the maximum is r/y_0.

3.102 Let Y = # of green marbles chosen in three draws. Then. $P(Y = 3) = \binom{5}{3} \Big/ \binom{10}{3}$ = 1/12.

3.103 Use the hypergeometric probability distribution with $N = 10$, $r = 4$, $n = 5$. $P(Y = 0) = \frac{1}{42}$.

3.104 Define the events: A: 1st four selected packets contain cocaine
B: 2nd two selected packets do not contain cocaine

Then, the desired probability is $P(A \cap B) = P(B|A)P(A)$. So,
$P(A) = \binom{15}{4} \Big/ \binom{20}{4} = .2817$ and $P(B|A) = \binom{5}{2} \Big/ \binom{16}{2} = .0833$. Thus,
$P(A \cap B) = .2817(.0833) = 0.0235$.

3.105 **a.** The random variable Y follows a hypergeometric distribution. The probability of being chosen on a trial is dependent on the outcome of previous trials.

b. $P(Y \geq 2) = P(Y = 2) + P(Y = 3) = \dfrac{\binom{5}{2}\binom{3}{1}}{\binom{8}{3}} + \dfrac{\binom{5}{3}}{\binom{8}{3}} = .5357 + .1786 = 0.7143$.

c. $\mu = 3(5/8) = 1.875$, $\sigma^2 = 3(5/8)(3/8)(5/7) = .5022$, so $\sigma = .7087$.

3.106 Using the results from Ex.103, $E(50Y) = 50E(Y) = 50[5\left(\frac{4}{10}\right)] = \$100$. Furthermore,
$V(50Y) = 2500V(Y) = 2500[5\,\frac{4}{10}\left(\frac{6}{10}\right)\left(\frac{5}{9}\right)] = 1666.67$.

3.107 The random variable Y follows a hypergeometric distribution with $N = 6$, $n = 2$, and $r = 4$.

3.108 Use the fact that P(at least one is defective) = 1 – P(none are defective). Then, we require P(none are defective) ≤ .2. If $n = 8$,

$$P(\text{none are defective}) = \left(\tfrac{17}{20}\right)\left(\tfrac{16}{19}\right)\left(\tfrac{15}{18}\right)\left(\tfrac{14}{17}\right)\left(\tfrac{13}{16}\right)\left(\tfrac{12}{15}\right)\left(\tfrac{11}{14}\right)\left(\tfrac{10}{13}\right) = 0.193.$$

3.109 Let Y = # of treated seeds selected.

a. $P(Y = 4) = \dfrac{\binom{5}{4}\binom{5}{0}}{\binom{10}{4}} = .0238$

b. $P(Y \leq 3) = 1 - P(Y = 4) = 1 - \dfrac{\binom{5}{4}\binom{5}{0}}{\binom{10}{4}} = 1 - .0238 = .9762$.

c. same answer as part (b) above.

3.110 **a.** $P(Y = 1) = \dfrac{\binom{4}{2}\binom{2}{1}}{\binom{6}{3}} = .6.$

b. $P(Y \geq 1) = p(1) + p(2) = \dfrac{\binom{4}{2}\binom{2}{1}}{\binom{6}{3}} + \dfrac{\binom{4}{1}\binom{2}{2}}{\binom{6}{3}} = .8$

c. $P(Y \leq 1) = p(0) + p(1) = .8.$

3.111 **a.** The probability function for Y is $p(y) = \dfrac{\binom{2}{y}\binom{8}{3-y}}{\binom{10}{3}}$, $y = 0, 1, 2$. In tabular form, this is

y	0	1	2
$p(y)$	14/30	14/30	2/30

b. The probability function for Y is $p(y) = \dfrac{\binom{4}{y}\binom{6}{3-y}}{\binom{10}{3}}$, $y = 0, 1, 2, 3$. In tabular form, this is

y	0	1	2	3
$p(y)$	5/30	15/30	9/30	1/30

3.112 Let Y = # of malfunctioning copiers selected. Then, Y is hypergeometric with probability function

$$p(y) = \frac{\binom{3}{y}\binom{5}{4-y}}{\binom{8}{4}},\ y = 0, 1, 2, 3.$$

a. $P(Y = 0) = p(0) = 1/14.$
b. $P(Y \geq 1) = 1 - P(Y = 0) = 13/14.$

3.113 The probability of an event as rare or rarer than one observed can be calculated according to the hypergeometric distribution, Let Y = # of black members. Then, Y is hypergeometric and $P(Y \leq 1) = \dfrac{\binom{8}{3}\binom{12}{5}}{\binom{20}{6}} + \dfrac{\binom{8}{0}\binom{12}{6}}{\binom{20}{6}} = .187$. This is nearly 20%, so it is not unlikely.

3.114 $\mu = 6(8)/20 = 2.4$, σ^2 $6(8/20)(12/20)(14/19) = 1.061$.

3.115 The probability distribution for Y is given by

y	0	1	2
$p(y)$	1/5	3/5	1/5

3.116 (Answers vary, but with n =100, the relative frequencies should be close to the probabilities in the table above.)

3.117 Let Y = # of improperly drilled gearboxes. Then, Y is hypergeometric with N = 20, n = 5, and r = 2.

a. $P(Y = 0) = .553$

b. The random variable T, the total time, is given by $T = 10Y + (5 - Y) = 9Y + 5$. Thus, $E(T) = 9E(Y) + 5 = 9[5(2/20)] + 5 = 9.5$.
$V(T) = 81V(T) = 81(.355) = 28.755$, $\sigma = 5.362$.

3.118 Let Y = # of aces in the hand. Then. $P(Y = 4 \mid Y \geq 3) = \dfrac{P(Y=4)}{P(Y=3)+P(Y=4)}$. Note that Y is a hypergeometric random variable. So, $P(Y=3) = \dfrac{\binom{4}{3}\binom{48}{2}}{\binom{52}{5}}$ = .001736 and

$P(Y=4) = \dfrac{\binom{4}{4}\binom{48}{1}}{\binom{52}{5}}$ = .00001847. Thus, $P(Y = 4 \mid Y \geq 3) = .0105$.

3.119 Let the event A = 2nd king is dealt on 5th card. The four possible outcomes for this event are {$KNNNK$, $NKNNK$, $NNKNK$, $NNNKK$}, where K denotes a king and N denotes a non–king. Each of these outcomes has probability: $\left(\frac{4}{52}\right)\left(\frac{48}{51}\right)\left(\frac{47}{50}\right)\left(\frac{46}{49}\right)\left(\frac{3}{48}\right)$. Then, the desired probability is $P(A) = 4\left(\frac{4}{52}\right)\left(\frac{48}{51}\right)\left(\frac{47}{50}\right)\left(\frac{46}{49}\right)\left(\frac{3}{48}\right) = .016$.

3.120 There are N animals in this population. After taking a sample of k animals, making and releasing them, there are $N - k$ unmarked animals. We then choose a second sample of size 3 from the N animals. There are $\binom{N}{3}$ ways of choosing this second sample and there are $\binom{N-k}{2}\binom{k}{1}$ ways of finding exactly one of the originally marked animals. For $k = 4$, the probability of finding just one marked animal is

$$P(Y = 1) = \frac{\binom{N-4}{2}\binom{4}{1}}{\binom{N}{3}} = \frac{12(N-4)(N-5)}{N(N-1)(N-2)}.$$

Calculating this for various values of N, we find that the probability is largest for N = 11 or N = 12 (the same probability is found: .503).

3.121 **a.** $P(Y = 4) = \frac{2^4}{4!}e^{-2} = .090$.

b. $P(Y \geq 4) = 1 - P(Y \leq 3) = 1 - .857 = .143$ (using Table 3, Appendix III).

c. $P(Y < 4) = P(Y \leq 3) = .857$.

d. $P(Y \geq 4 \mid Y \geq 2) = \dfrac{P(Y \geq 4)}{P(Y \geq 2)} = .143/.594 = .241$

3.122 Let Y = # of customers that arrive during the hour. Then, Y is Poisson with $\lambda = 7$.
a. $P(Y \le 3) = .0818$.
b. $P(Y \ge 2) = .9927$.
c. $P(Y = 5) = .1277$

3.123 If $p(0) = p(1)$, $e^{-\lambda} = \lambda e^{-\lambda}$. Thus, $\lambda = 1$. Therefore, $p(2) = \frac{1^2}{2!}e^{-1} = .1839$.

3.124 Using Table 3 in Appendix III, we find that if Y is Poisson with $\lambda = 6.6$, $P(Y \le 2) = .04$. Using this value of λ, $P(Y > 5) = 1 - P(Y \le 5) = 1 - .355 = .645$.

3.125 Let S = total service time = $10Y$. From Ex. 3.122, Y is Poisson with $\lambda = 7$. Therefore, $E(S) = 10E(Y) = 7$ and $V(S) = 100V(Y) = 700$. Also, $P(S > 150) = P(Y > 15) = 1 - \text{P}(Y \le 15) = 1 - .998 = .002$, and unlikely event.

3.126 **a.** Let Y = # of customers that arrive in a given two–hour time. Then, Y has a Poisson distribution with $\lambda = 2(7) = 14$ and $P(Y = 2) = \frac{14^2}{2!}e^{-14}$.

b. The same answer as in part a. is found.

3.127 Let Y = # of typing errors per page. Then, Y is Poisson with $\lambda = 4$ and $P(Y \le 4) = .6288$.

3.128 Note that over a one–minute period, Y = # of cars that arrive at the toll booth is Poisson with $\lambda = 80/60 = 4/3$. Then, $P(Y \ge 1) = 1 - P(Y = 0) = 1 - e^{-4/3} = .7364$.

3.129 Following the above exercise, suppose the phone call is of length t, where t is in minutes. Then, Y = # of cars that arrive at the toll booth is Poisson with $\lambda = 4t/3$. Then, we must find the value of t such that

$$P(Y = 0) = 1 - e^{-4t/3} \ge .4.$$

Therefore, $t \le -\frac{3}{4}\ln(.6) = .383$ minutes, or about $.383(60) = 23$ seconds.

3.130 Define: Y_1 = # of cars through entrance I, Y_2 = # of cars through entrance II. Thus, Y_1 is Poisson with $\lambda = 3$ and Y_2 is Poisson with $\lambda = 4$.

Then, $P(\text{three cars arrive}) = P(Y_1 = 0, Y_2 = 3) + P(Y_1 = 1, Y_2 = 2) + P(Y_1 = 2, Y_2 = 1) + {} + P(Y_1 = 3, Y_2 = 0)$.

By independence, $P(\text{three cars arrive}) = P(Y_1 = 0)P(Y_2 = 3) + P(Y_1 = 1)P(Y_2 = 2) + P(Y_1 = 2)P(Y_2 = 1) + P(Y_1 = 3)P(Y_2 = 0)$.
Using Poisson probabilities, this is equal to 0.0521

3.131 Let the random variable Y = # of knots in the wood. Then, Y has a Poisson distribution with $\lambda = 1.5$ and $P(Y \le 1) = .5578$.

3.132 Let the random variable Y = # of cars entering the tunnel in a two–minute period. Then, Y has a Poisson distribution with $\lambda = 1$ and $P(Y > 3) = 1 - P(Y \le 3) = 0.01899$.

3.133 Let X = # of two–minute intervals with more than three cars. Therefore, X is binomial with n = 10 and p = .01899 and $P(X \geq 1) = 1 - P(X = 0) = 1 - (1-.01899)^{10} = .1745$.

3.134 The probabilities are similar, even with a fairly small n.

y	$p(y)$, exact binomial	$p(y)$, Poisson approximation
0	.358	.368
1	.378	.368
2	.189	.184
3	.059	.061
4	.013	.015

3.135 Using the Poisson approximation, $\lambda \approx np = 100(.03) = 3$, so $P(Y \geq 1) = 1 - P(Y = 0) =$ .9524.

3.136 Let Y = # of *E. coli* cases observed this year. Then, Y has an approximate Poisson distribution with $\lambda \approx 2.4$.
a. $P(Y \geq 5) = 1 - P(Y \leq 4) = 1 - .904 = .096$.
b. $P(Y > 5) = 1 - P(Y \leq 5) = 1 - .964 = .036$. Since there is a small probability associated with this event, the rate probably has charged.

3.137 Using the Poisson approximation to the binomial with $\lambda \approx np = 30(.2) = 6$. Then, $P(Y \leq 3) = .1512$.

3.138 $E[Y(Y-1)] = \sum_{y=0}^{\infty} \frac{y(y-1)\lambda^y e^{-\lambda}}{y!} = \lambda^2 \sum_{y=0}^{\infty} \frac{y(y-1)\lambda^{y-2} e^{-\lambda}}{y!}$. Using the substitution $z = y - 2$, it is found that $E[Y(Y-1)] = \lambda^2$. Use this with $V(Y) = E[Y(Y-1)] + E(Y) - [E(Y)]^2 = \lambda$.

3.139 Note that if Y is Poisson with $\lambda = 2$, $E(Y) = 2$ and $E(Y^2) = V(Y) + [E(Y)]^2 = 2 + 4 = 6$. So, $E(X) = 50 - 2E(Y) - E(Y^2) = 50 - 2(2) - 6 = 40$.

3.140 Since Y is Poisson with $\lambda = 2$, $E(C) = E\left[100\left(\tfrac{1}{2}\right)^Y\right] = \sum_{y=0}^{\infty} \frac{100\left(\frac{2}{2}\right)^y e^{-2}}{y!} = 100e^{-1} \sum_{y=1}^{\infty} \frac{1^y e^{-1}}{y!} = 100e^{-1}$.

3.141 Similar to Ex. 3.139: $E(R) = E(1600 - 50Y^2) = 1600 - 50(6) = \$1300$.

3.142 **a.** $\frac{p(y)}{p(y-1)} = \frac{\frac{\lambda^y e^{-\lambda}}{y!}}{\frac{\lambda^{y-1} e^{-\lambda}}{(y-1)!}} = \frac{\lambda}{y}, y = 1, 2, \ldots$
b. Note that if $\lambda > y$, $p(y) > p(y-1)$. If $\lambda > y$, $p(y) > p(y-1)$. If $\lambda = y$ for some integer y, $p(y) = p(y-1)$.
c. Note that for λ a non–integer, part b. implies that for $y - 1 < y < \lambda$,
$$p(y-1) < p(y) > p(y+1).$$
Hence, $p(y)$ is maximized for y = largest integer less than λ. If λ is an integer, then $p(y)$ is maximized at both values $\lambda - 1$ and λ.

3.143 Since λ is a non–integer, $p(y)$ is maximized at $y = 5$.

3.144 Observe that with $\lambda = 6$, $p(5) = \frac{6^5 e^{-6}}{5!} = .1606$, $p(6) = \frac{6^6 e^{-6}}{6!} = .1606$.

3.145 Using the binomial theorem, $m(t) = E(e^{tY}) = \sum_{y=0}^{n} \binom{n}{y} (pe^t)^y q^{n-y} = (pe^t + q)^n$.

3.146 $\frac{d}{dt} m(t) = n(pe^t + q)^{n-1} pe^t$. At $t = 0$, this is $np = E(Y)$.

$\frac{d^2}{dt^2} m(t) = n(n-1)(pe^t + q)^{n-1}(pe^t)^2 + n(pe^t + q)^{n-1} pe^t$. At $t = 0$, this is $np^2(n-1) + np$.
Thus, $V(Y) = np^2(n-1) + np - (np)2 = np(1-p)$.

3.147 The moment–generating function is $m(t) = E(e^{tY}) = \sum_{y=1}^{n} pe^{ty} q^{y-1} = pe^t \sum_{y=0}^{\infty} (qe^t)^y = \dfrac{pe^t}{1 - qe^t}$.

3.148 $\frac{d}{dt} m(t) = \dfrac{pe^t}{(1 - qe^t)^2}$. At $t = 0$, this is $1/p = E(Y)$.

$\frac{d^2}{dt^2} m(t) = \dfrac{(1 - qe^t)^2 pe^t - 2pe^t(1 - qe^t)(-qe^t)}{(1 - qe^t)^4}$. At $t = 0$, this is $(1+q)/p^2$.
Thus, $V(Y) = (1+q)/p^2 - (1/p)^2 = q/p^2$.

3.149 This is the moment–generating function for the binomial with $n = 3$ and $p = .6$.

3.150 This is the moment–generating function for the geometric with $p = .3$.

3.151 This is the moment–generating function for the binomial with $n = 10$ and $p = .7$, so $P(Y \le 5) = .1503$.

3.152 This is the moment–generating function for the Poisson with $\lambda = 6$. So, $\mu = 6$ and $\sigma = \sqrt{6} \approx 2.45$. So, $P(|Y - \mu| \le 2\sigma) = P(\mu - 2\sigma \le Y \le \mu + 2\sigma) = P(1.1 \le Y \le 10.9) = P(2 \le Y \le 10) = .940$.

3.153 **a.** Binomial with $n = 5$, $p = .1$
b. If $m(t)$ is multiplied top and bottom by ½, this is a geometric mgf with $p = ½$.
c. Poisson with $\lambda = 2$.

3.154 **a.** Binomial mean and variance: $\mu = 1.667$, $\sigma^2 = 1.111$.
b. Geometric mean and variance: $\mu = 2$, $\sigma^2 = 2$.
c. Poisson mean and variance: $\mu = 2$, $\sigma^2 = 2$.

3.155 Differentiate to find the necessary moments:

a. $E(Y) = 7/3$.

b. $V(Y) = E(Y^2) – [E(Y)]^2 = 6 – (7/3)^2 = 5/9$.

c. Since $m(t) = E(e^{tY})$. Y can only take on values 1, 2, and 3 with probabilities 1/6, 2/6, and 3/6.

3.156 **a.** $m(0) = E(e^{0Y}) = E(1) = 1$.

b. $m_W(t) = E(e^{tW}) = E(e^{t3Y}) = E(e^{(3t)Y}) = m(3t)$.

c. $m_X(t) = E(e^{tX}) = E(e^{t(Y-2)}) = E(e^{-2t}e^{tY}) = e^{-2t}m(t)$.

3.157 **a.** From part b. in Ex. 3.156, the results follow from differentiating to find the necessary moments.

b. From part c. in Ex. 3.156, the results follow from differentiating to find the necessary moments.

3.158 The mgf for W is $m_W(t) = E(e^{tW}) = E(e^{t(aY+b)}) = E(e^{bt}e^{(at)Y}) = e^{bt}m(at)$.

3.159 From Ex. 3.158, the results follow from differentiating the mgf of W to find the necessary moments.

3.160 **a.** $E(Y^*) = E(n – Y) = n – E(Y) = n – np = n(1 – p) = nq$. $V(Y^*) = V(n – Y) = V(Y) = npq$.

b. $m_{Y^*}(t) = E(e^{tY^*}) = E(e^{t(n-Y)}) = E(e^{nt}e^{(-t)Y}) = e^{nt}m(-t) = (pe^t + q)^n$.

c. Based on the moment–generating function, Y* has a binomial distribution.

d. The random variable Y* = # of failures.

e. The classification of "success" and "failure" in the Bernoulli trial is arbitrary.

3.161 $m_{Y^*}(t) = E(e^{tY^*}) = E(e^{t(Y-1)}) = E(e^{-t}e^{tY}) = e^{-t}m(t) = \frac{p}{1-qe^t}$.

3.162 Note that $r^{(1)}(t) = \frac{m^{(1)}(t)}{m(t)}$, $r^{(2)}(t) = \frac{m^{(2)}(t)m(t) - [m^{(1)}(t)]^2}{(m(t))^2}$. Then, $r^{(1)}(0) = \frac{m^{(1)}(0)}{m(0)} = \frac{E(Y)}{1} = \mu$.

$r^{(2)}(0) = \frac{m^{(2)}(0)m(0) - [m^{(1)}(0)]^2}{(m(0))^2} = \frac{E(Y^2) - [E(Y)]^2}{1} = \sigma^2$.

3.163 Note that $r(t) = 5(e^t – 1)$. Then, $r^{(1)}(t) = 5e^t$ and $r^{(2)}(t) = 5e^t$. So, $r^{(1)}(0) = 5 = \mu = \lambda$ and $r^{(2)}(t) = 5e^t = \sigma^2 = \lambda$.

3.164 For the binomial, $P(t) = E(t^Y) = \sum_{y=0}^{n} \binom{n}{y} (pt)^y q^{n-y} = (q + pt)^n$. Differentiating with respect to t, $\frac{d}{dt} P(t)\Big|_{t=1} = np(q + pt)^{n-1}\Big|_{t=1} = np$.

3.165 For the Poisson, $P(t) = E(t^Y) = \sum_{y=0}^{\infty} \frac{\lambda^y e^{-\lambda} t^y}{y!} = \frac{e^{-\lambda}}{e^{-\lambda t}} \sum_{y=0}^{\infty} \frac{(t\lambda)^y e^{-\lambda t}}{y!} = e^{\lambda(t-1)}$. Differentiating with respect to t, $E(Y) = \frac{d}{dt} P(t)\Big|_{t=1} = \lambda e^{\lambda(t-1)}\Big|_{t=1} = \lambda$ and $\frac{d^2}{dt^2} P(t)\Big|_{t=1} = \lambda^2 e^{\lambda(t-1)}\Big|_{t=1} = \lambda^2$ = $E[Y(Y-1)] = E(Y^2) - E(Y)$. Thus, $V(Y) = \lambda$.

3.166 $E[Y(Y-1)(Y-2)] = \frac{d^3}{dt^3} P(t)\Big|_{t=1} = \lambda^3 e^{\lambda(t-1)}\Big|_{t=1} = \lambda^3 = E(Y^3) - 3E(Y^2) + 2E(Y)$. Therefore, $E(Y^3) = \lambda^3 + 3(\lambda^2 + \lambda) - 2\lambda = \lambda^3 + 3\lambda^2 + \lambda$.

3.167 **a.** The value 6 lies (11–6)/3 = 5/3 standard deviations below the mean. Similarly, the value 16 lies (16–11)/3 = 5/3 standard deviations above the mean. By Tchebysheff's theorem, at least $1 - 1/(5/3)^2 = 64\%$ of the distribution lies in the interval 6 to 16.

b. By Tchebysheff's theorem, $.09 = 1/k^2$, so $k = 10/3$. Since $\sigma = 3$, $k\sigma = (10/3)^3 = 10 = C$.

3.168 Note that Y has a binomial distribution with $n = 100$ and $p = 1/5 = .2$
- **a.** $E(Y) = 100(.2) = 20$.
- **b.** $V(Y) = 100(.2)(.8) = 16$, so $\sigma = 4$.
- **c.** The intervals are $20 \pm 2(4)$ or (12, 28), $20 + 3(4)$ or (8, 32).
- **d.** By Tchebysheff's theorem, $1 - 1/3^2$ or approximately 89% of the time the number of correct answers will lie in the interval (8, 32). Since a passing score of 50 is far from this range, receiving a passing score is very unlikely.

3.169 **a.** $E(Y) = -1(1/18) + 0(16/18) + 1(1/18) = 0$. $E(Y^2) = 1(1/18) + 0(16/18) + 1(1/18) = 2/18 = 1/9$. Thus, $V(Y) = 1/9$ and $\sigma = 1/3$.

b. $P(|Y - 0| \geq 1) = P(Y = -1) + P(Y = 1) = 1/18 + 1/18 = 2/18 = 1/9$. According to Tchebysheff's theorem, an upper bound for this probability is $1/3^2 = 1/9$.

c. Example: let X have probability distribution $p(-1) = 1/8$, $p(0) = 6/8$, $p(1) = 1/8$. Then, $E(X) = 0$ and $V(X) = 1/4$.

d. For a specified k, assign probabilities to the points –1, 0, and 1 as $p(-1) = p(1) = \frac{1}{2k^2}$ and $p(0) = 1 - \frac{1}{k}$.

3.170 Similar to Ex. 3.167: the interval (.48, 52) represents two standard deviations about the mean. Thus, the lower bound for this interval is 1 – ¼ = ¾. The expected number of coins is 400(¾) = 300.

3.171 Using Tchebysheff's theorem, $5/9 = 1 - 1/k^2$, so $k = 3/2$. The interval is $100 \pm (3/2)10$, or 85 to 115.

3.172 From Ex. 3.115, $E(Y) = 1$ and $V(Y) = .4$. Thus, $\sigma = .63$. The interval of interest is $1 \pm 2(.63)$, or $(-.26, 2.26)$. Since Y can only take on values 0, 1, or 2, 100% of the values will lie in the interval. According to Tchebysheff's theorem, the lower bound for this probability is 75%.

3.173 **a.** The binomial probabilities are $p(0) = 1/8$, $p(1) = 3/8$, $p(2) = 3/8$, $p(3) = 1/8$.

b. The graph represents a <u>symmetric distribution</u>.

c. $E(Y) = 3(1/2) = 1.5$, $V(Y) = 3(1/2)(1/2) = .75$. Thus, $\sigma = .866$.

d. For *one* standard deviation about the mean: $1.5 \pm .866$ or $(.634, 2.366)$
This traps the values 1 and 2, which represents 7/8 or 87.5% of the probability. This is consistent with the empirical rule.
For *two* standard deviations about the mean: $1.5 \pm 2(.866)$ or $(-.232, 3.232)$
This traps the values 0, 1, and 2, which represents 100% of the probability. This is consistent with both the empirical rule and Tchebysheff's theorem.

3.174 **a.** (Similar to Ex. 3.173) the binomial probabilities are $p(0) = .729$, $p(1) = .243$, $p(2) = .027$, $p(3) = .001$.

b. The graph represents a <u>skewed distribution</u>.

c. $E(Y) = 3(.1) = .3$, $V(Y) = 3(.1)(.9) = .27$. Thus, $\sigma = .520$.

d. For *one* standard deviation about the mean: $.3 \pm .520$ or $(-.220, .820)$
This traps the value 1, which represents 24.3% of the probability. This is not consistent with the empirical rule.
For *two* standard deviations about the mean: $.3 \pm 2(.520)$ or $(-.740, 1.34)$
This traps the values 0 and 1, which represents 97.2% of the probability. This is consistent with both the empirical rule and Tchebysheff's theorem.

3.175 **a.** The expected value is $120(.32) = 38.4$
b. The standard deviation is $\sqrt{120(.32)(.68)} = 5.11$.
c. It is quite likely, since 40 is close to the mean 38.4 (less than .32 standard deviations away).

3.176 Let Y represent the number of students in the sample who favor banning clothes that display gang symbols. If the teenagers are actually equally split, then $E(Y) = 549(.5) = 274.5$ and $V(Y) = 549(.5)(.5) = 137.25$. Now. Y/549 represents the proportion in the sample who favor banning clothes that display gang symbols, so $E(Y/549) = .5$ and $V(Y/549) = .5(.5)/549 = .000455$. Then, by Tchebysheff's theorem,

$$P(Y/549 \geq .85) \leq P(|Y/549 - .5| \geq .35) \leq 1/k^2,$$

where k is given by $k\sigma = .35$. From above, $\sigma = .02134$ so $k = 16.4$ and $1/(16.4)^2 = .0037$. This is a very unlikely result. It is also unlikely using the empirical rule. We assumed that the sample was selected randomly from the population.

3.177 For $C = 50 + 3Y$, $E(C) = 50 + 3(10) = \$80$ and $V(C) = 9(10) = 90$, so that $\sigma = 9.487$. Using Tchebysheff's theorem with $k = 2$, we have $P(|Y - 80| < 2(9.487)) \geq .75$, so that the required interval is $(80 - 2(9.487), 80 + 2(9.487))$ or $(61.03, 98.97)$.

3.178 Using the binomial, $E(Y) = 1000(.1) = 100$ and $V(Y) = 1000(.1)(.9) = 90$. Using the result that at least 75% of the values will fall within two standard deviation of the mean, the interval can be constructed as $100 \pm 2\sqrt{90}$, or $(81, 119)$.

3.179 Using Tchebysheff's theorem, observe that

$$P(Y \geq \mu + k\sigma) = P(Y - \mu \geq k\sigma) \leq P(|Y - \mu| \geq k\sigma) \leq \tfrac{1}{k^2}.$$

Therefore, to find $P(Y \geq 350) \leq 1/k^2$, we solve $150 + k(67.081) = 350$, so $k = 2.98$. Thus, $P(Y \geq 350) \leq 1/(2.98)^2 = .1126$, which is not highly unlikely.

3.180 Number of combinations = 26(26)(10)(10)(10)(10) = 6,760,000. Thus, $E(\text{winnings}) = 100{,}000(1/6{,}760{,}000) + 50{,}000(2/6{,}760{,}000) + 1000(10/6{,}760{,}000) = \$.031$, which is much less than the price of the stamp.

3.181 Note that $P(\text{acceptance}) = P(\text{observe no defectives}) = \binom{5}{0}p^0q^5$. Thus:

p = Fraction defective	P(acceptance)
0	1
.10	.5905
.30	.1681
.50	.0312
1.0	0

3.182 OC curves can be constructed using points given in the tables below.

a. Similar to Ex. 3.181: $P(\text{acceptance}) = \binom{10}{0}p^0q^{10}$. Thus,

p	0	.05	.10	.30	.50	1
P(acceptance)	1	.599	.349	.028	.001	0

b. Here, $P(\text{acceptance}) = \binom{10}{0}p^0q^{10} + \binom{10}{1}p^1q^9$. Thus,

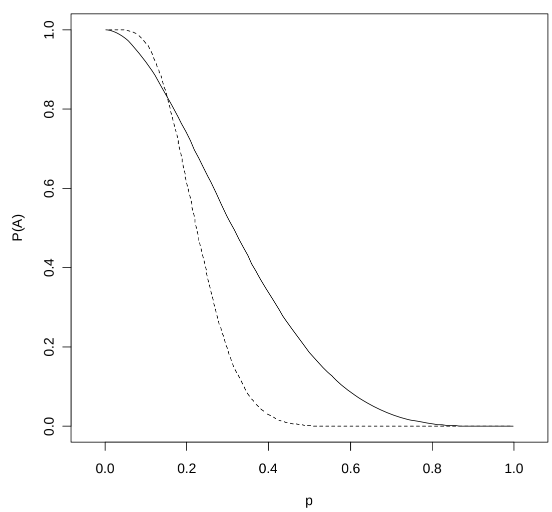
p	0	.05	.10	.30	.50	1
P(acceptance)	1	.914	.736	.149	.01	0

c. Here, $P(\text{acceptance}) = \binom{10}{0}p^0q^{10} + \binom{10}{1}p^1q^9 + \binom{10}{2}p^2q^8$. Thus,

p	0	.05	.10	.30	.50	1
P(acceptance)	1	.988	.930	.383	.055	0

3.183 Graph the two OC curves with $n = 5$ and $a = 1$ in the first case and $n = 25$ and $a = 5$ in the second case.

a. By graphing the OC curves, it is seen that if the defectives fraction ranges from $p = 0$ to $p = .10$, the seller would want the probability of accepting in this interval to be a high as possible. So, he would choose the second plan.
b. If the buyer wishes to be protected against accepting lots with a defective fraction greater than .3, he would want the probability of acceptance (when p > .3) to be as small as possible. Thus, he would also choose the second plan.



The above graph illustrates the two OC curves. The solid lie represents the first case and the dashed line represents the second case.

3.184 Let Y = # in the sample who favor garbage collect by contract to a private company. Then, Y is binomial with n = 25.
a. If $p = .80$, $P(Y \geq 22) = 1 - P(Y \leq 21) = 1 - .766 = .234$,
b. If $p = .80$, $P(Y = 22) = .1358$.
c. There is not strong evidence to show that the commissioner is incorrect.

3.185 Let Y = # of students who choose the numbers 4, 5, or 6. Then, Y is binomial with $n = 20$ and $p = 3/10$.
a. $P(Y \geq 8) = 1 - P(Y \leq 7) = 1 - .7723 = .2277$.
b. Given the result in part a, it is not an unlikely occurrence for 8 students to choose 4, 5 or 6.

3.186 The total cost incurred is $W = 30Y$. Then,

$$E(W) = 30E(Y) = 30(1/.3) = 100, \qquad V(W) = 900V(Y) = 900(.7/.3^2) = 7000.$$

Using the empirical rule, we can construct a interval of three standard deviations about the mean: $100 \pm 3\sqrt{7000}$, or (151, 351).

3.187 Let Y = # of rolls until the player stops. Then, Y is geometric with $p = 5/6$.

a. $P(Y = 3) = (1/6)^2(5/6) = .023$.

b. $E(Y) = 6/5 = 1.2$.

c. Let X = amount paid to player. Then, $X = 2^{Y-1}$.

$$E(X) = E(2^{Y-1}) = \sum_{y=1}^{\infty} 2^{y-1}q^{y-1}p = p\sum_{x=0}^{\infty}(2q)^x = \frac{p}{1-2q},$$ since $2q < 1$. With $p = 5/6$, this is \$1.25.

3.188 The result follows from $P(Y > 1 \mid Y \geq 1) = \dfrac{P(Y > 1)}{P(Y \geq 1)} = \dfrac{P(Y \geq 2)}{P(Y \geq 1)} = \dfrac{1 - P(Y = 1) - P(Y = 0)}{1 - P(Y = 0)}$.

3.189 The random variable Y = # of failures in 10,000 starts is binomial with n = 10,000 and $p = .00001$. Thus, $P(Y \geq 1) = 1 - P(Y = 0) = 1 - (.9999)^{10000} = .09516$.
Poisson approximation: $1 - e^{-.1} = .09516$.

3.190 Answers vary, but with $n = 100$, $\bar{y}$ should be quite close to $\mu = 1$.

3.191 Answers vary, but with $n = 100$, s^2 should be quite close to $\sigma^2 = .4$.

3.192 Note that $p(1) = p(2) = \ldots p(6) = 1/6$. From Ex. 3.22, $\mu = 3.5$ and $\sigma^2 = 2.9167$. The interval constructed of two standard deviations about the mean is (.08, 6.92) which contains 100% of the possible values for Y.

3.193 Let Y_1 = # of defectives from line I, Y_2 is defined similarly. Then, both Y_1 and Y_2 are binomial with $n = 5$ and defective probability p. In addition, $Y_1 + Y_2$ is also binomial with $n = 10$ and defective probability p. Thus,

$$P(Y_1 = 2 \mid Y_1 + Y_2 = 4) = \frac{P(Y_1 = 2)P(Y_1 = 2)}{P(Y_1 + Y_2 = 4)} = \frac{\binom{5}{2}p^2q^3\binom{5}{2}p^2q^3}{\binom{10}{4}p^5q^5} = \frac{\binom{5}{2}\binom{5}{2}}{\binom{10}{4}} = 0.476.$$

Notice that the probability does not depend on p.

3.194 The possible outcomes of interest are:

WLLLLLLLLLL, *LWLLLLLLLLL*, *LLWLLLLLLLL*

So the desired probability is $.1(.9)^{10} + .9(.1)(.9)^9 + (.9)^2(.1)(.9)^8 = 3(.1)(.9)^{10} = .104$.

3.195 Let Y = # of imperfections in one–square yard of weave. Then, Y is Poisson with $\lambda = 4$.
a. $P(Y \geq 1) = 1 - P(Y = 0) = 1 - e^{-4} = .982$.
b. Let W = # of imperfections in three–square yards of weave. Then, W is Poisson with $\lambda = 12$. $P(W \geq 1) = 1 - P(W = 0) = 1 - e^{-12}$.

3.196 For an 8–square yard bolt, let X = # of imperfections so that X is Poisson with $\lambda = 32$. Thus, $C = 10X$ is the cost to repair the weave and

$$E(C) = 10E(X) = \$320 \text{ and } V(C) = 100V(X) = 3200.$$

3.197 **a.** Let Y = # of samples with at least one bacteria colony. Then, Y is binomial with $n = 4$ and $p = P$(at least one bacteria colony) $= 1 - P$(no bacteria colonies) $= 1 - e^{-2} = .865$ (by the Poisson). Thus, $P(Y \geq 1) = 1 - P(Y = 0) = 1 - (.135)^4 = .9997$.

b. Following the above, we require $1 - (.135)^n = .95$ or $(.135)^n = .05$. Solving for n, we have $n = \frac{\ln(.05)}{\ln(.135)} = 1.496$, so take $n = 2$.

3.198 Let Y = # of neighbors for a seedling within an area of size A. Thus, Y is Poisson with $\lambda = A*d$, where for this problem $d = 4$ per square meter.
a. Note that "within 1 meter" denotes an area $A = \pi(1\ m)^2 = \pi\ m^2$. Thus, $P(Y = 0) = e^{-4\pi}$.
b. "Within 2 meters" denotes an area $A = \pi(2\ m)^2 = 4\pi\ m^2$. Thus,

$$P(Y \leq 3) = P(Y = 0) + P(Y = 1) + P(Y = 2) + P(Y = 3) = \sum_{y=0}^{3} \frac{(16\pi)^y e^{-16\pi}}{y!}.$$

3.199 **a.** Using the binomial model with $n = 1000$ and $p = 30/100{,}000$, let $\lambda \approx np = 1000(30/100000) = .300$ for the Poisson approximation.

b. Let Y = # of cases of IDD. $P(Y \geq 2) = 1 - P(Y = 0) - P(Y = 1) = 1 - .963 = .037$.

3.200 Note that

$$(q + pe^t)^n = \left[q + p(1 + t + \tfrac{t^2}{2!} + \tfrac{t^3}{3!} + \cdots)\right]^n = \left[1 + pt + p\tfrac{t^2}{2!} + p\tfrac{t^3}{3!} + \cdots\right]^n.$$

Expanding the above multinomial (but only showing the first three terms gives

$$(q + pe^t)^n = 1^n + (np)t1^{n-1} + \left[n(n-1)p^2 + np\right]\tfrac{t^2}{2!}1^{n-2} + \cdots$$

The coefficients agree with the first and second moments for the binomial distribution.

3.201 From Ex. 103 and 106, we have that $\mu = 100$ and $\sigma = \sqrt{1666.67} = 40.825$. Using an interval of two standard deviations about the mean, we obtain $100 \pm 2(40.825)$ or $(18.35, 181.65)$

3.202 Let W = # of drivers who wish to park and W' = # of cars, which is Poisson with mean λ.
a. Observe that

$$P(W=k) = \sum_{n=k}^{\infty} P(W=k \mid W'=n)P(W'=n) = \sum_{n=k}^{\infty} \frac{n!}{k!(n-k)!} p^k q^{n-k} e^{-\lambda} \frac{\lambda^n}{n!}$$

$$= \lambda^k e^{-\lambda}\left(\frac{p^k}{k!}\right)\sum_{n=k}^{\infty} \frac{q^{n-k}}{(n-k)!}\lambda^{n-k} = \frac{(\lambda p)^k}{k!}\sum_{j=0}^{\infty}\frac{q^j}{j!}\lambda^j = \frac{(\lambda p)^k}{k!}e^{-\lambda}e^{q\lambda}$$

$$= \frac{(\lambda p)^k}{k!}e^{-\lambda p},\ k = 0, 1, 2, \ldots .$$

Thus, $P(W=0) = e^{-\lambda p}$.

b. This is a Poisson distribution with mean λp.

3.203 Note that $Y(t)$ has a negative binomial distribution with parameters $r = k$, $p = e^{-\lambda t}$.
a. $E[Y(t)] = ke^{\lambda t}$, $V[Y(t)] = \frac{k(1-e^{-\lambda t})}{e^{-2\lambda t}} = k(e^{2\lambda t} - e^{\lambda t})$.

b. With $k = 2$, $\lambda = .1$, $E[Y(5)] = 3.2974$, $V[Y(5)] = 2.139$.

3.204 Let Y = # of left–turning vehicles arriving while the light is red. Then, Y is binomial with $n = 5$ and $p = .2$. Thus, $P(Y \le 3) = .993$.

3.205 One solution: let Y = # of tosses until 3 sixes occur. Therefore, Y is negative binomial where $r = 3$ and $p = 1/6$. Then, $P(Y = 9) = \binom{8}{2}\left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^6 = .0434127$. Note that this probability contains all events where a six occurs on the 9th toss. Multiplying the above probability by 1/6 gives the probability of observing 4 sixes in 10 trials, where a six occurs on the 9th and 10th trial: (.0424127)(1/6) = .007235.

3.206 Let Y represent the gain to the insurance company for a particular insured driver and let P be the premium charged to the driver. Given the information, the probability distribution for Y is given by:

y	$p(y)$
P	.85
P – 2400	.15(.80) = .12
P – 7200	.15(.12) = .018
P – 12,000	.15(.08) = .012

If the expected gain is 0 (breakeven), then:
$E(Y) = P(.85) + (P - 2400).12 + (P - 7200).018 + (P - 12000).012 = 0$, so P = \$561.60.

3.207 Use the Poisson distribution with $\lambda = 5$.
a. $p(2) = .084$, $P(Y \le 2) = .125$.
b. $P(Y > 10) = 1 - P(Y \le 10) = 1 - .986 = .014$, which represents an unlikely event.

3.208 If the general public was split 50–50 in the issue, then Y = # of people in favor of the proposition is binomial with $n = 2000$ and $p = .5$. Thus,

$$E(Y) = 2000(.5) = 1000 \text{ and } V(Y) = 2000(.5)(.5) = 500.$$

Since $\sigma = \sqrt{500} = 22.36$, observe that 1373 is $(1373 - 1000)/22.36 = 16.68$ standard deviations above the mean. Such a value is unlikely.

3.209 Let Y = # of contracts necessary to obtain the third sale. Then, Y is negative binomial with $r = 3$, $p = .3$. So, $P(Y < 5) = P(Y = 3) + P(Y = 4) = .3^3 + 3(.3)^3(.7) = .0837$.

3.210 In Example 3.22, $\lambda = \mu = 3$ and $\sigma^2 = 3$ and that $\sigma = \sqrt{3} = 1.732$. Thus, $P(|Y - 3| \le 2(1.732)) = P(-.464 \le Y \le 6.464) = \text{P}(\text{Y} \le 6) = .966$. This is consistent with the empirical rule (approximately 95%).

3.211 There are three scenarios:

- if she stocks two items, both will sell with probability 1. So, her profit is \$.40.
- if she stocks three items, two will sell with probability .1 (a loss of .60) and three will sell with probability .9. Thus, her expected profit is (–.60).1 + .60(.9) = \$.48.
- if she stocks four items, two will sell with probability .1 (a loss of 1.60), three will sell with probability .4 (a loss of .40), and four will sell with probability .5 (a gain of .80. Thus, her expected profit is (–1.60).1 + (–.40).4 + (.80).5 = \$.08

So, to maximize her expected profit, stock three items.

3.212 Note that: $\dfrac{\binom{r}{y}\binom{N-r}{n-y}}{\binom{N}{n}} = \frac{n!}{y!(n-y)!}\left[\left(\frac{r}{N}\right)\left(\frac{r-1}{N-1}\right)\left(\frac{r-2}{N-2}\right)\cdots\left(\frac{r-y+1}{N-y+1}\right)\right] \times \left[\left(\frac{N-r}{N-y}\right)\left(\frac{N-r-1}{N-y-1}\right)\left(\frac{N-r-2}{N-y-2}\right)\cdots\left(\frac{N-r-n-y+1}{N-n+1}\right)\right]$. In the first bracketed part, each quotient in parentheses has a limiting value of p. There are y such quotients. In the second bracketed part, each quotient in parentheses has a limiting value of $1 - p = q$. There are $n - y$ such quotients, Thus,

$$\frac{\binom{r}{y}\binom{N-r}{n-y}}{\binom{N}{n}} \to \binom{n}{y} p^y q^{n-y} \text{ as } N \to \infty$$

3.213 **a.** The probability is $p(10) = \dfrac{\binom{40}{10}\binom{60}{10}}{\binom{100}{20}} = .1192$ (found by `dhyper(10, 40, 60, 20)` in R).

b. The binomial approximation is $\left(\frac{20!}{10!10!}\right).4^{10}(.6)^{10} = .117$, a close value to the above (exact) answer.

3.214 Define: A = accident next year B = accident this year C = safe driver

Thus, $P(C) = .7$, $P(A|C) = .1 = P(B|C)$, and $P(A\,|\,\overline{C}) = P(B\,|\,\overline{C}) = .5$. From Bayes' rule,

$$P(C|B) = \frac{P(B\,|\,C)P(C)}{P(B\,|\,C)P(C) + P(B\,|\,\overline{C})P(\overline{C})} = \frac{.1(.7)}{.1(.7) + .5(.3)} = 7/22.$$

Now, we need $P(A|B)$. Note that since $C \cup \overline{C} = S$, this conditional probability is equal to

$P(A \cap (C \cup \overline{C})\,|\,B) = P(A \cap C\,|\,B) + P(A \cap \overline{C}\,|\,B) = P(A \bigcap C\,|\,B) + P(A \cap \overline{C}\,|\,B)$, or
$P(A\,|\,B) = P(C\,|\,B)P(A\,|\,C \cap B) + P(\overline{C}\,|\,B)P(A\,|\,\overline{C} \cap B)$ = 7/22(.1) + 15/22(.5) = .3727.

So, the premium should be 400(.3727) = \$149.09.

3.215 **a.** Note that for (2), there are two possible values for N_2, the number of tests performed: 1 and $k + 1$. If $N_2 = 1$, all of the k people are healthy and this probability is $(.95)^k$. Thus, $P(N_2 = k + 1) = 1 - (.95)^k$. Thus, $E(N_2) = 1(.95)^k + (k + 1)(1 - .95^k) = 1 + k(1 - .95^k)$. This expectation holds for each group, so that for n groups the expected number of tests is $n[1 + k(1 - .95^k)]$.

b. Writing the above as $g(k) = \frac{N}{k}[1 + k(1 - .95^k)]$, where $n = \frac{N}{k}$, we can minimize this with respect to k. Note that $g'(k) = \frac{1}{k^2} + (.95^k)\ln(.95)$, a strictly decreasing function. Since k must be an integer, it is found that $g(k)$ is minimized at $k = 5$ and $g(5) = .4262$.

c. The expected number of tests is $.4262N$, compared to the N tests is (1) is used. The savings is then $N - .4262N = .5738N$.

3.216 **a.** $P(Y = n) = \dfrac{\binom{r}{n}\binom{N-r}{n-n}}{\binom{N}{n}} = \frac{r!}{N!} \times \frac{(N-n)!}{(r-n)!} = \frac{r(r-1)(r-2)\cdots(r-n+1)}{N(N-1)(N-2)\cdots(N-n+1)}$.

b. Since for integers $a > b$, $\dfrac{\binom{a}{b}}{\binom{a}{b+1}} = \frac{b+1}{a-b}$, apply this result to find that

$$\frac{p(y|r_1)}{p(y+1|r_1)} = \frac{y+1}{r_1-y} \cdot \frac{N-r_1+n+y+1}{n-y} \text{ and } \frac{p(y|r_2)}{p(y+1|r_2)} = \frac{y+1}{r_2-y} \cdot \frac{N-r_{21}+n+y+1}{n-y}.$$

With $r_1 < r_2$, it follows that $\frac{p(y|r_1)}{p(y|r_{21})} > \frac{p(y+1|r_1)}{p(y+1|r_2)}$.

c. Note that from the binomial theorem, $(1+a)^{N_1} = \sum_{k=0}^{N_1}\binom{N_1}{k}y^k$. So,

$$(1+a)^{N_1}(1+a)^{N_2} = \left[\binom{N_1}{0} + \binom{N_1}{1}y + \cdots \binom{N_1}{N_1}y^{N_1}\right] \times \left[\binom{N_1}{0} + \binom{N_1}{1}y + \cdots \binom{N_1}{N_1}y^{N_1}\right] =$$

$(1+a)^{N_1+N_2} = \left[\binom{N_1+N_2}{0} + \binom{N_1+N_2}{1}y + \cdots \binom{N_1+N_2}{N_1+N_2}y^{N_1+N_2}\right]$. Since these are equal, the coefficient of every y^n must be equal. In the second equality, the coefficient is

$$\binom{N_1+N_2}{n}.$$

In the first inequality, the coefficient is given by the sum

$\binom{N_1}{0}\binom{N_2}{n} + \binom{N_1}{1}\binom{N_2}{n-1} + \cdots + \binom{N_1}{n}\binom{N_2}{0} = \sum_{k=0}^{n}\binom{N_1}{k}\binom{N_2}{n-k}$, thus the relation holds.

d. The result follows from part **c** above.

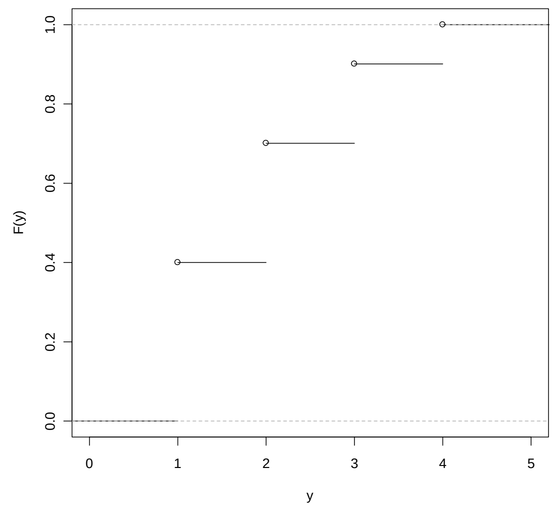
3.217 $E(Y) = \sum_{y=0}^{n} \frac{y\binom{r}{y}\binom{N-r}{n-y}}{\binom{N}{n}} = r\sum_{y=1}^{n}\left[\frac{(r-1)!}{(y-1)!(r-y)!}\right]\left[\frac{\binom{N-r}{n-y}}{\binom{N}{n}}\right] = r\sum_{y=1}^{n}\left[\frac{\binom{r-1}{y-1}\binom{N-r}{n-y}}{\binom{N}{n}}\right]$. In this sum, let $x = y - 1$:

$r\sum_{x=o}^{n-1}\left[\frac{\binom{r-1}{x}\binom{N-r}{n-x-1}}{\binom{N}{n}}\right] = r\sum_{x=o}^{n-1}\left[\frac{\binom{r-1}{x}\binom{N-r}{n-x-1}}{\left(\frac{N}{n}\right)\binom{N-1}{n-1}}\right] = \frac{nr}{N}$.

3.218 $E[Y(Y-1)] = \sum_{y=0}^{n} \frac{y(y-1)\binom{r}{y}\binom{N-r}{n-y}}{\binom{N}{n}} = \sum_{y=0}^{n}\left[\frac{r(r-1)(r-2)!}{y(y-1)(y-2)!(r-y)!}\right]\left[\frac{\binom{N-r}{n-y}}{\binom{N}{n}}\right] = r(r-1)\sum_{y=2}^{n}\left[\frac{\binom{r-2}{y-2}\binom{N-r}{n-y}}{\binom{N}{n}}\right]$. In this sum, let $x = y - 2$ to obtain the expectation $\frac{r(r-1)n(n-1)}{N(N-1)}$. From this result, the variance of the hypergeometric distribution can also be calculated.

Chapter 4: Continuous Variables and Their Probability Distributions

4.1 **a.** $F(y) = P(Y \le y) = \begin{cases} 0 & y < 1 \\ .4 & 1 \le y < 2 \\ .7 & 2 \le y < 3 \\ .9 & 3 \le y < 4 \\ 1 & y \ge 4 \end{cases}$



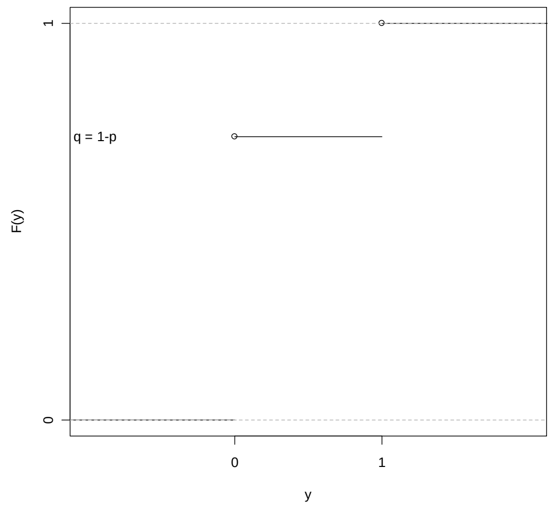
b. The graph is above.

4.2 **a.** $p(1) = .2$, $p(2) = (1/4)4/5 = .2$, $p(3) = (1/3)(3/4)(4/5) = 2.$, $p(4) = .2$, $p(5) = .2$.

b. $F(y) = P(Y \le y) = \begin{cases} 0 & y < 1 \\ .2 & 1 \le y < 2 \\ .4 & 2 \le y < 3 \\ .6 & 3 \le y < 4 \\ .8 & 4 \le y < 5 \\ 1 & y \ge 5 \end{cases}$

c. $P(Y < 3) = F(2) = .4$, $P(Y \le 3) = .6$, $P(Y = 3) = p(3) = .2$

d. No, since Y is a discrete random variable.



4.3 **a.** The graph is above.

b. It is easily shown that all three properties hold.

4.4 A binomial variable with $n = 1$ has the Bernoulli distribution.

4.5 For $y = 2, 3, \ldots, F(y) - F(y-1) = P(Y \le y) - P(Y \le y-1) = P(Y = y) = p(y)$. Also, $F(1) = P(Y \le 1) = P(Y = 1) = p(1)$.

4.6 **a.** $F(i) = P(Y \le i) = 1 - P(Y > i) = 1 - P(1^{\text{st}}\ i$ trials are failures$) = 1 - q^i$.

b. It is easily shown that all three properties hold.

4.7 **a.** $P(2 \le Y < 5) = P(Y \le 4) - P(Y \le 1) = .967 - .376 = 0.591$
$P(2 < Y < 5) = P(Y \le 4) - P(Y \le 2) = .967 - .678 = .289$.
Y is a discrete variable, so they are not equal.

b. $P(2 \le Y \le 5) = P(Y \le 5) - P(Y \le 1) = .994 - .376 = 0.618$
$P(2 < Y \le 5) = P(Y \le 5) - P(Y \le 2) = .994 - .678 = 0.316$.
Y is a discrete variable, so they are not equal.

c. Y is not a continuous random variable, so the earlier result do not hold.

4.8 **a.** The constant $k = 6$ is required so the density function integrates to 1.

b. $P(.4 \le Y \le 1) = .648$.

c. Same as part b. above.

d. $P(Y \le .4 \mid Y \le .8) = P(Y \le .4)/P(Y \le .8) = .352/.896 = 0.393$.

e. Same as part **d.** above.

4.9 **a.** Y is a discrete random variable because $F(y)$ is not a continuous function. Also, the set of possible values of Y represents a countable set.

b. These values are 2, 2.5, 4, 5.5, 6, and 7.

c. $p(2) = 1/8$, $p(2.5) = 3/16 – 1/8 = 1/16$, $p(4) = 1/2 – 3/16 = 5/16$, $p(5.5) = 5/8 – 1/2 = 1/8$, $p(6) = 11/16 – 5/8 = 1/16$, $p(7) = 1 – 11/16 = 5/16$.

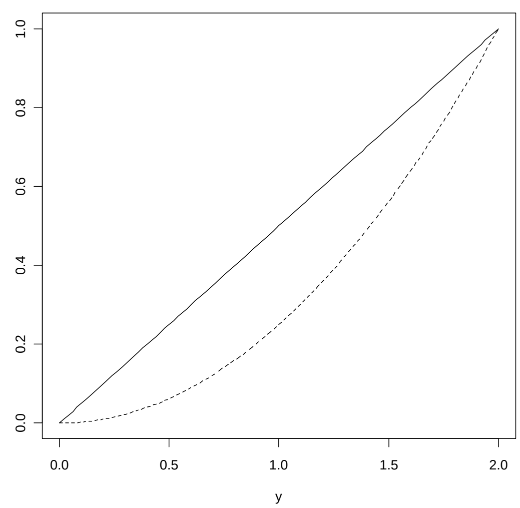
d. $P(Y \le \phi_{.5}) = F(\phi_{.5}) = .5$, so $\phi_{.5} = 4$.

4.10 **a.** $F(\phi_{.95}) = \int_0^{\phi_{.95}} 6y(1-y)dy = .95$, so $\phi_{.95} = 0.865$.

b. Since Y is a continuous random variable, $y_0 = \phi_{.95} = 0.865$.

4.11 **a.** $\int_0^2 cydy = \left[cy^2/2\right]_0^2 = 2c = 1$, so $c = 1/2$.

b. $F(y) = \int_{-\infty}^{y} f(t)dt = \int_0^y \frac{t}{2}dt = \frac{y^2}{4}$, $0 \le y \le 2$.



c. Solid line: $f(y)$; dashed line: $F(y)$

d. $P(1 \le Y \le 2) = F(2) – F(1) = 1 – .25 = .75$.

e. Note that $P(1 \le Y \le 2) = 1 – P(0 \le Y < 1)$. The region $(0 \le y < 1)$ forms a triangle (in the density graph above) with a base of 1 and a height of .5. So, $P(0 \le Y < 1) = \frac{1}{2}(1)(.5) = .25$ and $P(1 \le Y \le 2) = 1 – .25 = .75$.

4.12 **a.** $F(-\infty) = 0$, $F(\infty) = 1$, and $F(y_1) - F(y_2) = e^{-y_2^2} - e^{-y_1^2} > 0$ provided $y_1 > y_2$.

b. $F(\phi_{.3}) = 1 - e^{-\phi_{.3}^2} = .3$, so $\phi_{.3} = \sqrt{-\ln(.7)} = 0.5972$.

c. $f(y) = F'(y) = 2ye^{-y^2}$ for $y \geq 0$ and 0 elsewhere.

d. $P(Y \geq 200) = 1 - P(Y < 200) = 1 - P(Y \leq 200) = 1 - F(2) = e^{-4}$.

e. $P(Y > 100 \mid Y \leq 200) = P(100 < Y \leq 200)/P(Y \leq 200) = [F(2) - F(1)]/F(2) = \frac{e^{-1}-e^{-4}}{1-e^{-4}}$.

4.13 **a.** For $0 \leq y \leq 1$, $F(y) = \int_0^y t\,dt = y^2/2$. For $1 < y \leq 1.5$, $F(y) = \int_0^1 t\,dt + \int_1^y dt = 1/2 + y - 1$ $= y - 1/2$. Hence,

$$F(y) = \begin{cases} 0 & y < 0 \\ y^2/2 & 0 \leq y \leq 1 \\ y - 1/2 & 1 < y \leq 1.5 \\ 1 & y > 1.5 \end{cases}$$

b. $P(0 \leq Y \leq .5) = F(.5) = 1/8$.

c. $P(.5 \leq Y \leq 1.2) = F(1.2) - F(.5) = 1.2 - 1/2 - 1/8 = .575$.

4.14 **a.** A triangular distribution.

b. For $0 < y < 1$, $F(y) = \int_0^y t\,dt = y^2/2$. For $1 \leq y < 2$, $F(y) = \int_0^1 t\,dt + \int_1^y (2-t)dt = 2y - \frac{y^2}{2} - 1$.

c. $P(.8 \leq Y \leq 1.2) = F(1.2) - F(.8) = .36$.

d. $P(Y > 1.5 \mid Y > 1) = P(Y > 1.5)/P(Y > 1) = .125/.5 = .25$.

4.15 **a.** For $b \geq 0$, $f(y) \geq 0$. Also, $\int_{-\infty}^{\infty} f(y) = \int_b^{\infty} b/y^2 = -b/y\big]_b^{\infty} = 1$.

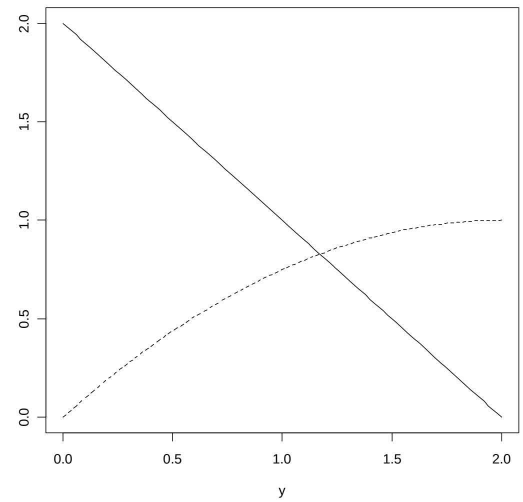
b. $F(y) = 1 - b/y$, for $y \geq b$, 0 elsewhere.

c. $P(Y > b + c) = 1 - F(b + c) = b/(b + c)$.

d. Applying part c., $P(Y > b + d \mid Y > b + c) = (b + c)/(b + d)$.

4.16 **a.** $\int_0^2 c(2-y)dy = c\left[2y - y^2/2\right]_0^2 = 2c = 1$, so $c = 1/2$.

b. $F(y) = y - y^2/4$, for $0 \le y \le 2$.

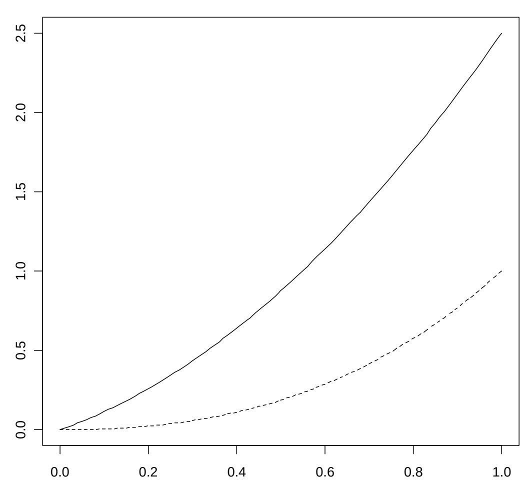


c. Solid line: $f(y)$; dashed line: $F(y)$

d. $P(1 \le Y \le 2) = F(2) - F(1) = 1/4$.

4.17 **a.** $\int_0^1 (cy^2 + y)dy = \left[cy^3/3 + y^2/2\right]_0^2 = 1$, $c = 3/2$.

b. $F(y) = y^3/2 + y^2/2$ for $0 \le y \le 1$.



c. Solid line: $f(y)$; dashed line: $F(y)$

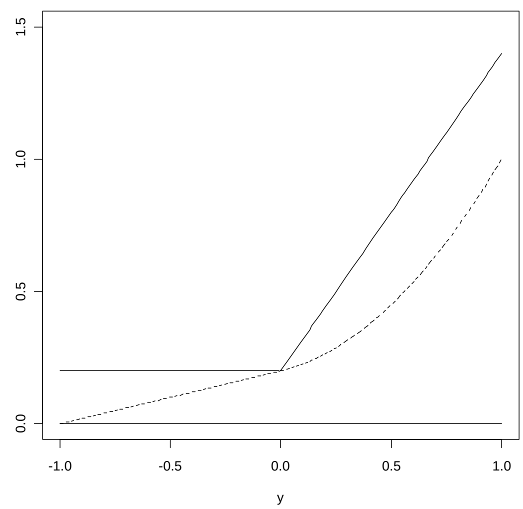
d. $F(-1) = 0$, $F(0) = 0$, $F(1) = 1$.

e. $P(Y < .5) = F(.5) = 3/16$.

f. $P(Y \geq .5 \mid Y \geq .25) = P(Y \geq .5)/P(Y \geq .25) = 104/123$.

4.18 **a.** $\int_{-1}^{0} .2dy + \int_{0}^{1} (.2 + cy)dy = .4 + c/2 = 1$, so $c = 1.2$.

b. $F(y) = \begin{cases} 0 & y \leq -1 \\ .2(1+y) & -1 < y \leq 0 \\ .2(1+y+3y^2) & 0 < y \leq 1 \\ 1 & y > 1 \end{cases}$



c. Solid line: $f(y)$; dashed line: $F(y)$

d. $F(-1) = 0$, $F(0) = .2$, $F(1) = 1$

e. $P(Y > .5 \mid Y > .1) = P(Y > .5)/P(Y > .1) = .55/.774 = .71$.

4.19 **a.** Differentiating $F(y)$ with respect to y, we have

$$f(y) = \begin{cases} 0 & y \leq 0 \\ .125 & 0 < y < 2 \\ .125y & 2 \leq y < 4 \\ 0 & y \geq 4 \end{cases}$$

b. $F(3) - F(1) = 7/16$

c. $1 - F(1.5) = 13/16$

d. $7/16/(9/16) = 7/9$.

4.20 From Ex. 4.16:

$$E(Y) = \int_0^2 .5y(2-y)dy = \frac{y^2}{2} + \frac{y^3}{6}\Big]_0^2 = 2/3, \; E(Y^2) = \int_0^2 .5y^2(2-y)dy = \frac{y^3}{3} + \frac{y^4}{8}\Big]_0^2 = 2/3.$$

So, $V(Y) = 2/3 - (2/3)^2 = 2/9$.

4.21 From Ex. 4.17:

$$E(Y) = \int_0^1 1.5y^3 + y^2)dy = \frac{3y^4}{8} + \frac{y^3}{3}\Big]_0^1 = 17/24 = .708.$$

$$E(Y^2) = \int_0^1 1.5y^4 + y^3)dy = \frac{3y^5}{10} + \frac{y^4}{4}\Big]_0^1 = 3/10 + 1/4 = .55.$$

So, $V(Y) = .55 - (.708)2 = .0487$.

4.22 From Ex. 4.18:

$$E(Y) = \int_{-1}^0 .2ydy + \int_0^1 (.2y + 1.2y^2)dy = .4, \; E(Y^2) = \int_{-1}^0 .2y^2dy + \int_0^1 (.2y^2 + 1.2y^3)dy = 1.3/3.$$

So, $V(Y) = 1.3/3 - (.4)2 = .2733$.

4.23 **1.** $E(c) = \int_{-\infty}^{\infty} cf(y)dy = c\int_{-\infty}^{\infty} f(y)dy = c(1) = c$.

2. $E[cg(Y)] = \int_{-\infty}^{\infty} cg(y)f(y)dy = c\int_{-\infty}^{\infty} g(y)f(y)dy = cE[g(Y)]$.

3. $E[g_1(Y) + g_2(Y) + \cdots g_k(Y)] = \int_{-\infty}^{\infty} [g_1(y) + g_2(y) + \cdots g_k(y)]f(y)dy$

$= \int_{-\infty}^{\infty} g_1(y)f(y)dy + \int_{-\infty}^{\infty} g_2(y)f(y)dy + \cdots \int_{-\infty}^{\infty} g_k(y)f(y)dy$

$= E[g_1(Y)] + E[g_2(Y)] + \cdots E[g_k(Y)]$.

4.24 $V(Y) = E\{[Y - E(Y)]^2\} = E\{Y^2 - 2YE(Y) + [E(Y)]^2\} = E(Y^2) - 2[E(Y)]^2 + [E(Y)]^2$
$= E(Y^2) - [E(Y)]^2 = \sigma^2$.

4.25 Ex. 4.19:

$$E(Y) = \int_0^2 .125ydy + \int_2^4 .125y^2dy = 31/12, \; E(Y^2) = \int_0^2 .125y^2dy + \int_2^4 .125y^3dy = 47/6.$$

So, $V(Y) = 47/6 - (31/12)^2 = 1.16$.

4.26 **a.** $E(aY + b) = \int_{-\infty}^{\infty} (ay + b)f(y)dy = \int_{-\infty}^{\infty} ayf(y)dy + \int_{-\infty}^{\infty} bf(y)dy = aE(Y) + b = a\mu + b$.

b. $V(aY + b) = E\{[aY + b - E(aY + b)]^2\} = E\{[aY + b - a\mu - b]^2\} = E\{a^2[Y - \mu]^2\}$
$= a^2V(Y) = a^2\sigma^2$.

4.27 First note that from Ex. 4.21, $E(Y) = .708$ and $V(Y) = .0487$. Then,
$E(W) = E(5 - .5Y) = 5 - .5E(Y) = 5 - .5(.708) = \$4.65$.
$V(W) = V(5 - .5Y) = .25V(Y) = .25(.0487) = .012$.

4.28 **a.** By using the methods learned in this chapter, $c = 105$.

b. $E(Y) = 105\int_0^1 y^3(1-y)^4\,dy = 3/8$.

4.29 $E(Y) = .5\int_{59}^{61} y\,dy = .5\frac{y^2}{2}\Big]_{59}^{61} = 60$, $E(Y^2) = .5\int_{59}^{61} y^2\,dy = .5\frac{y^3}{3}\Big]_{59}^{61} = 3600\tfrac{1}{3}$. Thus,
$V(Y) = 3600\tfrac{1}{3} - (60)^2 = \tfrac{1}{3}$.

4.30 **a.** $E(Y) = \int_0^1 2y^2\,dy = 2/3$, $E(Y^2) = \int_0^1 2y^3\,dy = 1/2$. Thus, $V(Y) = 1/2 - (2/3)^2 = 1/18$.

b. With $X = 200Y - 60$, $E(X) = 200(2/3) - 60 = 220/3$, $V(X) = 20000/9$.

c. Using Tchebysheff's theorem, a two standard deviation interval about the mean is given by $220/3 \pm 2\sqrt{20000/9}$ or $(-20.948, 167.614)$.

4.31 $E(Y) = \int_2^6 y\left(\tfrac{3}{32}\right)(y-2)(6-y)\,dy = 4$.

4.32 **a.** $E(Y) = \tfrac{3}{64}\int_0^4 y^3(4-y)\,dy = \tfrac{3}{64}\left[y^4 - \tfrac{y^5}{5}\right]_0^4 = 2.4$. $V(Y) = .64$.

b. $E(200Y) = 200(2.4) = \$480$, $V(200Y) = 200^2(.64) = 25{,}600$.

c. $P(200Y > 600) = P(Y > 3) = \tfrac{3}{64}\int_3^4 y^2(4-y)\,dy = .2616$, or about 26% of the time the cost will exceed \$600 (fairly common).

4.33 **a.** $E(Y) = \tfrac{3}{8}\int_5^7 y(7-y)^2\,dy = \tfrac{3}{8}\left[\tfrac{49}{2}y^2 - \tfrac{14}{3}y^3 + \tfrac{y^4}{4}\right]_5^7 = 5.5$

$E(Y^2) = \tfrac{3}{8}\int_5^7 y^2(7-y)^2\,dy = \tfrac{3}{8}\left[\tfrac{49}{3}y^3 - \tfrac{14}{4}y^4 + \tfrac{y^5}{5}\right]_5^7 = 30.4$, so $V(Y) = .15$.

b. Using Tchebysheff's theorem, a two standard deviation interval about the mean is given by $5.5 \pm 2\sqrt{.15}$ or $(4.725, 6.275)$. Since $Y \geq 5$, the interval is $(5, 6.275)$.

c. $P(Y < 5.5) = \tfrac{3}{8}\int_5^{5.5} (7-y)^2\,dy = .5781$, or about 58% of the time (quite common).

4.34 $E(Y)=\int_0^\infty yf(y)dy=\int_0^\infty\left(\int_0^y dt\right)f(y)dy=\int_0^\infty\left(\int_y^\infty f(y)dy\right)dt=\int_0^\infty P(Y>y)dy=\int_0^\infty[1-F(y)]dy.$

4.35 Let $\mu = E(Y)$. Then, $E[(Y-a)^2]=E[(Y-\mu+\mu-a)^2]$

$$= E[(Y-\mu)^2]-2E[(Y-\mu)(\mu-a)]+(\mu-a)^2$$
$$= \sigma^2+(\mu-a)^2.$$

The above quantity is minimized when $\mu = a$.

4.36 This is also valid for discrete random variables — the properties of expected values used in the proof hold for both continuous and discrete random variables.

4.37 $E(Y)=\int_{-\infty}^{\infty} yf(y)dy=\int_{-\infty}^{0} yf(y)dy+\int_0^\infty yf(y)dy$. In the first integral, let $w = -y$, Then,

$$E(Y)=-\int_0^\infty wf(-w)dy+\int_0^\infty yf(y)dy=-\int_0^\infty wf(w)dy+\int_0^\infty yf(y)dy=0.$$

4.38 **a.** $F(y)=\begin{cases}0 & y<0\\ \int_0^y 1dy=y & 0\le y\le 1\\ 1 & y>1\end{cases}$

b. $P(a \le Y \le a + b) = F(a + b) - F(a) = a + b - a = b$.

4.39 The distance Y is uniformly distributed on the interval A to B, If she is closer to A, she has landed in the interval (A, $\frac{A+B}{2}$). This is one half the total interval length, so the probability is .5. If her distance to A is more than three times her distance to B, she has landed in the interval ($\frac{3B+A}{4}$, B). This is one quarter the total interval length, so the probability is .25.

4.40 The probability of landing past the midpoint is 1/2 according to the uniform distribution. Let X = # parachutists that land past the midpoint of (A, B). Therefore, X is binomial with $n = 3$ and $p = 1/2$. $P(X = 1) = 3(1/2)^3 = .375$.

4.41 First find $E(Y^2)=\frac{1}{\theta_2-\theta_1}\int_{\theta_1}^{\theta_2} y^2dy=\frac{1}{\theta_2-\theta_1}\left[\frac{y^3}{3}\right]_{\theta_1}^{\theta_2}=\frac{\theta_2^3-\theta_1^3}{3(\theta_2-\theta_1)}=\frac{\theta_1^2+\theta_1\theta_2+\theta_2^2}{3}$. Thus,

$$V(Y)=\frac{\theta_1^2+\theta_1\theta_2+\theta_2^2}{3}-\left(\frac{\theta_2+\theta_1}{2}\right)^2=\frac{(\theta_2-\theta_1)^2}{12}.$$

4.42 The distribution function is $F(y) = \frac{y - \theta_1}{\theta_2 - \theta_1}$, for $\theta_1 \le y \le \theta_2$. For $F(\phi_{.5}) = .5$, then $\phi_{.5} = \theta_1 + .5(\theta_2 - \theta_1) = .5(\theta_2 + \theta_1)$. This is also the mean if the distribution.

4.43 Let $A = \pi R^2$, where R has a uniform distribution on the interval (0, 1). Then,

$$E(A) = \pi E(R^2) = \pi\int_0^1 r^2 dr = \frac{\pi}{3}$$

$$V(A) = \pi^2 V(R^2) = \pi^2[E(R^4) - \left(\frac{1}{3}\right)^2] = \pi^2\left[\int_0^1 r^4 dr - \left(\frac{1}{3}\right)^2\right] = \pi^2\left[\frac{1}{5} - \left(\frac{1}{3}\right)^2\right] = \frac{4\pi^2}{45}.$$

4.44 **a.** Y has a uniform distribution (constant density function), so $k = 1/4$.

b. $F(y) = \begin{cases} 0 & y < -2 \\ \int_{-2}^{y} \frac{1}{4} dy = \frac{y+2}{4} & -2 \le y \le 2 \\ 1 & y > 2 \end{cases}$

4.45 Let Y = low bid (in thousands of dollars) on the next intrastate shipping contract. Then, Y is uniform on the interval (20, 25).
a. $P(Y < 22) = 2/5 = .4$
b. $P(Y > 24) = 1/5 = .2$.

4.46 Mean of the uniform: $(25 + 20)/2 = 22.5$.

4.47 The density for Y = delivery time is $f(y) = \frac{1}{4}$, $1 \le y \le 5$. Also, $E(Y) = 3$, $V(Y) = 4/3$.
a. $P(Y > 2) = 3/4$.
b. $E(C) = E(c_0 + c_1Y^2) = c_0 + c_1E(Y^2) = c_0 + c_1[V(Y) + (E(Y))^2] = c_0 + c_1[4/3 + 9]$

4.48 Let Y = location of the selected point. Then, Y has a uniform distribution on the interval (0, 500).
a. $P(475 \le Y \le 500) = 1/20$
b. $P(0 \le Y \le 25) = 1/20$
c. $P(0 < Y < 250) = 1/2$.

4.49 If Y has a uniform distribution on the interval (0, 1), then $P(Y > 1/4) = 3/4$.

4.50 Let Y = time when the phone call comes in. Then, Y has a uniform distribution on the interval (0, 5). The probability is $P(0 < Y < 1) + P(3 < Y < 4) = .4$.

4.51 Let Y = cycle time. Thus, Y has a uniform distribution on the interval (50, 70). Then,

$P(Y > 65 \mid Y > 55) = P(Y > 65)/P(Y > 55) = .25/(.75) = 1/3$.

4.52 Mean and variance of a uniform distribution: $\mu = 60$, $\sigma^2 = (70–50)^2/12 = 100/3$.

4.53 Let Y = time when the defective circuit board was produced. Then, Y has an approximate uniform distribution on the interval (0, 8).
a. $P(0 < Y < 1) = 1/8$.
b. $P(7 < Y < 8) = 1/8$
c. $P(4 < Y < 5 \mid Y > 4) = P(4 < Y < 5)/P(Y > 4) = (1/8)/(1/2) = 1/4$.

4.54 Let Y = amount of measurement error. Then, Y is uniform on the interval (–.05, .05).
a. $P(-.01 < Y < .01) = .2$
b. $E(Y) = 0$, $V(Y) = (.05 + .05)^2/12 = .00083$.

4.55 Let Y = amount of measurement error. Then, Y is uniform on the interval (–.02, .05).
a. $P(-.01 < Y < .01) = 2/7$
b. $E(Y) = (-.02 + .05)/2 = .015$, $V(Y) = (.05 + .02)^2/12 = .00041$.

4.56 From Example 4.7, the arrival time Y has a uniform distribution on the interval (0, 30). Then, $P(25 < Y < 30 \mid Y > 10) = 1/6/(2/3) = 1/4$.

4.57 The volume of a sphere is given by $(4/3)\pi r^3 = (1/6)\pi d^3$, where r is the radius and d is the diameter. Let D = diameter such that D is uniform distribution on the interval (.01, .05).

Thus, $E(\frac{\pi}{6}D^3) = \frac{\pi}{6}\int_{.01}^{.05} d^3 \frac{1}{4} dd = .0000065\pi$. By similar logic used in Ex. 4.43, it can be found that $V(\frac{\pi}{6}D^3) = .0003525\pi^2$.

4.58 **a.** $P(0 \le Z \le 1.2) = .5 - .1151 = .3849$
b. $P(-.9 \le Z \le 0) = .5 - .1841 - .3159$.
c. $P(.3 \le Z \le 1.56) = .3821 - .0594 = .3227$.
d. $P(-.2 \le Z \le .2) = 1 - 2(.4207) = .1586$.
e. $P(-1.56 \le Z \le -.2) = .4207 - .0594 = .3613$
f. $P(0 \le Z \le 1.2) = .38493$. The desired probability is for a standard normal.

4.59 **a.** $z_0 = 0$.
b. $z_0 = 1.10$
c. $z_0 = 1.645$
d. $z_0 = 2.576$

4.60 The parameter σ must be positive, otherwise the density function could obtain a negative value (a violation).

4.61 Since the density function is symmetric about the parameter μ, $P(Y < \mu) = P(Y > \mu) = .5$. Thus, μ is the median of the distribution, regardless of the value of σ.

4.62 **a.** $P(Z^2 < 1) = P(-1 < Z < 1) = .6826$.
b. $P(Z^2 < 3.84146) = P(-1.96 < Z < 1.96) = .95$.

4.63 **a.** Note that the value 17 is (17 – 16)/1 = 1 standard deviation above the mean. So, $P(Z > 1) = .1587$.
b. The same answer is obtained.

4.64 **a.** Note that the value 450 is (450 – 400)/20 = 2.5 standard deviations above the mean. So, $P(Z > 2.5) = .0062$.
b. The probability is .00618.
c. The top scale is for the standard normal and the bottom scale is for a normal distribution with mean 400 and standard deviation 20.

4.65 For the standard normal, $P(Z > z_0) = .1$ if $z_0 = 1.28$. So, $y_0 = 400 + 1.28(20) = \$425.60$.

4.66 Let Y = bearing diameter, so Y is normal with $\mu = 3.0005$ and $\sigma = .0010$. Thus, Fraction of scrap = $P(Y > 3.002) + P(Y < 2.998) = P(Z > 1.5) + P(Z < -2.5) = .0730$.

4.67 In order to minimize the scrap fraction, we need the maximum amount in the specifications interval. Since the normal distribution is symmetric, the mean diameter should be set to be the midpoint of the interval, or $\mu = 3.000$ in.

4.68 The GPA 3.0 is (3.0 – 2.4)/.8 = .75 standard deviations above the mean. So, $P(Z > .75) = .2266$.

4.69 The z–score for 1.9 is (1.9 – 2.4)/.8 = –.625. Thus, $P(Z < -.625) = .2660$.

4.70 From Ex. 4.68, the proportion of students with a GPA greater than 3.0 is .2266. Let X = # in the sample with a GPA greater than 3.0. Thus, X is binomial with $n = 3$ and $p = .2266$. Then, $P(X = 3) = (.2266)^3 = .0116$.

4.71 Let Y = the measured resistance of a randomly selected wire.
a. $P(.12 \le Y \le .14) = P(\frac{.12-.13}{.005} \le Z \le \frac{.14-.13}{.005}) = P(-2 \le Z \le 2) = .9544$.
b. Let X = # of wires that do not meet specifications. Then, X is binomial with $n = 4$ and $p = .9544$. Thus, $P(X = 4) = (.9544)^4 = .8297$.

4.72 Let Y = interest rate forecast, so Y has a normal distribution with $\mu = .07$ and $\sigma = .026$.
a. $P(Y > .11) = P(Z > \frac{.11-.07}{.026}) = P(Z > 1.54) = .0618$.
b. $P(Y < .09) = P(Z > \frac{.09-.07}{.026}) = P(Z > .77) = .7794$.

4.73 Let Y = width of a bolt of fabric, so Y has a normal distribution with $\mu = 950$ mm and $\sigma = 10$ mm.
a. $P(947 \le Y \le 958) = P(\frac{947-950}{10} \le Z \le \frac{958-950}{10}) = P(-.3 \le Z \le .8) = .406$
b. It is necessary that $P(Y \le c) = .8531$. Note that for the standard normal, we find that $P(Z \le z_0) = .8531$ when $z_0 = 1.05$. So, $c = 950 + (1.05)(10) = 960.5$ mm.

4.74 Let Y = examination score, so Y has a normal distribution with $\mu = 78$ and $\sigma^2 = 36$.
a. $P(Y > 72) = P(Z > -1) = .8413$.
b. We seek c such that $P(Y > c) = .1$. For the standard normal, $P(Z > z_0) = .1$ when $z_0 = 1.28$. So $c = 78 + (1.28)(6) = 85.68$.
c. We seek c such that $P(Y > c) = .281$. For the standard normal, $P(Z > z_0) = .281$ when $z_0 = .58$. So, $c = 78 + (.58)(6) = 81.48$.
d. For the standard normal, $P(Z < -.67) = .25$. So, the score that cuts off the lowest 25% is given by $(-.67)(6) + 78 = 73.98$.
e. Similar answers are obtained.
f. $P(Y > 84 \mid Y > 72) = P(Y > 84)/P(Y > 72) = P(Z > 1)/P(Z > -1) = .1587/.8413 = .1886$.

4.75 Let Y = volume filled, so that Y is normal with mean μ and $\sigma = .3$ oz. They require that $P(Y > 8) = .01$. For the standard normal, $P(Z > z_0) = .01$ when $z_0 = 2.33$. Therefore, it must hold that $2.33 = (8 - \mu)/.3$, so $\mu = 7.301$.

4.76 It follows that $.95 = P(|Y - \mu| < 1) = P(|Z| < 1/\sigma)$, so that $1/\sigma = 1.96$ or $\sigma = 1/1.96 = .5102$.

4.77 **a.** Let Y = SAT math score. Then, $P(Y < 550) = P(Z < .7) = 0.758$.

b. If we choose the same percentile, $18 + 6(.7) = 22.2$ would be comparable on the ACT math test.

4.78 Easiest way: maximize the function $\ln f(y) = -\ln(\sigma\sqrt{2\pi}) - \frac{(y-\mu)^2}{2\sigma^2}$ to obtain the maximum at $y = \mu$ and observe that $f(\mu) = 1/(\sigma\sqrt{2\pi})$.

4.79 The second derivative of $f(y)$ is found to be $f''(y) = \left(\frac{1}{\sigma^3\sqrt{2\pi}}\right)e^{-(y-\mu)^2/2\sigma^2}\left[1 - \frac{(\mu-y)^2}{\sigma^2}\right]$. Setting this equal to 0, we must have that $\left[1 - \frac{(\mu-y)^2}{\sigma^2}\right] = 0$ (the other quantities are strictly positive). The two solutions are $y = \mu + \sigma$ and $\mu - \sigma$.

4.80 Observe that $A = L*W = |Y|\times 3|Y| = 3Y^2$. Thus, $E(A) = 3E(Y^2) = 3(\sigma^2 + \mu^2)$.

4.81 **a.** $\Gamma(1) = \int_0^\infty e^{-y}\,dy = -e^{-y}\Big]_0^\infty = 1$.

b. $\Gamma(\alpha) = \int_0^\infty y^{\alpha-1}e^{-y}\,dy = \left[-y^{\alpha-1}e^{-y}\right]_0^\infty + \int_0^\infty (\alpha-1)y^{\alpha-2}e^{-y}\,dy = (\alpha-1)\Gamma(\alpha-1)$.

4.82 From above we have $\Gamma(1) = 1$, so that $\Gamma(2) = 1\Gamma(1) = 1$, $\Gamma(3) = 2\Gamma(2) = 2(1)$, and generally $\Gamma(n) = (n-1)\Gamma(n-1) = (n-1)!$ $\Gamma(4) = 3! = 6$ and $\Gamma(7) = 6! = 720$.

4.83 Applet Exercise — the results should agree.

4.84 **a.** The larger the value of α, the more symmetric the density curve.
b. The location of the distribution centers are increasing with α.
c. The means of the distributions are increasing with α.

4.85 **a.** These are all exponential densities.
b. Yes, they are all skewed densities (decaying exponential).
c. The spread is increasing with β.

4.86 **a.** $P(Y < 3.5) = .37412$
b. $P(W < 1.75) = P(Y/2 < 1.75) = P(Y < 3.5) = .37412$.
c. They are identical.

4.87 **a.** For the gamma distribution, $\phi_{.05} = .70369$.
b. For the χ^2 distribution, $\phi_{.05} = .35185$.
c. The .05–quantile for the χ^2 distribution is exactly one–half that of the .05–quantile for the gamma distribution. It is due to the relationship stated in Ex. 4.86.

4.88 Let Y have an exponential distribution with β = 2.4.

a. $P(Y > 3) = \int_3^\infty \frac{1}{2.4}e^{-y/2.4}dy = e^{-3/2.4} = .2865.$

b. $P(2 \le Y \le 3) = \int_2^3 \frac{1}{2.4}e^{-y/2.4}dy = .1481.$

4.89 **a.** Note that $\int_2^\infty \frac{1}{\beta}e^{-y/\beta}dy = e^{-2/\beta} = .0821$, so β = .8

b. $P(Y \le 1.7) = 1 - e^{-1.7/.8} = .5075$

4.90 Let Y = magnitude of the earthquake which is exponential with β = 2.4. Let X = # of earthquakes that exceed 5.0 on the Richter scale. Therefore, X is binomial with $n = 10$ and $p = P(Y > 5) = \int_5^\infty \frac{1}{2.4}e^{-y/2.4}dy = e^{-5/2.4} = .1245$. Finally, the probability of interest is

$$P(X \ge 1) = 1 - P(X = 0) = 1 - (.8755)^{10} = 1 - .2646 = .7354.$$

4.91 Let Y = water demand in the early afternoon. Then, Y is exponential with β = 100 cfs.

a. $P(Y > 200) = \int_{200}^\infty \frac{1}{100}e^{-y/100}dy = e^{-2} = .1353.$

b. We require the 99th percentile of the distribution of Y:

$P(Y > \phi_{.99}) = \int_{\phi_{.99}}^\infty \frac{1}{100}e^{-y/100}dy = e^{-\phi_{.99}/100} = .01$. So, $\phi_{.99} = -100\ln(.01) = 460.52$ cfs.

4.92 The random variable Y has an exponential distribution with $\beta = 10$. The cost C is related to Y by the formula $C = 100 + 40Y + 3Y^2$. Thus,

$E(C) = E(100 + 40Y + 3Y^2) = 100 + 40(10) + 3E(Y^2) = 100 + 400 + 3(100 + 10^2) = 1100.$

To find $V(C)$, note that $V(C) = E(C^2) - [E(C)]^2$. Therefore,

$E(C^2) = E[(100 + 40Y + 3Y^2)^2] = 10{,}000 + 2200E(Y^2) + 9E(Y^4) + 8000E(Y) + 240E(Y^3).$

$$E(Y) = 10 \qquad E(Y^2) = 200$$

$$E(Y^3) = \int_0^\infty y^3 \tfrac{1}{100} e^{-y/100} dy = \Gamma(4)100^3 = 6000.$$

$$E(Y^4) = \int_0^\infty y^4 \tfrac{1}{100} e^{-y/100} dy = \Gamma(5)100^4 = 240{,}000.$$

Thus, $E(C^2) = 10{,}000 + 2200(200) + 9(240{,}000) + 8000(10) + 240(6000) = 4{,}130{,}000.$

So, $V(C) = 4{,}130{,}000 - (1100)^2 = 2{,}920{,}000.$

4.93 Let Y = time between fatal airplane accidents. So, Y is exponential with $\beta = 44$ days.

a. $P(Y \le 31) = \int_0^{31} \tfrac{1}{44} e^{-y/44} dy = 1 - e^{-31/44} = .5057.$

b. $V(Y) = 44^2 = 1936.$

4.94 Let Y = CO concentration in air samples. So, Y is exponential with $\beta = 3.6$ ppm.

a. $P(Y > 9) = \int_9^\infty \tfrac{1}{3.6} e^{-y/3.6} dy = e^{-9/3.6} = .0821$

b. $P(Y > 9) = \int_9^\infty \tfrac{1}{2.5} e^{-y/3.6} dy = e^{-9/2.5} = .0273$

4.95 **a.** For any $k = 1, 2, 3, \ldots$

$$P(X = k) = P(k - 1 \le Y < k) = P(Y < k) - P(Y \le k - 1) = 1 - e^{-k/\beta} - (1 - e^{-(k-1)/\beta})$$
$$= e^{-(k-1)/\beta} - e^{-k/\beta}.$$

b. $P(X = k) = e^{-(k-1)/\beta} - e^{-k/\beta} = e^{-(k-1)/\beta} - e^{-(k-1)/\beta}(e^{1/\beta}) = e^{-(k-1)/\beta}(1 - e^{1/\beta}) = [e^{-1/\beta}]^{k-1}(1 - e^{1/\beta}).$

Thus, X has a geometric distribution with $p = 1 - e^{1/\beta}$.

4.96 **a.** The density function $f(y)$ is in the form of a gamma density with $\alpha = 4$ and $\beta = 2$. Thus, $k = \dfrac{1}{\Gamma(4)2^4} = \dfrac{1}{96}$.

b. Y has a χ^2 distribution with $\nu = 2(4) = 8$ degrees of freedom.

c. $E(Y) = 4(2) = 8$, $V(Y) = 4(2^2) = 16$.

d. Note that $\sigma = \sqrt{16} = 4$. Thus, $P(|Y - 8| < 2(4)) = P(0 < Y < 16) = .95762$.

4.97 $P(Y > 4) = \int_4^\infty \frac{1}{4}e^{-y/4}dy = e^{-1} = .3679$.

4.98 We require the 95th percentile of the distribution of Y:

$$P(Y > \phi_{.95}) = \int_{\phi_{.95}}^\infty \tfrac{1}{4}e^{-y/4}dy = e^{-\phi_{.95}/4} = .05. \text{ So, } \phi_{.95} = -4\ln(.05) = 11.98.$$

4.99 **a.** $P(Y > 1) = \sum_{y=0}^{1} \frac{e^{-1}}{y!} = e^{-1} + e^{-1} = .7358$.

b. The same answer is found.

4.100 **a.** $P(X_1 = 0) = e^{-\lambda_1}$ and $P(X_2 = 0) = e^{-\lambda_2}$. Since $\lambda_2 > \lambda_1$, $e^{-\lambda_2} < e^{-\lambda_1}$.

b. The result follows from Ex. 4.100.

c. Since distribution function is a nondecreasing function, it follows from part b that

$$P(X_1 \le k) = P(Y > \lambda_1) > P(Y > \lambda_2) = P(X_2 \le k)$$

d. We say that X_2 is "stochastically greater" than X_1.

4.101 Let Y have a gamma distribution with $\alpha = .8$, $\beta = 2.4$.

a. $E(Y) = (.8)(2.4) = 1.92$
b. $P(Y > 3) = .21036$
c. The probability found in Ex. 4.88 (a) is larger. There is greater variability with the exponential distribution.
d. $P(2 \le Y \le 3) = P(Y > 2) - P(Y > 3) = .33979 - .21036 = .12943$.

4.102 Let Y have a gamma distribution with $\alpha = 1.5$, $\beta = 3$.

a. $P(Y > 4) = .44592$.
b. We require the 95th percentile: $\phi_{.95} = 11.72209$.

4.103 Let R denote the radius of a crater. Therefore, R is exponential w/ $\beta = 10$ and the area is $A = \pi R^2$. Thus,

$$E(A) = E(\pi R^2) = \pi E(R^2) = \pi(100 + 100) = 200\pi.$$

$$V(A) = E(A^2) - [E(A)]^2 = \pi^2[E(R^4) - 200^2] = \pi^2[240{,}000 - 200^2] = 200{,}000\pi^2,$$

where $E(R^4) = \int_0^\infty \frac{1}{10} r^4 e^{-r/10} dr = 10^4\Gamma(5) = 240.000.$

4.104 Y has an exponential distribution with $\beta = 100$. Then, $P(Y > 200) = e^{-200/100} = e^{-2}$. Let the random variable X = # of componential that operate in the equipment for more than 200 hours. Then, X has a binomial distribution and

P(equipment operates) = $P(X \geq 2) = P(X = 2) + P(X = 3) = 3(e^{-2})^2(1 - e^{-2}) + (e^{-2})^3 = .05.$

4.105 Let the random variable Y = four–week summer rainfall totals

a. $E(Y) = 1.6(2) = 3.2$, $V(Y) = 1.6(2^2) = 6.4$

b. $P(Y > 4) = .28955.$

4.106 Let Y = response time. If $\mu = 4$ and $\sigma^2 = 8$, then it is clear that $\alpha = 2$ and $\beta = 2$.

a. $f(y) = \frac{1}{4} y e^{-y/2}$, $y > 0$.

b. $P(Y < 5) = 1 - .2873 = .7127.$

4.107 **a.** Using Tchebysheff's theorem, two standard deviations about the mean is given by $4 \pm 2\sqrt{8} = 4 \pm 5.657$ or $(-1.657, 9.657)$, or simply $(0, 9.657)$ since Y must be positive.

b. $P(Y < 9.657) = 1 - .04662 = 0.95338.$

4.108 Let Y = annual income. Then, Y has a gamma distribution with $\alpha = 20$ and $\beta = 1000$.

a. $E(Y) = 20(1000) = 20{,}000$, $V(Y) = 20(1000)^2 = 20{,}000{,}000.$

b. The standard deviation $\sigma = \sqrt{20{,}000{,}000} = 4472.14$. The value 30,000 is $\frac{30{,}000-20{,}000}{4472.14}$ = 2.236 standard deviations above the mean. This represents a fairly extreme value.

c. $P(Y > 30{,}000) = .02187$

4.109 Let Y have a gamma distribution with $\alpha = 3$ and $\beta = 2$. Then, the loss $L = 30Y + 2Y^2$. Then,

$$E(L) = E(30Y + 2Y^2) = 30E(Y) + 2E(Y^2) = 30(6) + 2(12 + 6^2) = 276,$$
$$V(L) = E(L^2) - [E(L)]^2 = E(900Y^2 + 120Y^3 + 4Y^4) - 276^2.$$

$$E(Y^3) = \int_0^\infty \frac{y^5}{16} e^{-y/2} = 480 \qquad E(Y^4) = \int_0^\infty \frac{y^6}{16} e^{-y/2} = 5760$$

Thus, $V(L) = 900(48) + 120(480) + 4(5760) - 276^2 = 47{,}664.$

4.110 Y has a gamma distribution with α = 3 and β = .5. Thus, $E(Y) = 1.5$ and $V(Y) = .75$.

4.111 **a.** $E(Y^a) = \int_0^\infty y^a \frac{1}{\Gamma(\alpha)\beta^\alpha} y^{\alpha-1} e^{-y/\beta} dy = \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^\infty y^{a+\alpha-1} e^{-y/\beta} dy = \frac{1}{\Gamma(\alpha)\beta^\alpha} \frac{\Gamma(a+\alpha)\beta^{a+\alpha}}{1} = \frac{\beta^a \Gamma(a+\alpha)}{\Gamma(\alpha)}$.

b. For the gamma function Γ(t), we require $t > 0$.

c. $E(Y^1) = \frac{\beta^1\Gamma(1+\alpha)}{\Gamma(\alpha)} = \frac{\beta\alpha\Gamma(\alpha)}{\Gamma(\alpha)} = \alpha\beta.$

d. $E(\sqrt{Y}) = E(Y^{.5}) = \frac{\beta^{.5}\Gamma(.5+\alpha)}{\Gamma(\alpha)}$, α > 0.

e. $E(1/Y) = E(Y^{-1}) = \frac{\beta^{-1}\Gamma(-1+\alpha)}{\Gamma(\alpha)} = \frac{1}{\beta(\alpha-1)}$, α > 1.

$E(1/\sqrt{Y}) = E(Y^{-.5}) = \frac{\beta^{-.5}\Gamma(-.5+\alpha)}{\Gamma(\alpha)} = \frac{\Gamma(\alpha-.5)}{\sqrt{\beta}\Gamma(\alpha)}$, α > .5.

$E(1/Y^2) = E(Y^{-2}) = \frac{\beta^{-2}\Gamma(-2+\alpha)}{\Gamma(\alpha)} = \frac{\beta^{-2}\Gamma(\alpha-2)}{(\alpha-1)(\alpha-2)\Gamma(\alpha-2)} \frac{1}{\beta^2(\alpha-1)(\alpha-2)}$, α > 2.

4.112 The chi–square distribution with ν degrees of freedom is the same as a gamma distribution with α = ν/2 and β = 2.

a. From Ex. 4.111, $E(Y^a) = \frac{2^a\Gamma(a+\frac{\nu}{2})}{\Gamma(\frac{\nu}{2})}$.

b. As in Ex. 4.111 with α + a > 0 and α = ν/2, it must hold that ν > –2a

c. $E(\sqrt{Y}) = E(Y^{.5}) = \frac{\sqrt{2}\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})}$, ν > 0.

d. $E(1/Y) = E(Y^{-1}) = \frac{2^{-1}\Gamma(-1+\frac{\nu}{2})}{\Gamma(\frac{\nu}{2})} = \frac{1}{\nu-2}$, ν > 2.

$E(1/\sqrt{Y}) = E(Y^{-.5}) = \frac{\Gamma(\frac{\nu-1}{2})}{\sqrt{2}\Gamma(\frac{\nu}{2})}$, ν > 1.

$E(1/Y^2) = E(Y^{-2}) = \frac{1}{2^2(\frac{\nu}{2}-1)(\frac{\nu}{2}-2)} = \frac{1}{(\nu-2)(\nu-4)}$, α > 4.

4.113 Applet exercise.

4.114 **a.** This is the (standard) uniform distribution.
b. The beta density with α = 1, β = 1 is symmetric.
c. The beta density with α = 1, β = 2 is skewed right.
d. The beta density with α = 2, β = 1 is skewed left.
e. Yes.

4.115 **a.** The means of all three distributions are .5.
b. They are all symmetric.
c. The spread decreases with larger (and equal) values of α and β.
d. The standard deviations are .2236, .1900, and .1147 respectively. The standard deviations are decreasing which agrees with the density plots.
e. They are always symmetric when α = β.

4.116 **a.** All of the densities are skewed right.
b. The density obtains a more symmetric appearance.
c. They are always skewed right when $\alpha < \beta$ and $\alpha > 1$ and $\beta > 1$.

4.117 **a.** All of the densities are skewed left.
b. The density obtains a more symmetric appearance.
c. They are always skewed right when $\alpha > \beta$ and $\alpha > 1$ and $\beta > 1$.

4.118 **a.** All of the densities are skewed right (similar to an exponential shape).
b. The spread decreases as the value of β gets closer to 12.
c. The distribution with $\alpha = .3$ and $\beta = 4$ has the highest probability.
d. The shapes are all similar.

4.119 **a.** All of the densities are skewed left (a mirror image of those from Ex. 4.118).
b. The spread decreases as the value of α gets closer to 12.
c. The distribution with $\alpha = 4$ and $\beta = .3$ has the highest probability.
d. The shapes are all similar.

4.120 Yes, the mapping explains the mirror image.

4.121 **a.** These distributions exhibit a "U" shape.
b. The area beneath the curve is greater closer to "1" than "0".

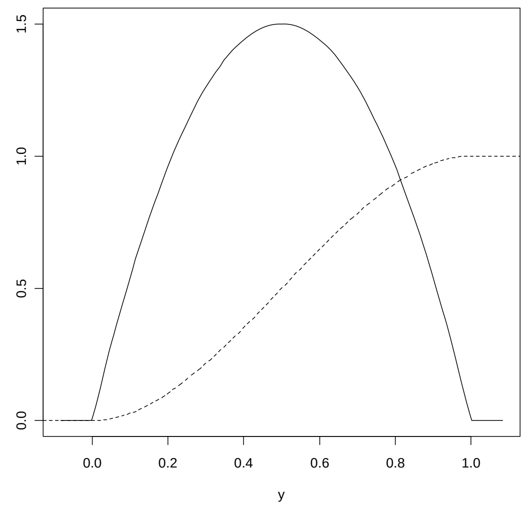
4.122 **a.** $P(Y > .1) = .13418$
b. $P(Y < .1) = 1 - .13418 = .86582$.
c. Values smaller than .1 have greater probability.
d. $P(Y < .1) = 1 - .45176 = .54824$
e. $P(Y > .9) = .21951$.
f. $P(0.1 < Y < 0.9) = 1 - .54824 - .21951 = .23225$.
g. Values of $Y < .1$ have the greatest probability.

4.123 **a.** The random variable Y follows the beta distribution with $\alpha = 4$ and $\beta = 3$, so the constant $k = \frac{\Gamma(4+3)}{\Gamma(4)\Gamma(3)} = \frac{6!}{3!2!} = 60$.
b. We require the 95th percentile of this distribution, so it is found that $\phi_{.95} = 0.84684$.

4.124 **a.** $P(Y > .4) = \int_{.4}^{1}(12y^2 - 12y^3)dy = \left[4y^3 - 3y^4\right]_{.4}^{1} = .8208$.
b. $P(Y > .4) = .82080$.

4.125 From Ex. 4.124 and using the formulas for the mean and variance of beta random variables, $E(Y) = 3/5$ and $V(Y) = 1/25$.

4.126 **a.** $F(y) = \int_0^y (6t - 6t^2)dt = 3y^2 - 2y^3$, $0 \le y \le 1$. $F(y) = 0$ for $y < 0$ and $F(y) = 1$ for $y > 1$.



b. Solid line: $f(y)$; dashed line: $F(y)$

c. $P(.5 \le Y \le .8) = F(.8) - F(.5) = 1.92 - 1.092 - .75 + .25 = .396$.

4.127 For $\alpha = \beta = 1$, $f(y) = \frac{\Gamma(2)}{\Gamma(1)\Gamma(1)} y^{1-1}(1-y)^{1-1} = 1,\ 0 \le y \le 1$, which is the uniform distribution.

4.128 The random variable Y = weekly repair cost (in hundreds of dollars) has a beta distribution with $\alpha = 1$ and $\beta = 3$. We require the 90$^{\text{th}}$ percentile of this distribution:

$$P(Y > \phi_{.9}) = \int_{\phi_{.9}}^{1} 3(1-y)^2\,dy = (1-\phi_{.9})^3 = .1.$$

Therefore, $\phi_{.9} = 1 - (.1)^{1/3} = .5358$. So, the budgeted cost should be \$53.58.

4.129 $E(C) = 10 + 20E(Y) + 4E(Y^2) = 10 + 20\left(\frac{1}{3}\right) + 4\left(\frac{2}{9*4} + \frac{1}{9}\right) = \frac{52}{3}$

$V(C) = E(C^2) - [E(C)]^2 = E[(10 + 20Y + 4Y^2)^2] - \left(\frac{52}{3}\right)^2$

$E[(10 + 20Y + 4Y^2)^2] = 100 + 400E(Y) + 480E(Y^2) + 160E(Y^3) + 16E(Y^4)$

Using mathematical expectation, $E(Y^3) = \frac{1}{10}$ and $E(Y^4) = \frac{1}{15}$. So,

$V(C) = E(C^2) - [E(C)]^2 = (100 + 400/3 + 480/6 + 160/10 + 16/15) - (52/3)^2 = 29.96.$

4.130 To find the variance $\sigma^2 = E(Y^2) - \mu^2$:

$$E(Y^2) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\int_0^1 y^{\alpha+1}(1-y)^{\beta-1}\,dy = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\frac{\Gamma(\alpha+2)\Gamma(\beta)}{\Gamma(\alpha+2+\beta)} = \frac{(\alpha+1)\alpha}{(\alpha+\beta)(\alpha+\beta+1)}$$

$$\sigma^2 = \frac{(\alpha+1)\alpha}{(\alpha+\beta)(\alpha+\beta+1)} - \left(\frac{\alpha}{\alpha+\beta}\right)^2 = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}.$$

4.131 This is the same beta distribution used in Ex. 4.129.

a. $P(Y < .5) = \int_0^{.5} 2(1-y)dy = 2y - y^2\Big]_0^{.5} = .75$

b. $E(Y) = 1/3$, $V(Y) = 1/18$, so $\sigma = 1/\sqrt{18} = .2357$.

4.132 Let Y = proportion of weight contributed by the fine powders

a. $E(Y) = .5$, $V(Y) = 9/(36*7) = 1/28$

b. $E(Y) = .5$, $V(Y) = 4/(16*5) = 1/20$

c. $E(Y) = .5$, $V(Y) = 1/(4*3) = 1/12$

d. Case (a) will yield the most homogenous blend since the variance is the smallest.

4.133 The random variable Y has a beta distribution with $\alpha = 3$, $\beta = 5$.

a. The constant $c = \frac{\Gamma(3+5)}{\Gamma(3)\Gamma(5)} = \frac{7!}{2!4!} = 105$.

b. $E(Y) = 3/8$.

c. $V(Y) = 15/(64*9) = 5/192$, so $\sigma = .1614$.

d. $P(Y > .375 + 2(.1614)) = P(Y > .6978) = .02972$.

4.134 **a.** If $\alpha = 4$ and $\beta = 7$, then we must find

$P(Y \le .7) = F(.7) = \sum_{i=4}^{10}\binom{10}{i}(.7)^i(.3)^{10-i}$ = $P(4 \le X \le 10)$, for the random variable X distributed as binomial with $n = 10$ and $p = .7$. Using Table I in Appendix III, this is .989.

b. Similarly, $F(.6) = P(12 \le X \le 25)$, for the random variable X distributed as binomial with $n = 25$ and $p = .6$. Using Table I in Appendix III, this is .922.

c. Similar answers are found.

4.135 **a.** $P(Y_1 = 0) = (1 - p_1)^n > P(Y_2 = 0) = (1 - p_2)^n$, since $p_1 < p_2$.

b. $P(Y_1 \le k) = 1 - P(Y_1 \ge k + 1) = 1 - \sum_{i=k+1}^{n}\binom{n}{i}p_1^i(1-p_1)^{n-i} = 1 - \int_0^{p_1}\frac{t^k(1-t)^{n-k-1}}{B(k+1,n-k)}$

$= 1 - P(X \le p_1) = P(X > p_1)$, where is X beta with parameters $k + 1$, $n - k$.

c. From part b, we see the integrands for $P(Y_1 \le k)$ and $P(Y_2 \le k)$ are identical but since $p_1 < p_2$, the regions of integration are different. So, Y_2 is "stochastically greater" than Y_1.

4.136 **a.** Observing that the exponential distribution is a special case of the gamma distribution, we can use the gamma moment–generating function with $\alpha = 1$ and $\beta = \theta$:

$$m(t) = \frac{1}{1-\theta t},\ t < 1/\theta.$$

b. The first two moments are found by $m'(t) = \frac{\theta}{(1-\theta t)^2}$, $E(Y) = m'(0) = \theta$.

$m''(t) = \frac{2\theta}{(1-\theta t)^3}$, $E(Y^2) = m''(0) = 2\theta^2$. So, $V(Y) = 2\theta^2 - \theta^2 = \theta^2$.

4.137 The mgf for U is $m_U(t) = E(e^{tU}) = E(e^{t(aY+b)}) = E(e^{bt}e^{(at)Y}) = e^{bt}m(at)$. Thus,

$$m'_U(t) = be^{bt}m(at) + ae^{bt}m'(at).\ \text{So, } m'_U(0) = b + am'(0) = b + a\mu = E(U).$$

$$m''_U(t) = b^2e^{bt}m(at) + abe^{bt}m'(at) + abe^{bt}m'(at) + a^2e^{bt}m''(at),\ \text{so}$$
$$m''_U(0) = b^2 + 2ab\mu + a^2E(Y^2) = E(U^2).$$

Therefore, $V(U) = b^2 + 2ab\mu + a^2E(Y^2) - (b + a\mu)^2 = a^2[E(Y^2) - \mu^2] = a^2\sigma^2$.

4.138 **a.** For $U = Y - \mu$, the mgf $m_U(t)$ is given in Example 4.16. To find the mgf for $Y = U + \mu$, use the result in Ex. 4.137 with $a = 1$, $b = -\mu$:

$$m_Y(t) = e^{-\mu t}m_U(t) = e^{\mu t + \sigma^2 t^2/2}$$

b. $m'_Y(t) = (\mu + t\sigma^2)e^{\mu t + \sigma^2 t^2/2}$, so $m'_Y(0) = \mu$
$m''_Y(t) = (\mu + t\sigma^2)^2 e^{\mu t + \sigma^2 t^2/2} + \sigma^2 e^{\mu t + \sigma^2 t^2/2}$, so $m''_Y(0) = \mu^2 + \sigma^2$. Finally, $V(Y) = \sigma^2$.

4.139 Using Ex. 4.137 with $a = -3$ and $b = 4$, it is trivial to see that the mgf for X is

$$m_X(t) = e^{4t}m(-3t) = e^{(4-3\mu)t + 9\sigma^2 t^2/2}.$$

By the uniqueness of mgfs, X is normal with mean $4 - 3\mu$ and variance $9\sigma^2$.

4.140 **a.** Gamma with $\alpha = 2$, $\beta = 4$
b. Exponential with $\beta = 3.2$
c. Normal with $\mu = -5$, $\sigma^2 = 12$

4.141 $m(t) = E(e^{tY}) = \int_{\theta_1}^{\theta_2} \frac{e^{ty}}{\theta_2 - \theta_1}dy = \frac{e^{t\theta_2} - e^{t\theta_1}}{t(\theta_2 - \theta_1)}$.

4.142 **a.** $m_Y(t) = \frac{e^t - 1}{t}$

b. From the cited exercises, $m_W(t) = \frac{e^{at} - 1}{at}$. From the uniqueness property of mgfs, W is uniform on the interval (0, a).

c. The mgf for X is $m_X(t) = \frac{e^{-at} - 1}{-at}$, which implies that X is uniform on the interval (–a, 0).

d. The mgf for V is $m_V(t) = e^{bt}\frac{e^{at} - 1}{at} = \frac{e^{(b+a)t} - e^{bt}}{at}$, which implies that V is uniform on the interval (b, b + a).

4.143 The mgf for the gamma distribution is $m(t) = (1 - \beta t)^{-\alpha}$. Thus,

$$m'(t) = \alpha\beta(1 - \beta t)^{-\alpha - 1}, \text{ so } m'(0) = \alpha\beta = E(Y)$$

$$m''(t) = (\alpha + 1)\alpha\beta^2(1 - \beta t)^{-\alpha - 2}, \text{ so } m''(0) = (\alpha + 1)\alpha\beta^2 = E(Y^2). \text{ So,}$$

$$V(Y) = (\alpha + 1)\alpha\beta^2 - (\alpha\beta)^2 = \alpha\beta^2.$$

4.144 **a.** The density shown is a normal distribution with μ = 0 and σ^2 = 1. Thus, $k = 1/\sqrt{2\pi}$.

b. From Ex. 4.138, the mgf is $m(t) = e^{t^2/2}$.

c. $E(Y)$ = 0 and $V(Y)$ = 1.

4.145 **a.** $E(e^{3T/2}) = \int_{-\aleph}^{0} e^{3y/2}e^y dy = \frac{2}{5}e^{5y/2}\Big]_{-\infty}^{0} = \frac{2}{5}$.

b. $m(t) = E(e^{tY}) = \int_{-\infty}^{0} e^{ty}e^y dy = \frac{1}{t+1},\ t > -1$.

c. By using the methods with mgfs, $E(Y)$ = –1, $E(Y^2)$ = 2, so $V(Y)$ = 2 – $(–1)^2$ = 1.

4.146 We require $P(|Y - \mu| \le k\sigma) \ge .90 = 1 - 1/k^2$. Solving for k, we see that k = 3.1622. Thus, the necessary interval is $|Y - 25{,}000| \le (3.1622)(4000)$ or $12{,}351 \le 37{,}649$.

4.147 We require $P(|Y - \mu| \le .1) \ge .75 = 1 - 1/k^2$. Thus, k = 2. Using Tchebysheff's inequality, 1 = $k\sigma$ and so σ = 1/2.

4.148 In Exercise 4.16, μ = 2/3 and σ = $\sqrt{2/9}$ = .4714. Thus,

$$P(|Y - \mu| \le 2\sigma) = P(|Y - 2/3| \le .9428) = P(-.2761 \le Y \le 1.609) = F(1.609) = .962.$$

Note that the negative portion of the interval in the probability statement is irrelevant since Y is non–negative. According to Tchsebysheff's inequality, the probability is at least 75%. The empirical rule states that the probability is approximately 95%. The above probability is closest to the empirical rule, even though the density function is far from mound shaped.

4.149 For the uniform distribution on (θ_1, θ_2), $\mu = \frac{\theta_1+\theta_2}{2}$ and $\sigma^2 = \frac{(\theta_2-\theta_1)^2}{12}$. Thus,

$$2\sigma = \frac{(\theta_2-\theta_1)}{\sqrt{3}}.$$

The probability of interest is

$$P(|Y-\mu| \le 2\sigma) = P(\mu - 2\sigma \le Y \le \mu + 2\sigma) = P\left(\frac{\theta_1+\theta_2}{2} - \frac{(\theta_2-\theta_1)}{\sqrt{3}} \le Y \le \frac{\theta_1+\theta_2}{2} + \frac{(\theta_2-\theta_1)}{\sqrt{3}}\right)$$

It is not difficult to show that the range in the last probability statement is greater than the actual interval that Y is restricted to, so

$$P\left(\frac{\theta_1+\theta_2}{2} - \frac{(\theta_2-\theta_1)}{\sqrt{3}} \le Y \le \frac{\theta_1+\theta_2}{2} + \frac{(\theta_2-\theta_1)}{\sqrt{3}}\right) = P(\theta_1 \le Y \le \theta_2) = 1.$$

Note that Tchebysheff's theorem is satisfied, but the probability is greater than what is given by the empirical rule. The uniform is not a mound–shaped distribution.

4.150 For the exponential distribution, $\mu = \beta$ and $\sigma^2 = \beta^2$. Thus, $2\sigma = 2\beta$. The probability of interest is

$$P(|Y-\mu| \le 2\sigma) = P(\mu - 2\sigma \le Y \le \mu + 2\sigma) = P(-\beta \le Y \le 3\beta) = P(0 \le Y \le 3\beta)$$

This is simply $F(3\beta) = 1 - e^{-3\beta} = .9502$. The empirical rule and Tchebysheff's theorem are both valid.

4.151 From Exercise 4.92, $E(C) = 1000$ and $V(C) = 2{,}920{,}000$ so that the standard deviation is $\sqrt{2{,}920{,}000} = 1708.80$. The value 2000 is only $(2000 - 1100)/1708.8 = .53$ standard deviations above the mean. Thus, we would expect C to exceed 2000 fair often.

4.152 We require $P(|L-\mu| \le k\sigma) \ge .89 = 1 - 1/k^2$. Solving for k, we have $k = 3.015$. From Ex. 4.109, $\mu = 276$ and $\sigma = 218.32$. The interval is

$$|L - 276| \le 3.015(218.32) \text{ or } (-382.23, 934.23)$$

Since L must be positive, the interval is (0, 934.23)

4.153 From Ex. 4.129, it is shown that $E(C) = \frac{52}{3}$ and $V(C) = 29.96$, so, the standard deviation is $\sqrt{29.96} = 5.474$. Thus, using Tchebysheff's theorem with $k = 2$, the interval is

$$|Y - \tfrac{52}{3}| \le 2(5.474) \text{ or } (6.38, 28.28)$$

4.154 **a.** $\mu = 7$, $\sigma^2 = 2(7) = 14$.
b. Note that $\sigma = \sqrt{14} = 3.742$. The value 23 is $(23 - 7)/3.742 = 4.276$ standard deviations above the mean, which is unlikely.
c. With $\alpha = 3.5$ and $\beta = 2$, $P(Y > 23) = .00170$.

4.155 The random variable Y is uniform over the interval (1, 4). Thus, $f(y) = \frac{1}{3}$ for $1 \le y \le 4$ and $f(y) = 0$ elsewhere. The random variable C = cost of the delay is given as

$$C = g(Y) = \begin{cases} 100 & 1 \le y \le 2 \\ 100 + 20(Y-2) & 2 < y \le 4 \end{cases}$$

Thus, $E(C) = E[g(Y)] = \int_{-\infty}^{\infty} g(y)f(y)dy = \int_1^2 \frac{100}{3}dy + \int_2^4 [100 + 20(y-2)]\frac{1}{3}dy$ = \$113.33.

4.156 Note that Y is a discrete random variable with probability .2 + .1 = .3 and it is continuous with probability 1 – .3 = .7. Hence, by using Definition 4.15, we can write Y as a mixture of two random variables X_1 and X_2. The random variable X_1 is discrete and can assume two values with probabilities $P(X_1 = 3) = .2/.3 = 2/3$ and $P(X_1 = 6) = .1/.3 = 1/3$. Thus, $E(X_1) = 3(2/3) + 6(1/3) = 4$. The random variable X_2 is continuous and follows a gamma distribution (as given in the problem) so that $E(X_2) = 2(2) = 4$. Therefore,

$$E(Y) = .3(4) + .7(4) = 4.$$

4.157 **a.** The distribution function for X is $F(x) = \begin{cases} 0 & x < 0 \\ \int_0^x \frac{1}{100}e^{-t/100}dx = 1 - e^{-x/100} & 0 \le x < 200 \\ 1 & x \ge 200 \end{cases}$.

b. $E(X) = \int_0^{200} x\frac{1}{100}e^{-x/100}dx + .1353(200) = 86.47$, where $.1353 = P(Y > 200)$.

4.158 The distribution for V is gamma with $\alpha = 4$ and $\beta = 500$. Since there is one discrete point at 0 with probability .02, using Definition 4.15 we have that $c_1 = .02$ and $c_2 = .98$. Denoting the kinetic energy as $K = \frac{m}{2}V^2$ we can solve for the expected value:

$$E(K) = (.98)\tfrac{m}{2}E(V^2) = (.98)\tfrac{m}{2}\{V(V) + [E(V)]^2\} = (.98)\tfrac{m}{2}\{4(500)^2 + 2000^2\} = 2{,}450{,}000m.$$

4.159 **a.** The distribution function has jumps at two points: $y = 0$ (of size .1) and $y = .5$ (of size .15). So, the discrete component of $F(y)$ is given by

$$F_1(y) = \begin{cases} 0 & y < 0 \\ \frac{.1}{.1+.15} = .4 & 0 \le y < 5 \\ 1 & y \ge .5 \end{cases}$$

The continuous component of $F(y)$ can then by determined:

$$F_2(y) = \begin{cases} 0 & y < 0 \\ 4y^2/3 & 0 \le y < .5 \\ (4y-1)/3 & .5 \le y < 1 \\ 1 & y \ge 1 \end{cases}$$

b. Note that $c_1 = .1 + .15 = .25$. So, $F(y) = 0.25F_1(y) + 0.75F_2(y)$.

c. First, observe that $f_2(y) = F_2'(y) = \begin{cases} 8y/3 & 0 \le y < .5 \\ 4/3 & y \ge .5 \end{cases}$. Thus,

$E(Y) = .25(.6)(.5) + \int_0^{.5} 8y^2/3dy + \int_{.5}^{1} 4y/3dy = .533$. Similarly, $E(Y^2) = .3604$ so that $V(Y) = .076$.

4.160 a. $F(y) = \int_{-1}^{y} \frac{2}{\pi(1+y^2)} dy = \frac{2}{\pi}\tan^{-1}(y) + \frac{1}{2}, -1 \le y \le 1, F(y) = 0 \text{ if } y < 0, F(y) = 1 \text{ if } y > 1$.

b. Find $E(Y)$ directly using mathematical expectation, or observe that $f(y)$ is symmetric about 0 so using the result from Ex. 4.27, $E(Y) = 0$.

4.161 Here, μ = 70 and σ = 12 with the normal distribution. We require $\phi_{.9}$, the 90th percentile of the distribution of test times. Since for the standard normal distribution, $P(Z < z_0) = .9$ for $z_0 = 1.28$, thus

$$\phi_{.9} = 70 + 12(1.28) = 85.36.$$

4.162 Here, μ = 500 and σ = 50 with the normal distribution. We require $\phi_{.01}$, the 1st percentile of the distribution of light bulb lives. For the standard normal distribution, $P(Z < z_0) = .01$ for $z_0 = –2.33$, thus

$$\phi_{.01} = 500 + 50(-2.33) = 383.5$$

4.163 Referring to Ex. 4.66, let X = # of defective bearings. Thus, X is binomial with $n = 5$ and $p = P(\text{defective}) = .073$. Thus,

$$P(X > 1) = 1 - P(X = 0) = 1 - (.927)^5 = .3155.$$

4.164 Let Y = lifetime of a drill bit. Then, Y has a normal distribution with μ = 75 hours and σ = 12 hours.

a. $P(Y < 60) = P(Z < -1.25) = .1056$

b. $P(Y \ge 60) = 1 - P(Y < 60) = 1 - .1056 = .8944$.

c. $P(Y > 90) = P(Z > 1.25) = .1056$

4.165 The density function for Y is in the form of a gamma density with α = 2 and β = .5.

a. $c = \frac{1}{\Gamma(2)(.5)^2} = 4$.

b. $E(Y) = 2(.5) = 1$, $V(Y) = 2(.5)^2 = .5$.

c. $m(t) = \frac{1}{(1-.5t)^2}$, $t < 2$.

4.166 In Example 4.16, the mgf is $m(t) = e^{t^2\sigma^2/2}$. The infinite series expansion of this is

$$m(t) = 1 + \left(\tfrac{t^2\sigma^2}{2}\right) + \left(\tfrac{t^2\sigma^2}{2}\right)^2 \tfrac{1}{2!} + \left(\tfrac{t^2\sigma^2}{2}\right)^3 \tfrac{1}{3!} + \cdots = 1 + \tfrac{t^2\sigma^2}{2} + \tfrac{t^4\sigma^4}{8} + \tfrac{t^6\sigma^6}{48} + \cdots$$

Then, μ_1 = coefficient of t, so $\mu_1 = 0$
μ_2 = coefficient of $t^2/2!$, so $\mu_2 = \sigma^2$
μ_3 = coefficient of $t^3/3!$, so $\mu_3 = 0$
μ_4 = coefficient of $t^4/4!$, so $\mu_4 = 3\sigma^4$

4.167 For the beta distribution,

$$E(Y^k) = \int_0^1 y^k \tfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} y^{\alpha-1}(1-y)^{\beta-1}\,dy = \tfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 y^{k+\alpha-1}(1-y)^{\beta-1}\,dy = \tfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \tfrac{\Gamma(k+\alpha)\Gamma(\beta)}{\Gamma(k+\alpha+\beta)}.$$

Thus, $E(Y^k) = \frac{\Gamma(\alpha+\beta)\Gamma(k+\alpha)}{\Gamma(\alpha)\Gamma(k+\alpha+\beta)}$.

4.168 Let T = length of time until the first arrival. Thus, the distribution function for T is given by

$$F(t) = P(T \le t) = 1 - P(T > t) = 1 - P[\text{no arrivals in } (0, t)] = 1 - P[N = 0 \text{ in } (0, t)]$$

The probability $P[N = 0$ in $(0, t)]$ is given by $\frac{(\lambda t)^0 e^{-\lambda t}}{0!} = e^{-\lambda t}$. Thus, $F(t) = 1 - e^{-\lambda t}$ and

$$f(t) = \lambda e^{-\lambda t},\ t > 0.$$

This is the exponential distribution with $\beta = 1/\lambda$.

4.169 Let Y = time between the arrival of two call, measured in hours. To find $P(Y > .25)$, note that $\lambda t = 10$ and $t = 1$. So, the density function for Y is given by $f(y) = 10e^{-10y}$, $y > 0$. Thus,

$$P(Y > .25) = e^{-10(.25)} = e^{-2.5} = .082.$$

4.170 a. Similar to Ex. 4.168, the second arrival will occur after time t if either one arrival has occurred in (0, t) or no arrivals have occurred in $(0, t)$. Thus:

$P(U > t) = P[\text{one arrival in } (0, t)] + P[\text{no arrivals in } (0, t)] = \frac{(\lambda t)^0 e^{-\lambda t}}{0!} + \frac{(\lambda t)^1 e^{-\lambda t}}{1!}$. So,

$$F(t) = 1 - P(U > t) = 1 - \tfrac{(\lambda t)^0 e^{-\lambda t}}{0!} + \tfrac{(\lambda t)^1 e^{-\lambda t}}{1!} = 1 - (\lambda t + 1)e^{-\lambda t}.$$

The density function is given by $f(t) = F'(t) = \lambda^2 t e^{-\lambda t}$, $t > 0$. This is a gamma density with $\alpha = 2$ and $\beta = 1/\lambda$.

b. Similar to part a, but let X = time until the k^{th} arrival. Thus, $P(X > t) = \sum_{n=0}^{k-1} \frac{(\lambda t)^n e^{-\lambda t}}{n!}$. So,

$$F(t) = 1 - \sum_{n=0}^{k-1} \frac{(\lambda t)^n e^{-\lambda t}}{n!}.$$

The density function is given by

$$f(t) = F'(t) = -\left[-\lambda e^{-\lambda t}\sum_{n=0}^{k-1}\frac{(\lambda t)^n}{n!} + e^{-\lambda t}\sum_{n=1}^{k-1}\frac{\lambda^n t^{n-1}}{(n-1)!}\right] = \lambda e^{-\lambda t}\left[\sum_{n=0}^{k-1}\frac{(\lambda t)^n}{n!} - \sum_{n=1}^{k-1}\frac{(\lambda t)^{n-1}}{(n-1)!}\right].$$ Or,

$f(t) = \frac{\lambda^k t^{k-1}e^{-\lambda t}}{(k-1)!}$, $t > 0$. This is a gamma density with $\alpha = k$ and $\beta = 1/\lambda$.

4.171 From Ex. 4.169, W = waiting time follow an exponential distribution with $\beta = 1/2$.

a. $E(W) = 1/2$, $V(W) = 1/4$.

b. P(at least one more customer waiting) = 1 – P(no customers waiting in three minutes) $= 1 - e^{-6}$.

4.172 Twenty seconds is 1/5 a minute. The random variable Y = time between calls follows an exponential distribution with $\beta = .25$. Thus:

$$P(Y > 1/5) = \int_{1/5}^{\infty} 4e^{-4y}\,dy = e^{-4/5}.$$

4.173 Let R = distance to the nearest neighbor. Then,

$$P(R > r) = P(\text{no plants in a circle of radius } r)$$

Since the number of plants in a area of one unit has a Poisson distribution with mean λ, the number of plants in a area of πr^2 units has a Poisson distribution with mean $\lambda\pi r^2$. Thus,

$$F(r) = P(R \le r) = 1 - P(R > r) = 1 - e^{-\lambda\pi r^2}.$$

So, $f(r) = F'(t) = 2\lambda\pi r e^{-\lambda\pi r^2}$, $r > 0$.

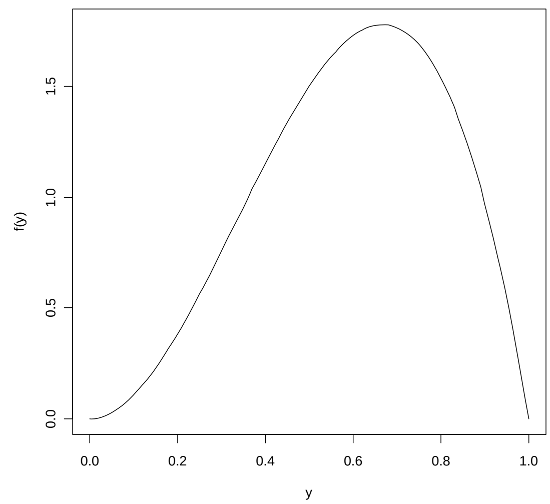
4.174 Let Y = interview time (in hours). The second applicant will have to wait only if the time to interview the first applicant exceeds 15 minutes, or .25 hour. So,

$$P(Y > .25) = \int_{.25}^{\infty} 2e^{-2y}\,dy = e^{-.5} = .61.$$

4.175 From Ex. 4.11, the median value will satisfy $F(y) = y^2/2 = .5$. Thus, the median is given by $\sqrt{2} = 1.414$.

4.176 The distribution function for the exponential distribution with mean β is $F(y) = 1 - e^{-y/\beta}$. Thus, we require the value y such that $F(y) = 1 - e^{-y/\beta} = .5$. Solving for y, this is $\beta\ln(2)$.

4.177 **a.** 2.07944 = 3ln(2)
b. 3.35669 < 4, the mean of the gamma distribution.
c. 46.70909 < 50, the mean of the gamma distribution.
d. In all cases the median is *less* than the mean, indicating right skewed distributions.



4.178 The graph of this beta density is above.

a. Using the relationship with binomial probabilities,
$P(.1 \le Y \le .2) = 4(.2)^3(.8) + (.2)^4 - 4(.1)^3(.9) - (.1)^4 = .0235$.
b. $P(.1 \le Y \le .2) = .9963 - .9728 = .0235$, which is the same answer as above.
c. $\phi_{.05} = .24860$, $\phi_{.95} = .90239$.
d. $P(\phi_{.05} \le Y \le \phi_{.95}) = .9$.

4.179 Let X represent the grocer's profit. In general, her profit (in cents) on a order of 100k pounds of food will be $X = 1000Y - 600k$ as long as $Y < k$. But, when $Y \ge k$ the grocer's profit will be $X = 1000k - 600k = 400k$. Define the random variable Y' as

$$Y' = \begin{cases} Y & 0 \le Y < k \\ k & Y \ge k \end{cases}$$

Then, we can write $g(Y') = X = 1000Y' + 600k$. The random variable Y' has a mixed distribution with one discrete point at k. Therefore,

$$c_1 = P(Y' = k) = P(Y \ge k) = \int_k^1 3y^2 dy = 1 - k^3, \text{ so that } c_2 = k^3.$$

Thus, $F_2(y) = \begin{cases} 0 & 0 \le y < k \\ 1 & y \ge k \end{cases}$ and $F_1(y) = P(Y_2 \le y \mid 0 \le Y' < k) = \dfrac{\int_0^y 3t^2 dt}{k^3} = \frac{y^3}{k^3}$, $0 \le y < k$.

Thus, from Definition 4.15,

$$E(X) = E[g(Y')] = c_1 E[g(Y_1)] + c_2 E[g(Y_2)] = (1 - k^3)400k + k^3 \int_0^k (1000y - 600k)\tfrac{3y^2}{k^3} dy,$$

or $E(X) = 400k - 250k^2$. This is maximized at $k = (.4)^{1/3} = .7368$. (2nd derivative is –.)

4.180 **a.** Using the result of Ex. 4.99, $P(Y \le 4) = 1 - \sum_{y=0}^{2} \frac{4^y e^{-4}}{y!} = .7619$.

b. A similar result is found.

4.181 The mgf for Z is $m_Z(t) = E(e^{Zt}) = E(e^{(\frac{y-\mu}{\sigma})t}) = e^{-\frac{\mu}{\sigma}t} m_Y(t/\sigma) = e^{t^2/2}$, which is a mgf for a normal distribution with μ = 0 and σ = 1.

4.182 **a.** $P(Y \le 4) = P(X \le \ln 4) = P[Z \le (\ln 4 - 4)/1] = P(Z \le -2.61) = .0045$.
b. $P(Y > 8) = P(X > \ln 8) = P[Z > (\ln 8 - 4)/1] = P(Z > -1.92) = .9726$.

4.183 **a.** $E(Y) = e^{3+16/2} = e^{11}$ (598.74 g), $V(Y) = e^{22}(e^{16} - 1)10^{-4}$.

b. With $k = 2$, the interval is given by $E(Y) \pm 2\sqrt{V(Y)}$ or 598.74 ± 3,569,038.7. Since the weights must be positive, the final interval is (0, 3,570,236.1)
c. $P(Y < 598.74) = P(\ln Y < 6.3948) = P[Z < (6.3948 - 3)/4] = P(Z < .8487) = .8020$

4.184 The mgf for Y is $m_Y(t) = E(e^{tY}) = \frac{1}{2}\int_{\infty}^{0} e^{ty}e^{y}dy + \frac{1}{2}\int_{0}^{\infty} e^{ty}e^{-y}dy = \frac{1}{2}\int_{\infty}^{0} e^{(t+1)y}dy + \frac{1}{2}\int_{0}^{\infty} e^{-y(1-t)}dy$.

This simplifies to $m_Y(t) = \frac{1}{1-t^2}$. Using this, $E(Y) = m'(t)\big|_{t=0} = \frac{2t}{(1-t^2)}\big|_{t=0} = 0$.

4.185 **a.** $\int_{-\infty}^{\infty} f(y)dy = a\int_{-\infty}^{\infty} f_1(y)dy + (1-a)\int_{-\infty}^{\infty} f_2(y)dy = a + (1-a) = 1$.

b. i. $E(Y) = \int_{-\infty}^{\infty} yf(y)dy = a\int_{-\infty}^{\infty} yf_1(y)dy + (1-a)\int_{-\infty}^{\infty} yf_2(y)dy = a\mu_1 + (1-a)\mu_2$

ii. $E(Y_2) = a\int_{-\infty}^{\infty} y^2 f_1(y)dy + (1-a)\int_{-\infty}^{\infty} y^2 f_2(y)dy = a(\mu_1^2 + \sigma_1^2) + (1-a)(\mu_2^2 + \sigma_2^2)$. So, $V(Y) = E(Y^2) - [E(Y)]^2 = a(\mu_1^2 + \sigma_1^2) + (1-a)(\mu_2^2 + \sigma_2^2) - [a\mu_1 + (1-a)\mu_2]^2$, which simplifies to $a\sigma_1^2 + (1-a)\sigma_2^2 + a(1-a)[\mu_1 - \mu_2]^2$

4.186 For $m = 2$, $E(Y) = \int_0^{\infty} y\frac{2y}{\alpha}e^{-y^2/\alpha}dy$. Let $u = y^2/\alpha$. Then, $dy = \frac{\sqrt{\alpha}}{2\sqrt{u}}du$. Then,
$E(Y) = \int_0^{\infty} \frac{2y^2}{\alpha}e^{-y^2/\alpha}dy = \sqrt{\alpha}\int_0^{\infty} u^{1/2}e^{-u}du = \sqrt{\alpha}\Gamma(3/2) = \frac{\sqrt{\alpha}\Gamma(1/2)}{2}$. Using similar methods, it can be shown that $E(Y^2) = \alpha$ so that $V(Y) = \alpha - \left[\frac{\sqrt{\alpha}\Gamma(1/2)}{2}\right]^2 = \alpha\left[1 - \frac{\pi}{4}\right]$, since it will be shown in Ex. 4.196 that $\Gamma(1/2) = \sqrt{\pi}$.

4.187 The density for Y = the life length of a resistor (in thousands of hours) is

$$f(y) = \frac{2ye^{-y^2/10}}{10},\ y > 0.$$

a. $P(Y > 5) = \int_5^\infty \frac{2ye^{-y^2/10}}{10}dy = -e^{-y^2/10}\Big]_5^\infty = e^{-2.5} = .082.$

b. Let X = # of resistors that burn out prior to 5000 hours. Thus, X is a binomial random variable with $n = 3$ and $p = .082$. Then, $P(X = 1) = 3(1 - .082)(.082)^2 = .0186$.

4.188 **a.** This is the exponential distribution with β = α.
b. Using the substitution u = y^m/α in the integrals below, we find:

$$E(Y) = \int_0^\infty \tfrac{m}{\alpha}y^m e^{-y^m/\alpha}dy = \alpha^{1/m}\int_0^\infty u^{1/m}e^{-u}du = \alpha^{1/m}\Gamma(1+1/m)$$

$$E(Y^2) = \int_0^\infty \tfrac{m}{\alpha}y^{m+1} e^{-y^m/\alpha}dy = \alpha^{2/m}\int_0^\infty u^{2/m}e^{-u}du = \alpha^{1/m}\Gamma(1+2/m).$$

Thus,
$V(Y) = \alpha^{2/m}[\Gamma(1+2/m) + \Gamma^2(1+1/m)]$.

4.189 Since this density is symmetric about 0, so using the result from Ex. 4.27, $E(Y) = 0$. Also, it is clear that $V(Y) = E(Y^2)$. Thus,

$E(Y^2) = \int_{-1}^{1} \frac{1}{B(1/2,(n-2)/2)}y^2(1-y^2)^{(n-4)/2}dy = \frac{B(3/2,(n-2)/2)}{B(1/2,(n-2)/2)} = \frac{1}{n-1} = V(Y)$. This equality follows after making the substitution $u = y^2$.

4.190 **a.** For the exponential distribution, $f(t) = \frac{1}{\beta}e^{-t/\beta}$ and $F(t) = 1 - e^{-t/\beta}$. Thus, $r(t) = 1/\beta$.

b. For the Weibull, $f(y) = \frac{my^{m-1}}{\alpha}e^{-y^m/\alpha}$ and $F(y) = 1 - e^{-y^m/\alpha}$. Thus, $r(t) = \frac{my^{m-1}}{\alpha}$, which is an increasing function of t when $m > 1$.

4.191 **a.** $G(y) = P(Y \le y \mid Y \ge c) = \frac{P(c \le Y \le y)}{P(Y \ge c)} = \frac{F(y) - F(c)}{1 - F(c)}$.

b. Refer to the properties of distribution functions; namely, show $G(-\infty) = 0$, $G(\infty) = 1$, and for constants a and b such that $a \le b$, $G(a) \le G(b)$.

c. It is given that $F(y) = 1 - e^{-y^2/3}$. Thus, by using the result in part b above,

$$P(Y \le 4 \mid Y \ge 2) = \frac{1 - e^{-4^2/3} - (1 - e^{-2^2/2}}{e^{-2^2/2}} = 1 - e^{-4}.$$

4.192 **a.** $E(V) = 4\pi\left(\frac{m}{2\pi KT}\right)^{3/2}\int_0^\infty v^3 e^{-v^2(m/2KT)}dv$. To evaluate this integral, let $u = v^2\left(\frac{m}{2KT}\right)$ so that

$dv = \sqrt{\frac{2KT}{m}}\frac{1}{2\sqrt{u}}du$ to obtain $E(V) = 2\sqrt{\frac{2KT}{m\pi}}\int_0^\infty ue^{-u}du = 2\sqrt{\frac{2KT}{m\pi}}\Gamma(2) = 2\sqrt{\frac{2KT}{m\pi}}$.

b. $E(\frac{1}{2}mV^2) = \frac{1}{2}mE(V^2) = 2\pi m\left(\frac{m}{2\pi KT}\right)^{3/2}\int_0^\infty v^4 e^{-v^2(m/2KT)}dv$. Here, let $u = v^2\left(\frac{m}{2KT}\right)$ so that $dv = \sqrt{\frac{2KT}{m}}\frac{1}{2\sqrt{u}}du$ to obtain $E(\frac{1}{2}mV^2) = \frac{2KT}{\sqrt{\pi}}\Gamma(\frac{5}{2}) = \frac{3}{2}KT$ (here, we again used the result from Ex. 4.196 where it is shown that $\Gamma(1/2) = \sqrt{\pi}$).

4.193 For $f(y) = \frac{1}{100}e^{-y/100}$, we have that $F(y) = 1 - e^{-y/100}$. Thus,

$$E(Y \mid Y \geq 50) = \frac{1}{e^{-1/2}}\int_{50}^\infty \frac{ye^{-y/100}}{100}dy = 150.$$

Note that this value is 50 + 100, where 100 is the (unconditional) mean of Y. This illustrates the memoryless property of the exponential distribution.

4.194 $\left[\frac{1}{\sqrt{2\pi}}\int_{-\infty}^\infty e^{-(1/2)uy^2}dy\right]\left[\frac{1}{\sqrt{2\pi}}\int_{-\infty}^\infty e^{-(1/2)ux^2}dx\right] = \frac{1}{2\pi}\int_{-\infty}^\infty\int_{-\infty}^\infty e^{-(1/2)u(x^2+y^2)}dxdy$. By changing to polar coordinates, $x^2 + y^2 = r^2$ and $dxdy = rdrd\theta$. Thus, the desired integral becomes

$$\frac{1}{2\pi}\int_0^{2\pi}\int_0^\infty e^{-(1/2)ur^2}rdrd\theta = \frac{1}{u}.$$

Note that the result proves that the standard normal density integrates to 1 with $u = 1$.

4.195 **a.** First note that $W = (Z^2 + 3Z)^2 = Z^4 + 6Z^3 + 9Z^2$. The odd moments of the standard normal are equal to 0, and $E(Z^2) = V(Z) + [E(Z)]^2 = 1 + 0 = 1$. Also, using the result in Ex. 4.199, $E(Z^4) = 3$ so that $E(W) = 3 + 9(1) = 12$.

b. Applying Ex. 4.198 and the result in part a:

$$P(W \leq w) \geq 1 - \frac{E(W)}{w} = .9,$$

so that $w = 120$.

4.196 $\Gamma(1/2) = \int_0^\infty y^{-1/2}e^{-y}dy = \int_0^\infty \sqrt{2}e^{-(1/2)x^2}dx = \sqrt{2}\sqrt{2\pi}\int_0^\infty \frac{1}{\sqrt{2\pi}}e^{-(1/2)x^2}dx = 2\sqrt{\pi}\left[\frac{1}{2}\right] = \sqrt{\pi}$ (relating the last integral to that $P(Z > 0)$, where Z is a standard normal random variable).

4.197 a. Let $y = \sin^2\theta$, so that $dy = 2\sin\theta\cos\theta d\theta$. Thus,

$$B(\alpha,\beta) = \int_0^1 y^{\alpha-1}(1-y)^{\beta-1}dy = 2\int_0^{\pi/2}\sin^{2\alpha-2}\theta(1-\sin^2\theta)^{\beta-1}d\theta = 2\int_0^{\pi/2}\sin^{2\alpha-1}\theta\cos^{2\beta-1}d\theta,$$ using the trig identity $1 - \sin^2\theta = \cos^2\theta$.

b. Following the text, $\Gamma(\alpha)\Gamma(\beta)=\int_0^\infty y^{\alpha-1}e^{-y}dy\int_0^\infty z^{\beta-1}e^{-z}dz=\int_0^\infty\int_0^\infty y^{\alpha-1}z^{\beta-1}e^{-y-z}dydz$. Now, use the transformation $y=r^2\cos^2\theta,\ x=r^2\sin^2\theta$ so that $dydz=4r^3\cos\theta\sin\theta$. Following this and using the result in part a, we find

$$\Gamma(\alpha)\Gamma(\beta)=B(\alpha,\beta)\int_0^\infty r^{2(\alpha+\beta-1)}e^{-r^2}2rdr.$$

A final transformation with $x = r^2$ gives $\Gamma(\alpha)\Gamma(\beta)=B(\alpha,\beta)\Gamma(\alpha+\beta)$, proving the result.

4.198 Note that

$$E[|g(Y)|]=\int_{-\infty}^{\infty}|g(y)|f(y)dy\ge\int_{|g(y)|>k}^{\infty}|g(y)|f(y)dy>\int_{|g(y)|>k}^{\infty}kf(y)dy=kP(|g(Y)|>k),$$

Since $|g(y)|>k$ for this integral. Therefore,

$$P(|g(Y)|\le k)\ge 1-E(|g(Y)|)/k.$$

4.199 **a.** Define $g(y)=y^{2i-1}e^{-y^2/2}$ for positive integer values of i. Observe that $g(-y)=-g(y)$ so that $g(y)$ is an odd function. The expected value $E(Z^{2i-1})$ can be written as $E(Z^{2i-1})=\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}g(y)dy$ which is thus equal to 0.

b. Now, define $h(y)=y^{2i}e^{-y^2/2}$ for positive integer values of i. Observe that $h(-y)=h(y)$ so that $h(y)$ is an even function. The expected value $E(Z^{2i})$ can be written as $E(Z^{2i})=\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}h(y)dy=\int_0^{\infty}\frac{2}{\sqrt{2\pi}}h(y)dy$. Therefore, the integral becomes

$$E(Z^{2i})=\int_0^\infty\frac{2}{\sqrt{2\pi}}y^{2i}e^{-y^2/2}dy=\frac{1}{\sqrt{\pi}}\int_0^\infty 2^i w^{i-1/2}e^{-w}dw=\frac{1}{\sqrt{\pi}}2^i\Gamma(i+1/2).$$

In the last integral, we applied the transformation $w = z^2/2$.

c.

$$E(Z^2)=\tfrac{1}{\sqrt{\pi}}2^1\Gamma(1+1/2)=\tfrac{1}{\sqrt{\pi}}2^1(1/2)\sqrt{\pi}=1$$
$$E(Z^4)=\tfrac{1}{\sqrt{\pi}}2^2\Gamma(2+1/2)=\tfrac{1}{\sqrt{\pi}}2^2(3/2)(1/2)\sqrt{\pi}=3$$
$$E(Z^6)=\tfrac{1}{\sqrt{\pi}}2^3\Gamma(3+1/2)=\tfrac{1}{\sqrt{\pi}}2^3(5/2)(3/2)(1/2)\sqrt{\pi}=15$$
$$E(Z^8)=\tfrac{1}{\sqrt{\pi}}2^4\Gamma(4+1/2)=\tfrac{1}{\sqrt{\pi}}2^4(7/2)(5/2)(3/2)(1/2)\sqrt{\pi}=105.$$

d. The result follows from:

$$\prod_{j=i}^{i}(2j-1)=\prod_{j=i}^{i}2(j-1/2)=2^i\prod_{j=i}^{i}(j-1/2)=2^i\Gamma(i+1/2)\left(\tfrac{1}{\sqrt{\pi}}\right)=E(Z^{2i}).$$

4.200 **a.** $E(Y^a) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 y^{a+\alpha-1}(1-y)^{\beta-1}dy = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(a+\alpha)\Gamma(\beta)}{\Gamma(a+\alpha+\beta)} = \frac{\Gamma(\alpha+\beta)\Gamma(a+\alpha)}{\Gamma(\alpha)\Gamma(a+\alpha+\beta)}$.

b. The value $\alpha + a$ must be positive in the beta density.

c. With $a = 1$, $E(Y^1) = \frac{\Gamma(\alpha+\beta)\Gamma(1+\alpha)}{\Gamma(\alpha)\Gamma(1+\alpha+\beta)} = \frac{\alpha}{\alpha+\beta}$.

d. With $a = 1/2$, $E(Y^{1/2}) = \frac{\Gamma(\alpha+\beta)\Gamma(1/2+\alpha)}{\Gamma(\alpha)\Gamma(1/2+\alpha+\beta)}$.

e. With $a = -1$, $E(Y^{-1}) = \frac{\Gamma(\alpha+\beta)\Gamma(\alpha-1)}{\Gamma(\alpha)\Gamma(\alpha+\beta-1)} = \frac{\alpha+\beta-1}{\alpha-1}$, $\alpha > 1$

With $a = -1/2$, $E(Y^{-1/2}) = \frac{\Gamma(\alpha+\beta)\Gamma(\alpha-1/2)}{\Gamma(\alpha)\Gamma(\alpha+\beta-1/2)}$, $\alpha > 1/2$

With $a = -2$, $E(Y^{-2}) = \frac{\Gamma(\alpha+\beta)\Gamma(\alpha-2)}{\Gamma(\alpha)\Gamma(\alpha+\beta-2)} = \frac{(\alpha+\beta-1)(\alpha+\beta-2)}{(\alpha-1)(\alpha-2)}$, $\alpha > 2$.

Chapter 5: Multivariate Probability Distributions

5.1 **a.** The sample space S gives the possible values for Y_1 and Y_2:

S	AA	AB	AC	BA	BB	BC	CA	CB	CC
(y_1, y_2)	(2, 0)	(1, 1)	(1, 0)	(1, 1)	(0, 2)	(1, 0)	(1, 0)	(0, 1)	(0, 0)

Since each sample point is equally likely with probably 1/9, the joint distribution for Y_1 and Y_2 is given by

		y_1		
		0	1	2
	0	1/9	2/9	1/9
y_2	1	2/9	2/9	0
	2	1/9	0	0

b. $F(1, 0) = p(0, 0) + p(1, 0) = 1/9 + 2/9 = 3/9 = 1/3$.

5.2 **a.** The sample space for the toss of three balanced coins w/ probabilities are below:

Outcome	HHH	HHT	HTH	HTT	THH	THT	TTH	TTT
(y_1, y_2)	(3, 1)	(3, 1)	(2, 1)	(1, 1)	(2, 2)	(1, 2)	(1, 3)	(0, –1)
probability	1/8	1/8	1/8	1/8	1/8	1/8	1/8	1/8

		y_1			
		0	1	2	3
	–1	1/8	0	0	0
y_2	1	0	1/8	2/8	1/8
	2	0	1/8	1/8	0
	3	0	1/8	0	0

b. $F(2, 1) = p(0, -1) + p(1, 1) + p(2, 1) = 1/2$.

5.3 Note that using material from Chapter 3, the joint probability function is given by

$$p(y_1, y_2) = P(Y_1 = y_1, Y_2 = y_2) = \frac{\binom{4}{y_1}\binom{3}{y_2}\binom{2}{3-y_1-y_2}}{\binom{9}{3}}, \text{ where } 0 \le y_1,\ 0 \le y_2, \text{ and } y_1 + y_2 \le 3.$$

In table format, this is

		y_1			
		0	1	2	3
	0	0	3/84	6/84	1/84
y_2	1	4/84	24/84	12/84	0
	2	12/84	18/84	0	0
	3	4/84	0	0	0

5.4 **a.** All of the probabilities are at least 0 and sum to 1.
b. $F(1, 2) = P(Y_1 \le 1, Y_2 \le 2) = 1$. Every child in the experiment either survived or didn't and used either 0, 1, or 2 seatbelts.

5.5 **a.** $P(Y_1 \le 1/2, Y_2 \le 1/3) = \int_0^{1/2}\int_0^{1/3} 3y_1 dy_1 dy_2 = .1065$.

b. $P(Y_2 \le Y_1/2) = \int_0^1 \int_0^{y_1/2} 3y_1 dy_1 dy_2 = .5$.

5.6 **a.** $P(Y_1 - Y_2 > .5) = P(Y_1 > .5 + Y_2) = \int_0^{.5}\int_{y_2+.5}^{1} 1 dy_1 dy_2 = \int_0^{.5} [y_1]_{y_2+.5}^{1} dy_2 = \int_0^{.5} (.5 - y_2) dy_2 = .125$.

b. $P(Y_1Y_2 < .5) = 1 - P(Y_1Y_2 > .5) = 1 - P(Y_1 > .5/Y_2) = 1 - \int_{.5}^{1}\int_{.5/y_2}^{1} 1 dy_1 dy_2 = 1 - \int_{.5}^{1} (1 - .5/y_2) dy_2$

$= 1 - [.5 + .5\ln(.5)] = .8466$.

5.7 **a.** $P(Y_1 < 1, Y_2 > 5) = \int_0^1 \int_5^{\infty} e^{-(y_1+y_2)} dy_1 dy_2 = \left[\int_0^1 e^{-y_1} dy_1\right]\left[\int_5^{\infty} e^{-y_2} dy_2\right] = \left[1 - e^{-1}\right]e^{-5} = .00426$.

b. $P(Y_1 + Y_2 < 3) = P(Y_1 < 3 - Y_2) = \int_0^3 \int_0^{3-y_2} e^{-(y_1+y_2)} dy_1 dy_2 = 1 - 4e^{-3} = .8009$.

5.8 **a.** Since the density must integrate to 1, evaluate $\int_0^1\int_0^1 ky_1y_2 dy_1 dy_2 = k/4 = 1$, so $k = 4$.

b. $F(y_1, y_2) = P(Y_1 \le y_1, Y_2 \le y_2) = 4\int_0^{y_2}\int_0^{y_1} t_1 t_2 dt_1 dt_2 = y_1^2 y_2^2$, $0 \le y_1 \le 1$, $0 \le y_2 \le 1$.

c. $P(Y_1 \le 1/2, Y_2 \le 3/4) = (1/2)^2(3/4)^2 = 9/64$.

5.9 **a.** Since the density must integrate to 1, evaluate $\int_0^1 \int_0^{y_2} k(1 - y_2) dy_1 dy_2 = k/6 = 1$, so $k = 6$.

b. Note that since $Y_1 \le Y_2$, the probability must be found in two parts (drawing a picture is useful):

$P(Y_1 \le 3/4, Y_2 \ge 1/2) = \int_{1/2}^{1}\int_{1/2}^{1} 6(1 - y_2) dy_1 dy_2 + \int_{1/2}^{3/4}\int_{y_1}^{1} 6(1 - y_2) dy_2 dy_1 = 24/64 + 7/64 = 31/64$.

5.10 **a.** Geometrically, since Y_1 and Y_2 are distributed uniformly over the triangular region, using the area formula for a triangle $k = 1$.

b. This probability can also be calculated using geometric considerations. The area of the triangle specified by $Y_1 \ge 3Y_2$ is 2/3, so this is the probability.

5.11 The area of the triangular region is 1, so with a uniform distribution this is the value of the density function. Again, using geometry (drawing a picture is again useful):

a. $P(Y_1 \le 3/4, Y_2 \le 3/4) = 1 - \text{P}(Y_1 > 3/4) - \text{P}(Y_2 > 3/4) = 1 - \frac{1}{2}\left(\frac{1}{2}\right)\left(\frac{1}{4}\right) - \frac{1}{2}\left(\frac{1}{4}\right)\left(\frac{1}{4}\right) = \frac{29}{32}$.

b. $P(Y_1 - Y_2 \ge 0) = P(Y_1 \ge Y_2)$. The region specified in this probability statement represents 1/4 of the total region of support, so $P(Y_1 \ge Y_2) = 1/4$.

5.12 Similar to Ex. 5.11:

a. $P(Y_1 \le 3/4, Y_2 \le 3/4) = 1 - \text{P}(Y_1 > 3/4) - \text{P}(Y_2 > 3/4) = 1 - \frac{1}{2}\left(\frac{1}{4}\right)\left(\frac{1}{4}\right) - \frac{1}{2}\left(\frac{1}{4}\right)\left(\frac{1}{4}\right) = \frac{7}{8}$.

b. $P(Y_1 \le 1/2, Y_2 \le 1/2) = \int_0^{1/2}\int_0^{1/2} 2\,dy_1 dy_2 = 1/2.$

5.13 **a.** $F(1/2, 1/2) = \int_0^{1/2}\int_{y_1-1}^{1/2} 30 y_1 y_2^2\, dy_2 dy_1 = \frac{9}{16}$.

b. Note that:

$F(1/2,\,2) = F(1/2,\,1) = P(Y_1 \le 1/2, Y_2 \le 1) = P(Y_1 \le 1/2, Y_2 \le 1/2) + P(Y_1 \le 1/2, Y_2 > 1/2)$

So, the first probability statement is simply $F(1/2,\,1/2)$ from part a. The second probability statement is found by

$$P(Y_1 \le 1/2, Y_2 > 1/2) = \int_{1/2}^{1}\int_{0}^{1-y_2} 30 y_1 y_2^2\, dy_2 dy = \frac{4}{16}.$$

Thus, $F(1/2,\,2) = \frac{9}{16} + \frac{4}{16} = \frac{13}{16}$.

c. $P(Y_1 > Y_2) = 1 - P(Y_1 \le Y_2) = 1 - \int_0^{1/2}\int_{y_1}^{1-y_1} 30 y_1 y_2^2\, dy_2 dy_1 = 1 - \frac{11}{32} = \frac{21}{32} = .65625.$

5.14 **a.** Since $f(y_1, y_2) \ge 0$, simply show $\int_0^1\int_{y_1}^{2-y_1} 6 y_1^2 y_2\, dy_2 dy_1 = 1$.

b. $P(Y_1 + Y_2 < 1) = P(Y_2 < 1 - Y_1) = \int_0^{.5}\int_{y_1}^{1-y_1} 6 y_1^2 y_2\, dy_2 dy_1 = 1/16$.

5.15 **a.** $P(Y_1 < 2, Y_2 > 1) = \int_1^2\int_1^{y_1} e^{-y_1}\, dy_2 dy_1 = \int_1^2\int_{y_2}^{2} e^{-y_1}\, dy_1 dy_2 = e^{-1} - 2e^{-2}$.

b. $P(Y_1 \ge 2Y_2) = \int_0^{\infty}\int_{2y_2}^{\infty} e^{-y_1}\, dy_1 dy_2 = 1/2$.

c. $P(Y_1 - Y_2 \ge 1) = P(Y_1 \ge Y_2 + 1) = \int_0^{\infty}\int_{y_2+1}^{\infty} e^{-y_1}\, dy_1 dy_2 = e^{-1}$.

5.16 **a.** $P(Y_1 < 1/2, Y_2 > 1/4) = \int_{1/4}^{1}\int_{0}^{1/2}(y_1 + y_2)dy_1dy_2 = 21/64 = .328125.$

b. $P(Y_1 + Y_2 \le 1) = P(Y_1 \le 1 - Y_2) = \int_{0}^{1}\int_{0}^{1-y_2}(y_1 + y_2)dy_1dy_2 = 1/3$.

5.17 This can be found using integration (polar coordinates are helpful). But, note that this is a bivariate uniform distribution over a circle of radius 1, and the probability of interest represents 50% of the support. Thus, the probability is .50.

5.18 $$P(Y_1 > 1, Y_2 > 1) = \int_{1}^{\infty}\int_{1}^{\infty}\tfrac{1}{8}y_1e^{-(y_1+y_2)/2}dy_1dy_2 = \left[\int_{1}^{\infty}\tfrac{1}{4}y_1e^{-y_1/2}dy_1\right]\left[\int_{1}^{\infty}\tfrac{1}{2}e^{-y_2/2}dy_2\right] = \tfrac{3}{2}e^{-\frac{1}{2}}\left(e^{-\frac{1}{2}}\right) = \tfrac{3}{2}e^{-1}$$

5.19 **a.** The marginal probability function is given in the table below.

y_1	0	1	2
$p_1(y_1)$	4/9	4/9	1/9

b. No, evaluating binomial probabilities with $n = 3$, $p = 1/3$ yields the same result.

5.20 **a.** The marginal probability function is given in the table below.

y_2	–1	1	2	3
$p_2(y_2)$	1/8	4/8	2/8	1/8

b. $P(Y_1 = 3 \mid Y_2 = 1) = \frac{P(Y_1=3, Y_2=1)}{P(Y_2=1)} = \frac{1/8}{4/8} = 1/4$.

5.21 **a.** The marginal distribution of Y_1 is hypergeometric with $N = 9$, $n = 3$, and $r = 4$.

b. Similar to part a, the marginal distribution of Y_2 is hypergeometric with $N = 9$, $n = 3$, and $r = 3$. Thus,

$$P(Y_1 = 1 \mid Y_2 = 2) = \frac{P(Y_1=1, Y_2=2)}{P(Y_2=2)} = \frac{\binom{4}{1}\binom{3}{2}\binom{2}{0}}{\binom{9}{3}} \Bigg/ \frac{\binom{3}{2}\binom{6}{1}}{\binom{9}{3}} = 2/3.$$

c. Similar to part b,

$$P(Y_3 = 1 \mid Y_2 = 1) = P(Y_1 = 1 \mid Y_2 = 1) = \frac{P(Y_1=1, Y_2=1)}{P(Y_2=1)} = \frac{\binom{3}{1}\binom{2}{1}\binom{4}{1}}{\binom{9}{3}} \Bigg/ \frac{\binom{3}{1}\binom{6}{2}}{\binom{9}{3}} = 8/15.$$

5.22 **a.** The marginal distributions for Y_1 and Y_2 are given in the margins of the table.

b. $P(Y_2 = 0 \mid Y_1 = 0) = .38/.76 = .5$ $\quad P(Y_2 = 1 \mid Y_1 = 0) = .14/.76 = .18$
$P(Y_2 = 2 \mid Y_1 = 0) = .24/.76 = .32$

c. The desired probability is $P(Y_1 = 0 \mid Y_2 = 0) = .38/.55 = .69$.

5.23 **a.** $f_2(y_2) = \int_{y_2}^{1} 3y_1 dy_1 = \frac{3}{2} - \frac{3}{2}y_2^2$, $0 \le y_2 \le 1$.

b. Defined over $y_2 \le y_1 \le 1$, with the constant $y_2 \ge 0$.

c. First, we have $f_1(y_1) = \int_0^{y_1} 3y_1 dy_2 = 3y_2^2$, $0 \le y_1 \le 1$. Thus,

$f(y_2 \mid y_1) = 1/y_1$, $0 \le y_2 \le y_1$. So, conditioned on $Y_1 = y_1$, we see Y_2 has a uniform distribution on the interval $(0, y_1)$. Therefore, the probability is simple:

$$P(Y_2 > 1/2 \mid Y_1 = 3/4) = (3/4 - 1/2)/(3/4) = 1/3.$$

5.24 **a.** $f_1(y_1) = 1, 0 \le y_1 \le 1$, $f_2(y_2) = 1, 0 \le y_2 \le 1$.

b. Since both Y_1 and Y_2 are uniformly distributed over the interval (0, 1), the probabilities are the same: .2

c. $0 \le y_2 \le 1$.

d. $f(y_1 \mid y_2) = f(y_1) = 1, 0 \le y_1 \le 1$

e. $P(.3 < Y_1 < .5 \mid Y_2 = .3) = .2$

f. $P(.3 < Y_2 < .5 \mid Y_2 = .5) = .2$

g. The answers are the same.

5.25 **a.** $f_1(y_1) = e^{-y_1}$, $y_1 > 0$, $f_2(y_2) = e^{-y_2}$, $y_2 > 0$. These are both exponential density functions with $\beta = 1$.

b. $P(1 < Y_1 < 2.5) = P(1 < Y_2 < 2.5) = e^{-1} - e^{-2.5} = .2858$.

c. $y_2 > 0$.

d. $f(y_1 \mid y_2) = f_1(y_1) = e^{-y_1}$, $y_1 > 0$.

e. $f(y_2 \mid y_1) = f_2(y_2) = e^{-y_2}$, $y_2 > 0$.

f. The answers are the same.

g. The probabilities are the same.

5.26 **a.** $f_1(y_1) = \int_0^1 4y_1y_2 dy_2 = 2y_1$, $0 \le y_1 \le 1$; $f(y_2) = 2y_2$, $0 \le y_2 \le 1$.

b. $P(Y_1 \le 1/2 | Y_2 \ge 3/4) = \dfrac{\int_0^{1/2}\int_{3/4}^{1} 4y_1y_2 dy_1 dy_2}{\int_{3/4}^{1} 2y_2 dy_2} = \int_0^{1/2} 2y_1 dy_1 = 1/4$.

c. $f(y_1 \mid y_2) = f_1(y_1) = 2y_1$, $0 \le y_1 \le 1$.

d. $f(y_2 \mid y_1) = f_2(y_2) = 2y_2$, $0 \le y_2 \le 1$.

e. $P(Y_1 \le 3/4 | Y_2 = 1/2) = P(Y_1 \le 3/4) = \int_0^{3/4} 2y_1 dy_1 = 9/16$.

5.27 **a.** $f_1(y_1) = \int_{y_1}^{1} 6(1-y_2)dy_2 = 3(1-y_1)^2$, $0 \le y_1 \le 1$;

$f_2(y_2) = \int_{0}^{y_2} 6(1-y_2)dy_1 = 6y_2(1-y_2)$, $0 \le y_2 \le 1$.

b. $P(Y_2 \le 1/2 | Y_1 \le 3/4) = \dfrac{\int_0^{1/2}\int_0^{y_2} 6(1-y_2)dy_1dy_2}{\int_0^{3/4} 3(1-y_1)^2 dy_1} = 32/63.$

c. $f(y_1 \mid y_2) = 1/y_2$, $0 \le y_1 \le y_2 \le 1$.

d. $f(y_2 \mid y_1) = 2(1-y_2)/(1-y_1)^2$, $0 \le y_1 \le y_2 \le 1$.

e. From part **d**, $f(y_2 \mid 1/2) = 8(1-y_2)$, $1/2 \le y_2 \le 1$. Thus, $P(Y_2 \ge 3/4 \mid Y_1 = 1/2) = 1/4$.

5.28 Referring to Ex. 5.10:

a. First, find $f_2(y_2) = \int_{2y_2}^{2} 1dy_1 = 2(1-y_2)$, $0 \le y_2 \le 1$. Then, $P(Y_2 \ge .5) = .25$.

b. First find $f(y_1 \mid y_2) = \frac{1}{2(1-y_2)}$, $2y_2 \le y_1 \le 2$. Thus, $f(y_1 \mid .5) = 1$, $1 \le y_1 \le 2$ — the conditional distribution is uniform on (1, 2). Therefore, $P(Y_1 \ge 1.5 \mid Y_2 = .5) = .5$

5.29 Referring to Ex. 5.11:

a. $f_2(y_2) = \int_{y_2-1}^{1-y_2} 1dy_1 = 2(1-y_2)$, $0 \le y_2 \le 1$. In order to find $f_1(y_1)$, notice that the limits of integration are different for $0 \le y_1 \le 1$ and $-1 \le y_1 \le 0$. For the first case:
$f_1(y_1) = \int_0^{1-y_1} 1dy_2 = 1-y_1$, for $0 \le y_1 \le 1$. For the second case, $f_1(y_1) = \int_0^{1+y_1} 1dy_2 = 1+y_1$, for $-1 \le y_1 \le 0$. This can be written as $f_1(y_1) = 1 - |y_1|$, for $-1 \le y_1 \le 1$.

b. The conditional distribution is $f(y_2 \mid y_1) = \frac{1}{1-|y_1|}$, for $0 \le y_1 \le 1 - |y_1|$. Thus,
$f(y_2 \mid 1/4) = 4/3$. Then, $P(Y_2 > 1/2 \mid Y_1 = 1/4) = \int_{1/2}^{3/4} 4/3 dy_2 = 1/3$.

5.30 **a.** $P(Y_1 \ge 1/2, Y_2 \le 1/4) = \int_0^{1/4}\int_{1/2}^{1-y_2} 2dy_1dy_2 = \frac{3}{16}$. And, $P(Y_2 \le 1/4) = \int_0^{1/4} 2(1-y_2)dy_2 = \frac{7}{16}$.
Thus, $P(Y_1 \ge 1/2 \mid Y_2 \le 1/4) = \frac{3}{7}$.

b. Note that $f(y_1 \mid y_2) = \frac{1}{1-y_2}$, $0 \le y_1 \le 1-y_2$. Thus, $f(y_1 \mid 1/4) = 4/3$, $0 \le y_1 \le 3/4$.
Thus, $P(Y_2 > 1/2 \mid Y_1 = 1/4) = \int_{1/2}^{3/4} 4/3 dy_2 = 1/3$.

5.31 **a.** $f_1(y_1) = \int_{y_1-1}^{1-y_1} 30y_1y_2^2dy_2 = 20y_1(1-y_1)^2$, $0 \le y_1 \le 1$.

b. This marginal density must be constructed in two parts:

$$f_2(y_2) = \begin{cases} \int_0^{1+y_2} 30y_1y_2^2dy_1 = 15y_2^2(1+y_2) & -1 \le y_2 \le 0 \\ \int_0^{1-y_2} 30y_1y_2^2dy_1 = 5y_2^2(1-y_2) & 0 \le y_2 \le 1 \end{cases}.$$

c. $f(y_2 \mid y_1) = \frac{3}{2}y_2^2(1-y_1)^{-3}$, for $y_1 - 1 \le y_2 \le 1 - y_1$.

d. $f(y_2 \mid .75) = \frac{3}{2}y_2^2(.25)^{-3}$, for $-.25 \le y_2 \le .25$, so $P(Y_2 > 0 \mid Y_1 = .75) = .5$.

5.32 **a.** $f_1(y_1) = \int_{y_1}^{2-y_1} 6y_1^2y_2dy_2 = 12y_1^2(1-y_1)$, $0 \le y_1 \le 1$.

b. This marginal density must be constructed in two parts:

$$f_2(y_2) = \begin{cases} \int_0^{y_2} 6y_1^2y_2dy_1 = 2y_2^4 & 0 \le y_2 \le 1 \\ \int_0^{2-y_2} 6y_1^2y_2dy_1 = 2y_2(2-y_2)^3 & 1 \le y_2 \le 2 \end{cases}.$$

c. $f(y_2 \mid y_1) = \frac{1}{2}y_2/(1-y_1)$, $y_1 \le y_2 \le 2 - y_1$.

d. Using

the density found in part **c**, $P(Y_2 < 1.1 \mid Y_1 = .6) = \frac{1}{2}\int_{.6}^{11} y_2/.4dy_2 = .53$

5.33 Refer to Ex. 5.15:

a. $f_1(y_1) = \int_0^{y_1} e^{-y_1}dy_2 = y_1e^{-y_1}$, $y_1 \ge 0$. $f_2(y_2) = \int_{y_2}^{\infty} e^{-y_1}dy_1 = e^{-y_2}$, $y_2 \ge 0$.

b. $f(y_1 \mid y_2) = e^{-(y_1-y_2)}$, $y_1 \ge y_2$.

c. $f(y_2 \mid y_1) = 1/y_1$, $0 \le y_2 \le y_1$.

d. The density functions are different.

e. The marginal and conditional probabilities can be different.

5.34 **a.** Given $Y_1 = y_1$, Y_2 has a uniform distribution on the interval $(0, y_1)$.

b. Since $f_1(y_1) = 1$, $0 \le y_1 \le 1$, $f(y_1, y_2) = f(y_2 \mid y_1)f_1(y_1) = 1/y_1$, $0 \le y_2 \le y_1 \le 1$.

c. $f_2(y_2) = \int_{y_2}^{1} 1/y_1\, dy_1 = -\ln(y_2)$, $0 \le y_2 \le 1$.

5.35 With $Y_1 = 2$, the conditional distribution of Y_2 is uniform on the interval (0, 2). Thus, $P(Y_2 < 1 \mid Y_1 = 2) = .5$.

5.36 **a.** $f_1(y_1) = \int_0^1 (y_1 + y_2)dy_2 = y_1 + \frac{1}{2}$, $0 \le y_1 \le 1$. Similarly $f_2(y_2) = y_2 + \frac{1}{2}$, $0 \le y_2 \le 1$.

b. First, $P(Y_2 \ge \frac{1}{2}) = \int_{1/2}^1 (y_2 + \frac{1}{2}) = \frac{5}{8}$, and $P(Y_1 \ge \frac{1}{2}, Y_2 \ge \frac{1}{2}) = \int_{1/2}^1 \int_{1/2}^1 (y_1 + y_2)dy_1dy_2 = \frac{3}{8}$.
Thus, $P(Y_1 \ge \frac{1}{2} \mid Y_2 \ge \frac{1}{2}) = \frac{3}{5}$.

c. $P(Y_1 > .75 \mid Y_2 = .5) = \dfrac{\int_{.75}^1 \left(y_1 + \frac{1}{2}\right)dy_1}{\frac{1}{2} + \frac{1}{2}} = .34375.$

5.37 Calculate $f_2(y_2) = \int_0^\infty \frac{y_1}{8} e^{-(y_1+y_2)/2} dy_1 = \frac{1}{2} e^{-y_2/2}$, $y_2 > 0$. Thus, Y_2 has an exponential distribution with $\beta = 2$ and $P(Y_2 > 2) = 1 - F(2) = e^{-1}$.

5.38 This is the identical setup as in Ex. 5.34.

a. $f(y_1, y_2) = f(y_2 \mid y_1)f_1(y_1) = 1/y_1$, $0 \le y_2 \le y_1 \le 1$.

b. Note that $f(y_2 \mid 1/2) = 1/2$, $0 \le y_2 \le 1/2$. Thus, $P(Y_2 < 1/4 \mid Y_1 = 1/2) = 1/2$.

c. The probability of interest is $P(Y_1 > 1/2 \mid Y_2 = 1/4)$. So, the necessary conditional density is $f(y_1 \mid y_2) = f(y_1, y_2)/f_2(y_2) = \frac{1}{y_1(-\ln y_2)}$, $0 \le y_2 \le y_1 \le 1$. Thus,

$$P(Y_1 > 1/2 \mid Y_2 = 1/4) = \int_{1/2}^1 \frac{1}{y_1 \ln 4} dy_1 = 1/2.$$

5.39 The result follows from:

$$P(Y_1 = y_1 \mid W = w) = \frac{P(Y_1 = y_1, W = w)}{P(W = w)} = \frac{P(Y_1 = y_1, Y_1 + Y_2 = w)}{P(W = w)} = \frac{P(Y_1 = y_1, Y_2 = w - y_1)}{P(W = w)}.$$

Since Y_1 and Y_2 are independent, this is

$$P(Y_1 = y_1 \mid W = w) = \frac{P(Y_1 = y_1)P(Y_2 = w - y_1)}{P(W = w)} = \frac{\frac{\lambda_1^{y_1} e^{-\lambda_1}}{y_1!}\left(\frac{\lambda_2^{w-y_1} e^{-\lambda_2}}{(w-y_1)!}\right)}{\frac{(\lambda_1+\lambda_2)^w e^{-(\lambda_1+\lambda_2)}}{w!}}$$

$$= \binom{w}{y_1}\left(\frac{\lambda_1}{\lambda_1 + \lambda_2}\right)^{y_1}\left(1 - \frac{\lambda_1}{\lambda_1 + \lambda_2}\right)^{w-y_1}.$$

This is the binomial distribution with $n = w$ and $p = \dfrac{\lambda_1}{\lambda_1 + \lambda_2}$.

5.40 As the Ex. 5.39 above, the result follows from:

$$P(Y_1 = y_1 \mid W = w) = \frac{P(Y_1 = y_1, W = w)}{P(W = w)} = \frac{P(Y_1 = y_1, Y_1 + Y_2 = w)}{P(W = w)} = \frac{P(Y_1 = y_1, Y_2 = w - y_1)}{P(W = w)}.$$

Since Y_1 and Y_2 are independent, this is (all terms involving p_1 and p_2 drop out)

$$P(Y_1 = y_1 \mid W = w) = \frac{P(Y_1 = y_1)P(Y_2 = w - y_1)}{P(W = w)} = \frac{\binom{n_1}{y_1}\binom{n_2}{w - y_1}}{\binom{n_1 + n_2}{w}}, \qquad \begin{matrix} 0 \le y_1 \le n_1 \\ 0 \le w - y_1 \le n_2 \end{matrix}.$$

5.41 Let Y = # of defectives in a random selection of three items. Conditioned on p, we have

$$P(Y = y \mid p) = \binom{3}{y} p^y (1-p)^{3-y},\ y = 0, 1, 2, 3.$$

We are given that the proportion of defectives follows a uniform distribution on (0, 1), so the unconditional probability that $Y = 2$ can be found by

$$P(Y = 2) = \int_0^1 P(Y = 2, p)dp = \int_0^1 P(Y = 2 \mid p) f(p)dp = \int_0^1 3p^2(1-p)^{3-1}dp = 3\int_0^1 (p^2 - p^3)dp$$
$$= 1/4.$$

5.42 (Similar to Ex. 5.41) Let Y = # of defects per yard. Then,

$$p(y) = \int_0^\infty P(Y = y, \lambda)d\lambda = \int_0^\infty P(Y = y \mid \lambda) f(\lambda)d\lambda = \int_0^\infty \tfrac{\lambda^y e^{-\lambda}}{y!} e^{-\lambda} d\lambda = \left(\tfrac{1}{2}\right)^{y+1},\ y = 0, 1, 2, \ldots.$$

Note that this is essentially a geometric distribution (see Ex. 3.88).

5.43 Assume $f(y_1 \mid y_2) = f_1(y_1)$. Then, $f(y_1, y_2) = f(y_1 \mid y_2) f_2(y_2) = f_1(y_1) f_2(y_2)$ so that Y_1 and Y_2 are independent. Now assume that Y_1 and Y_2 are independent. Then, there exists functions g and h such that $f(y_1, y_2) = g(y_1)h(y_2)$ so that

$$1 = \int\int f(y_1, y_2)dy_1 dy_2 = \int g(y_1)dy_1 \times \int h(y_2)dy_2.$$

Then, the marginals for Y_1 and Y_2 can be defined by

$$f_1(y_1) = \int \frac{g(y_1)h(y_2)}{\int g(y_1)dy_1 \times \int h(y_2)dy_2} dy_2 = \frac{g(y_1)}{\int g(y_1)dy_1}, \text{ so } f_2(y_2) = \frac{h(y_2)}{\int h(y_2)dy_2}.$$

Thus, $f(y_1, y_2) = f_1(y_1) f_2(y_2)$. Now it is clear that

$$f(y_1 \mid y_2) = f(y_1, y_2)/ f_2(y_2) = f_1(y_1) f_2(y_2)/ f_2(y_2) = f_1(y_1),$$

provided that $f_2(y_2) > 0$ as was to be shown.

5.44 The argument follows exactly as Ex. 5.43 with integrals replaced by sums and densities replaced by probability mass functions.

5.45 No. Counterexample: $P(Y_1 = 2, Y_2 = 2) = 0 \ne P(Y_1 = 2)P(Y_2 = 2) = (1/9)(1/9)$.

5.46 No. Counterexample: $P(Y_1 = 3, Y_2 = 1) = 1/8 \ne P(Y_1 = 3)P(Y_2 = 1) = (1/8)(4/8)$.

5.47 Dependent. For example: $P(Y_1 = 1, Y_2 = 2) \neq P(Y_1 = 1)P(Y_2 = 2)$.

5.48 Dependent. For example: $P(Y_1 = 0, Y_2 = 0) \neq P(Y_1 = 0)P(Y_2 = 0)$.

5.49 Note that $f_1(y_1) = \int_0^{y_1} 3y_1 dy_2 = 3y_1^2,\ 0 \le y_1 \le 1$, $f_2(y_2) = \int_{y_1}^{1} 3y_1 dy_1 = \frac{3}{2}[1 - y_2^2],\ 0 \le y_2 \le 1$.
Thus, $f(y_1, y_2) \neq f_1(y_1) f_2(y_2)$ so that Y_1 and Y_2 are dependent.

5.50 **a.** Note that $f_1(y_1) = \int_0^1 1 dy_2 = 1,\ 0 \le y_1 \le 1$ and $f_2(y_2) = \int_0^1 1 dy_1 = 1,\ 0 \le y_2 \le 1$. Thus, $f(y_1, y_2) = f_1(y_1) f_2(y_2)$ so that Y_1 and Y_2 are independent.

b. Yes, the conditional probabilities are the same as the marginal probabilities.

5.51 **a.** Note that $f_1(y_1) = \int_0^\infty e^{-(y_1+y_2)} dy_2 = e^{-y_1},\ y_1 > 0$ and $f_2(y_2) = \int_0^\infty e^{-(y_1+y_2)} dy_1 = e^{-y_2},\ y_2 > 0$.
Thus, $f(y_1, y_2) = f_1(y_1) f_2(y_2)$ so that Y_1 and Y_2 are independent.

b. Yes, the conditional probabilities are the same as the marginal probabilities.

5.52 Note that $f(y_1, y_2)$ can be factored and the ranges of y_1 and y_2 do not depend on each other so by Theorem 5.5 Y_1 and Y_2 are independent.

5.53 The ranges of y_1 and y_2 depend on each other so Y_1 and Y_2 cannot be independent.

5.54 The ranges of y_1 and y_2 depend on each other so Y_1 and Y_2 cannot be independent.

5.55 The ranges of y_1 and y_2 depend on each other so Y_1 and Y_2 cannot be independent.

5.56 The ranges of y_1 and y_2 depend on each other so Y_1 and Y_2 cannot be independent.

5.57 The ranges of y_1 and y_2 depend on each other so Y_1 and Y_2 cannot be independent.

5.58 Following Ex. 5.32, it is seen that $f(y_1, y_2) \neq f_1(y_1) f_2(y_2)$ so that Y_1 and Y_2 are dependent.

5.59 The ranges of y_1 and y_2 depend on each other so Y_1 and Y_2 cannot be independent.

5.60 From Ex. 5.36, $f_1(y_1) = y_1 + \frac{1}{2},\ 0 \le y_1 \le 1$, and $f_2(y_2) = y_2 + \frac{1}{2},\ 0 \le y_2 \le 1$. But, $f(y_1, y_2) \neq f_1(y_1) f_2(y_2)$ so Y_1 and Y_2 are dependent.

5.61 Note that $f(y_1, y_2)$ can be factored and the ranges of y_1 and y_2 do not depend on each other so by Theorem 5.5, Y_1 and Y_2 are independent.

5.62 Let X, Y denote the number on which person A, B flips a head on the coin, respectively. Then, X and Y are geometric random variables and the probability that the stop on the same number toss is:

$$P(X=1,Y=1)+P(X=2,Y=2)+\cdots=P(X=1)P(Y=1)+P(X=2)P(Y=2)+\cdots$$

$$=\sum_{i=1}^{\infty}P(X=i)P(Y=i)=\sum_{i=1}^{\infty}p(1-p)^{i-1}p(1-p)^{i-1}=p^2\sum_{k=0}^{\infty}[(1-p)^2]^k=\frac{p^2}{1-(1-p)^2}.$$

5.63 $P(Y_1>Y_2, Y_1<2Y_2)=\int_0^{\infty}\int_{y_1/2}^{y_1}e^{-(y_1+y_2)}dy_2dy_1=\frac{1}{6}$ and $P(Y_1<2Y_2)=\int_0^{\infty}\int_{y_1/2}^{\infty}e^{-(y_1+y_2)}dy_2dy_1=\frac{2}{3}$. So, $P(Y_1>Y_2 \mid Y_1<2Y_2)=1/4$.

5.64 $P(Y_1>Y_2, Y_1<2Y_2)=\int_0^{1}\int_{y_1/2}^{y_1}1dy_2dy_1=\frac{1}{4}$, $P(Y_1<2Y_2)=1-P(Y_1\geq 2Y_2)=1-\int_0^{1}\int_0^{y_1/2}1dy_2dy_1=\frac{3}{4}$. So, $P(Y_1>Y_2 \mid Y_1<2Y_2)=1/3$.

5.65 **a.** The marginal density for Y_1 is $f_1(y_1)=\int_0^{\infty}[(1-\alpha(1-2e^{-y_1})(1-2e^{-y_2})]e^{-y_1-y_2}dy_2$

$$=e^{-y_1}\left[\int_0^{\infty}e^{-y_2}dy_2-\alpha(1-2e^{-y_1})\int_0^{\infty}(e^{-y_2}-2e^{-2y_2})dy_2\right].$$

$$=e^{-y_1}\left[\int_0^{\infty}e^{-y_2}dy_2-\alpha(1-2e^{-y_1})(1-1)\right]=e^{-y_1},$$

which is the exponential density with a mean of 1.

b. By symmetry, the marginal density for Y_2 is also exponential with $\beta = 1$.

c. When $\alpha = 0$, then $f(y_1,y_2)=e^{-y_1-y_2}=f_1(y_1)f_2(y_2)$ and so Y_1 and Y_2 are independent. Now, suppose Y_1 and Y_2 are independent. Then, $E(Y_1Y_2) = E(Y_1)E(Y_2) = 1$. So,

$$E(Y_1Y_2)=\int_0^{\infty}\int_0^{\infty}y_1y_2[(1-\alpha(1-2e^{-y_1})(1-2e^{-y_2})]e^{-y_1-y_2}dy_1dy_2$$

$$=\int_0^{\infty}\int_0^{\infty}y_1y_2e^{-y_1-y_2}dy_1dy_2-\alpha\left[\int_0^{\infty}y_1(1-2e^{-y_1})e^{-y_1}dy_1\right]\times\left[\int_0^{\infty}y_2(1-2e^{-y_2})e^{-y_2}dy_2\right]$$

$=1-\alpha\left(1-\frac{1}{2}\right)\left(1-\frac{1}{2}\right)=1-\alpha/4$. This equals 1 only if $\alpha = 0$.

5.66 **a.** Since $F_2(\infty)=1$, $F(y_1,\infty)=F_1(y_1)\cdot 1\cdot[1-\alpha\{1-F_1(y_1)\}\{1-1\}]=F_1(y_1)$.

b. Similarly, it is $F_2(y_2)$ from $F(y_1,y_2)$

c. If $\alpha = 0$, $F(y_1,y_2)=F_1(y_1)F_2(y_2)$, so by Definition 5.8 they are independent.

d. If $\alpha \neq 0$, $F(y_1,y_2)\neq F_1(y_1)F_2(y_2)$, so by Definition 5.8 they are not independent.

5.67 $P(a < Y_1 \le b, c < Y_2 \le d) = F(b,d) - F(b,c) - F(a,d) + F(a,c)$
$= F_1(b)F_2(d) - F_1(b)F_2(c) - F_1(a)F_2(d) + F_1(a)F_2(c)$
$= F_1(b)[F_2(d) - F_2(c)] - F_1(a)[F_2(d) - F_2(c)]$
$= [F_1(b) - F_1(a)] \times [F_2(d) - F_2(c)]$
$= P(a < Y_1 \le b) \times P(c < Y_2 \le d)$.

5.68 Given that $p_1(y_1) = \binom{2}{y_1}(.2)^{y_1}(.8)^{2-y_1}$, y_1 = 0, 1, 2, and $p_2(y_2) = (.3)^{y_2}(.7)^{1-y_1}$, y_2 = 0, 1:

a. $p(y_1, y_2) = p_1(y_1)p_2(y_2) = \binom{2}{y_1}(.2)^{y_1}(.8)^{2-y_1}(.3)^{y_2}(.7)^{1-y_1}$, y_1 = 0, 1, 2 and y_2 = 0, 1.

b. The probability of interest is $P(Y_1 + Y_2 \le 1) = p(0, 0) + p(1, 0) + p(0, 1)$ = .864.

5.69 **a.** $f(y_1, y_2) = f_1(y_1)f_2(y_2) = (1/9)e^{-(y_1+y_2)/3}$, $y_1 > 0$, $y_2 > 0$.

b. $P(Y_1 + Y_2 \le 1) = \int_0^1 \int_0^{1-y_2} (1/9)e^{-(y_1+y_2)/3} dy_1 dy_2 = 1 - \frac{4}{3}e^{-1/3}$ = .0446.

5.70 With $f(y_1, y_2) = f_1(y_1)f_2(y_2) = 1$, $0 \le y_1 \le 1$, $0 \le y_2 \le 1$,

$$P(Y_2 \le Y_1 \le Y_2 + 1/4) = \int_0^{1/4}\int_0^{y_1} 1 dy_2 dy_1 + \int_{1/4}^{1}\int_{y_1-1/4}^{y_1} 1 dy_2 dy_1 = 7/32.$$

5.71 Assume uniform distributions for the call times over the 1–hour period. Then,

a. $P(Y_1 \le 1/2, Y_2 \le 1/2) = P(Y_1 \le 1/2 P(Y_2 \le 1/2) = (1/2)(1/2) = 1/4$.

b. Note that 5 minutes = 1/12 hour. To find $P(|Y_1 - Y_2| \le 1/12)$, we must break the region into three parts in the integration:

$$P(|Y_1 - Y_2| \le 1/12) = \int_0^{1/12}\int_0^{y_1+1/12} 1 dy_2 dy_1 + \int_{1/12}^{11/12}\int_{y_1-1/12}^{y_1+1/12} 1 dy_2 dy_1 + \int_{11/12}^{1}\int_{y_1-1/12}^{1} 1 dy_2 dy_1 = 23/144.$$

5.72 **a.** $E(Y_1) = 2(1/3) = 2/3$.
b. $V(Y_1) = 2(1/3)(2/3) = 4/9$
c. $E(Y_1 - Y_2) = E(Y_1) - E(Y_2) = 0$.

5.73 Use the mean of the hypergeometric: $E(Y_1) = 3(4)/9 = 4/3$.

5.74 The marginal distributions for Y_1 and Y_2 are uniform on the interval (0, 1). And it was found in Ex. 5.50 that Y_1 and Y_2 are independent. So:

a. $E(Y_1 - Y_2) = E(Y_1) - E(Y_2) = 0$.
b. $E(Y_1Y_2) = E(Y_1)E(Y_2) = (1/2)(1/2) = 1/4$.
c. $E(Y_1^2 + Y_2^2) = E(Y_1^2) + E(Y_2^2) = (1/12 + 1/4) + (1/12 + 1/4) = 2/3$
d. $V(Y_1Y_2) = V(Y_1)V(Y_2) = (1/12)(1/12) = 1/144$.

5.75 The marginal distributions for Y_1 and Y_2 are exponential with $\beta = 1$. And it was found in Ex. 5.51 that Y_1 and Y_2 are independent. So:

a. $E(Y_1 + Y_2) = E(Y_1) + E(Y_2) = 2$, $V(Y_1 + Y_2) = V(Y_1) + V(Y_2) = 2$.

b. $P(Y_1 - Y_2 > 3) = P(Y_1 > 3 + Y_2) = \int_0^\infty \int_{3+y_2}^\infty e^{-y_1-y_2} dy_1 dy_2$ =(1/2)e^{-3} = .0249.

c. $P(Y_1 - Y_2 < -3) = P(Y_1 > Y_2 - 3) = \int_0^\infty \int_{3+y_1}^\infty e^{-y_1-y_2} dy_2 dy_1$ =(1/2)e^{-3} = .0249.

d. $E(Y_1 - Y_2) = E(Y_1) - E(Y_2) = 0$, $V(Y_1 - Y_2) = V(Y_1) + V(Y_2) = 2$.

e. They are equal.

5.76 From Ex. 5.52, we found that Y_1 and Y_2 are independent. So,

a. $E(Y_1) = \int_0^1 2y_1^2 dy_1 = 2/3$.

b. $E(Y_1^2) = \int_0^1 2y_1^3 dy_1 = 2/4$, so $V(Y_1) = 2/4 - 4/9 = 1/18$.

c. $E(Y_1 - Y_2) = E(Y_1) - E(Y_2) = 0$.

5.77 Following Ex. 5.27, the marginal densities can be used:

a. $E(Y_1) = \int_0^1 3y_1(1-y_1)^2 dy_1 = 1/4$, $E(Y_2) = \int_0^1 6y_2(1-y_2)dy_2 = 1/2$.

b. $E(Y_1^2) = \int_0^1 3y_1^2(1-y_1)^2 dy_1 = 1/10$, $V(Y_1) = 1/10 - (1/4)^2 = 3/80$,

$E(Y_2^2) = \int_0^1 6y_2^2(1-y_2)dy_2 = 3/10$, $V(Y_2) = 3/10 - (1/2)^2 = 1/20$.

c. $E(Y_1 - 3Y_2) = E(Y_1) - 3\cdot E(Y_2) = 1/4 - 3/2 = -5/4$.

5.78 **a.** The marginal distribution for Y_1 is $f_1(y_1) = y_1/2$, $0 \le y_1 \le 2$. $E(Y_1) = 4/3$, $V(Y_1) = 2/9$.
b. Similarly, $f_2(y_2) = 2(1 - y_2)$, $0 \le y_2 \le 1$. So, $E(Y_2) = 1/3$, $V(Y_1) = 1/18$.
c. $E(Y_1 - Y_2) = E(Y_1) - E(Y_2) = 4/3 - 1/3 = 1$.
d. $V(Y_1 - Y_2) = E[(Y_1 - Y_2)^2] - [E(Y_1 - Y_2)]^2 = E(Y_1^2) - 2E(Y_1Y_2) + E(Y_2^2) - 1$.

Since $E(Y_1Y_2) = \int_0^1 \int_{2y_2}^2 y_1 y_2 dy_1 dy_2 = 1/2$, we have that

$$V(Y_1 - Y_2) = [2/9 + (4/3)^2] - 1 + [1/18 + (1/3)^2] - 1 = 1/6.$$

Using Tchebysheff's theorem, two standard deviations about the mean is (.19, 1.81).

5.79 Referring to Ex. 5.16, integrating the joint density over the two regions of integration:

$$E(Y_1Y_2) = \int_{-1}^{0}\int_{0}^{1+y_1} y_1 y_2 dy_2 dy_1 + \int_{0}^{1}\int_{0}^{1-y_1} y_1 y_2 dy_2 dy_1 = 0$$

5.80 From Ex. 5.36, $f_1(y_1) = y_1 + \frac{1}{2}$, $0 \le y_1 \le 1$, and $f_2(y_2) = y_2 + \frac{1}{2}$, $0 \le y_2 \le 1$. Thus, $E(Y_1) = 7/12$ and $E(Y_2) = 7/12$. So, $E(30Y_1 + 25Y_2) = 30(7/12) + 25(7/12) = 32.08$.

5.81 Since Y_1 and Y_2 are independent, $E(Y_2/Y_1) = E(Y_2)E(1/Y_1)$. Thus, using the marginal densities found in Ex. 5.61,

$$E(Y_2/Y_1) = E(Y_2)E(1/Y_1) = \frac{1}{2}\int_{0}^{\infty} y_2 e^{-y_2/2} dy_2 \left[\frac{1}{4}\int_{0}^{\infty} e^{-y_1/2} dy_1\right] = 2(\tfrac{1}{2}) = 1.$$

5.82 The marginal densities were found in Ex. 5.34. So,

$$E(Y_1 - Y_2) = E(Y_1) - E(Y_2) = 1/2 - \int_{0}^{1} -y_2 \ln(y_2) dy_2 = 1/2 - 1/4 = 1/4.$$

5.83 From Ex. 3.88 and 5.42, $E(Y) = 2 - 1 = 1$.

5.84 All answers use results proven for the geometric distribution and independence:

a. $E(Y_1) = E(Y_2) = 1/p$, $E(Y_1 - Y_2) = E(Y_1) - E(Y_2) = 0$.
b. $E(Y_1^2) = E(Y_2^2) = (1 - p)/p^2 + (1/p)^2 = (2 - p)/p^2$. $E(Y_1Y_2) = E(Y_1)E(Y_2) = 1/p^2$.
c. $E[(Y_1 - Y_2)^2] = E(Y_1^2) - 2E(Y_1Y_2) + E(Y_2^2) = 2(1 - p)/p^2$.
$V(Y_1 - Y_2) = V(Y_1) + V(Y_2) = 2(1 - p)/p^2$.
d. Use Tchebysheff's theorem with $k = 3$.

5.85 **a.** $E(Y_1) = E(Y_2) = 1$ (both marginal distributions are exponential with mean 1)
b. $V(Y_1) = V(Y_2) = 1$
c. $E(Y_1 - Y_2) = E(Y_1) - E(Y_2) = 0$.
d. $E(Y_1Y_2) = 1 - \alpha/4$, so $\text{Cov}(Y_1, Y_2) = -\alpha/4$.
e. $V(Y_1 - Y_2) = V(Y_1) + V(Y_2) - 2\text{Cov}(Y_1, Y_2) = 1 + \alpha/2$. Using Tchebysheff's theorem with $k = 2$, the interval is $(-2\sqrt{2+\alpha/2},\ 2\sqrt{2+\alpha/2})$.

5.86 Using the hint and Theorem 5.9:

a. $E(W) = E(Z)E(Y_1^{-1/2}) = 0E(Y_1^{-1/2}) = 0$. Also, $V(W) = E(W^2) - [E(W)]^2 = E(W^2)$.
Now, $E(W^2) = E(Z^2)E(Y_1^{-1}) = 1 \cdot E(Y_1^{-1}) = E(Y_1^{-1}) = \frac{1}{\nu_1 - 2}$, $\nu_1 > 2$ (using Ex. 4.82).

b. $E(U) = E(Y_1)E(Y_2^{-1}) = \frac{\nu_1}{\nu_2 - 2}$, $\nu_2 > 2$, $V(U) = E(U^2) - [E(U)]^2 = E(Y_1^2)E(Y_2^{-2}) - \left(\frac{\nu_1}{\nu_2-2}\right)^2$

$= \nu_1(\nu_1 + 2)\frac{1}{(\nu_2-2)(\nu_2-4)} - \left(\frac{\nu_1}{\nu_2-2}\right)^2 = \frac{2\nu_1(\nu_{\cdot} + \nu_2 - 2)}{(\nu_2-2)^2(\nu_2-4)}$, $\nu_2 > 4$.

5.87 **a.** $E(Y_1 + Y_2) = E(Y_1) + E(Y_2) = \nu_1 + \nu_2$.
b. By independence, $V(Y_1 + Y_2) = V(Y_1) + V(Y_2) = 2\nu_1 + 2\nu_2$.

5.88 It is clear that $E(Y) = E(Y_1) + E(Y_2) + \ldots + E(Y_6)$. Using the result that Yi follows a geometric distribution with success probability $(7 - i)/6$, we have

$$E(Y) = \sum_{i=1}^{6} \frac{6}{7-i} = 1 + 6/5 + 6/4 + 6/3 + 6/2 + 6 = 14.7.$$

5.89 $\text{Cov}(Y_1, Y_2) = E(Y_1Y_2) - E(Y_1)E(Y_2) = \sum_{y_1}\sum_{y_2} y_1 y_2 p(y_1, y_2) - [2(1/3)]^2 = 2/9 - 4/9 = -2/9.$

As the value of Y_1 increases, the value of Y_2 tends to decrease.

5.90 From Ex. 5.3 and 5.21, $E(Y_1) = 4/3$ and $E(Y_2) = 1$. Thus,

$$E(Y_1Y_2) = 1(1)\tfrac{24}{84} + 2(1)\tfrac{12}{84} + 1(2)\tfrac{18}{84} = 1$$

So, $\text{Cov}(Y_1, Y_2) = E(Y_1Y_2) - E(Y_1)E(Y_2) = 1 - (4/3)(1) = -1/3$.

5.91 From Ex. 5.76, $E(Y_1) = E(Y_2) = 2/3$. $E(Y_1Y_2) = \int_0^1\int_0^1 4y_1^2y_2^2\,dy_1dy_2 = 4/9$. So,

$\text{Cov}(Y_1, Y_2) = E(Y_1Y_2) - E(Y_1)E(Y_2) = 4/9 - 4/9 = 0$ as expected since Y_1 and Y_2 are independent.

5.92 From Ex. 5.77, $E(Y_1) = 1/4$ and $E(Y_2) = 1/2$. $E(Y_1Y_2) = \int_0^1\int_0^{y_2} 6y_1y_2(1 - y_2)\,dy_1dy_2 = 3/20$.

So, $\text{Cov}(Y_1, Y_2) = E(Y_1Y_2) - E(Y_1)E(Y_2) = 3/20 - 1/8 = 1/40$ as expected since Y_1 and Y_2 are dependent.

5.93 **a.** From Ex. 5.55 and 5.79, $E(Y_1Y_2) = 0$ and $E(Y_1) = 0$. So,

$$\text{Cov}(Y_1, Y_2) = E(Y_1Y_2) - E(Y_1)E(Y_2) = 0 - 0E(Y_2) = 0.$$

b. Y_1 and Y_2 are dependent.
c. Since $\text{Cov}(Y_1, Y_2) = 0$, $\rho = 0$.
d. If $\text{Cov}(Y_1, Y_2) = 0$, Y_1 and Y_2 are not necessarily independent.

5.94 **a.**

$$\begin{aligned}\text{Cov}(U_1, U_2) &= E[(Y_1 + Y_2)(Y_1 - Y_2)] - E(Y_1 + Y_2)E(Y_1 - Y_2)\\ &= E(Y_1^2) - E(Y_2^2) - [E(Y_1)]^2 - [E(Y_2)]^2\\ &= (\sigma_1^2 + \mu_1^2) - (\sigma_2^2 + \mu_2^2) - (\mu_1^2 - \mu_2^2) = \sigma_1^2 - \sigma_2^2.\end{aligned}$$

b. Since $V(U_1) = V(U_2) = \sigma_1^2 + \sigma_2^2$ (Y_1 and Y_2 are uncorrelated), $\rho = \dfrac{\sigma_1^2 - \sigma_2^2}{\sigma_1^2 + \sigma_2^2}$.

c. If $\sigma_1^2 = \sigma_2^2$, U_1 and U_2 are uncorrelated.

5.95 Note that the marginal distributions for Y_1 and Y_2 are

y_1	–1	0	1
$p_1(y_1)$	1/3	1/3	1/3

y_2	0	1
$p_2(y_2)$	2/3	1/3

So, Y_1 and Y_2 not independent since $p(-1, 0) \neq p_1(-1)p_2(0)$. However, $E(Y_1) = 0$ and $E(Y_1Y_2) = (-1)(0)1/3 + (0)(1)(1/3) + (1)(0)(1/3) = 0$, so $\text{Cov}(Y_1, Y_2) = 0$.

5.96 **a.** $\text{Cov}(Y_1, Y_2) = E[(Y_1 - \mu_1)(Y_2 - \mu_2)] = E[(Y_2 - \mu_2)(Y_1 - \mu_1)] = \text{Cov}(Y_2, Y_1)$.
b. $\text{Cov}(Y_1, Y_1) = E[(Y_1 - \mu_1)(Y_1 - \mu_1)] = E[(Y_1 - \mu_1)^2] = V(Y_1)$.

5.97 **a.** From Ex. 5.96, $\text{Cov}(Y_1, Y_1) = V(Y_1) = 2$.
b. If $\text{Cov}(Y_1, Y_2) = 7$, $\rho = 7/4 > 1$, impossible.
c. With $\rho = 1$, $\text{Cov}(Y_1, Y_2) = 1(4) = 4$ (a perfect positive linear association).
d. With $\rho = -1$, $\text{Cov}(Y_1, Y_2) = -1(4) = -4$ (a perfect negative linear association).

5.98 Since $\rho^2 \leq 1$, we have that $-1 \leq \rho \leq 1$ or $-1 \leq \dfrac{\text{Cov}(Y_1, Y_2)}{\sqrt{V(Y_1)}\sqrt{V(Y_2)}} \leq 1$.

5.99 Since $E(c) = c$, $\text{Cov}(c, Y) = E[(c - c)(Y - \mu)] = 0$.

5.100 **a.** $E(Y_1) = E(Z) = 0$, $E(Y_2) = E(Z^2) = 1$.
b. $E(Y_1Y_2) = E(Z^3) = 0$ (odd moments are 0).
c. $\text{Cov}(Y_1, Y_1) = E(Z^3) - E(Z)E(Z^2) = 0$.
d. $P(Y_2 > 1 \mid Y_1 > 1) = P(Z^2 > 1 \mid Z > 1) = 1 \neq P(Z^2 > 1)$. Thus, Y_1 and Y_2 are dependent.

5.101 **a.** $\text{Cov}(Y_1, Y_2) = E(Y_1Y_2) - E(Y_1)E(Y_2) = 1 - \alpha/4 - (1)(1) = -\dfrac{\alpha}{4}$.
b. This is clear from part a.
c. We showed previously that Y_1 and Y_2 are independent only if $\alpha = 0$. If $\rho = 0$, if must be true that $\alpha = 0$.

5.102 The quantity $3Y_1 + 5Y_2$ = dollar amount spend per week. Thus:
$E(3Y_1 + 5Y_2) = 3(40) + 5(65) = 445$.
$E(3Y_1 + 5Y_2) = 9V(Y_1) + 25V(Y_2) = 9(4) + 25(8) = 236$.

5.103 $E(3Y_1 + 4Y_2 - 6Y_3) = 3E(Y_1) + 4E(Y_2) - 6E(Y_3) = 3(2) + 4(-1) - 6(-4) = -22$,
$V(3Y_1 + 4Y_2 - 6Y_3) = 9V(Y_1) + 16V(Y_2) + 36E(Y_3) + 24\text{Cov}(Y_1, Y_2) - 36\text{Cov}(Y_1, Y_3) -$
$48\text{Cov}(Y_2, Y_3) = 9(4) + 16(6) + 36(8) + 24(1) - 36(-1) - 48(0) = 480$.

5.104 **a.** Let $X = Y_1 + Y_2$. Then, the probability distribution for X is

x	1	2	3
$p(x)$	7/84	42/84	35/84

Thus, $E(X) = 7/3$ and $V(X) = .3889$.

b. $E(Y_1 + Y_2) = E(Y_1) + E(Y_2) = 4/3 + 1 = 7/3$. We have that $V(Y_1) = 10/18$, $V(Y_2) = 42/84$, and $\text{Cov}(Y_1, Y_1) = -1/3$, so

$V(Y_1 + Y_2) = V(Y_1) + V(Y_2) + 2\text{Cov}(Y_2, Y_3) = 10/18 + 42/84 - 2/3 = 7/18 = .3889.$

5.105 Since Y_1 and Y_2 are independent, $V(Y_1 + Y_2) = V(Y_1) + V(Y_1) = 1/18 + 1/18 = 1/9$.

5.106 $V(Y_1 - 3Y_2) = V(Y_1) + 9V(Y_2) - 6\text{Cov}(Y_1, Y_2) = 3/80 + 9(1/20) - 6(1/40) = 27/80 = .3375$.

5.107 Since $E(Y_1) = E(Y_2) = 1/3$, $V(Y_1) = V(Y_2) = 1/18$ and $E(Y_1Y_2) = \int_0^1 \int_0^{1-y_2} 2y_1y_2 dy_1 dy_2 = 1/12$, we have that $\text{Cov}(Y_{1,} Y_1) = 1/12 - 1/9 = -1/36$. Therefore,

$$E(Y_1 + Y_2) = 1/3 + 1/3 = 2/3 \text{ and } V(Y_1 + Y_2) = 1/18 + 1/18 + 2(-1/36) = 1/18.$$

5.108 From Ex. 5.33, Y_1 has a gamma distribution with $\alpha = 2$ and $\beta = 1$, and Y_2 has an exponential distribution with $\beta = 1$. Thus, $E(Y_1 + Y_2) = 2(1) + 1 = 3$. Also, since

$$E(Y_1Y_2) = \int_0^\infty \int_0^{y_1} y_1 y_2 e^{-y_1} dy_2 dy_1 = 3, \ \text{Cov}(Y_{1,} Y_1) = 3 - 2(1) = 1,$$

$$V(Y_1 - Y_2) = 2(1)^2 + 1^2 - 2(1) = 1.$$

Since a value of 4 minutes is four three standard deviations above the mean of 1 minute, this is not likely.

5.109 We have $E(Y_1) = E(Y_2) = 7/12$. Intermediate calculations give $V(Y_1) = V(Y_2) = 11/144$. Thus, $E(Y_1Y_2) = \int_0^1 \int_0^1 y_1 y_2 (y_1 + y_2) dy_1 dy_2 = 1/3$, $\text{Cov}(Y_{1,} Y_1) = 1/3 - (7/12)^2 = -1/144$.

From Ex. 5.80, $E(30Y_1 + 25Y_2) = 32.08$, so

$$V(30Y_1 + 25Y_2) = 900V(Y_1) + 625V(Y_2) + 2(30)(25)\,\text{Cov}(Y_{1,} Y_1) = 106.08.$$

The standard deviation of $30Y_1 + 25Y_2$ is $\sqrt{106.08} = 10.30$. Using Tchebysheff's theorem with $k = 2$, the interval is (11.48, 52.68).

5.110 **a.** $V(1 + 2Y_1) = 4V(Y_1)$, $V(3 + 4Y_2) = 16V(Y_2)$, and $\text{Cov}(1 + 2Y_1, 3 + 4Y_2) = 8\text{Cov}(Y_1, Y_2)$.

So, $\dfrac{8\text{Cov}(Y_1, Y_2)}{\sqrt{4V(Y_1)}\sqrt{16V(Y_2)}} = \rho = .2$.

b. $V(1 + 2Y_1) = 4V(Y_1)$, $V(3 - 4Y_2) = 16V(Y_2)$, and $\text{Cov}(1 + 2Y_1, 3 - 4Y_2) = -8\text{Cov}(Y_1, Y_2)$.

So, $\dfrac{-8\text{Cov}(Y_1, Y_2)}{\sqrt{4V(Y_1)}\sqrt{16V(Y_2)}} = -\rho = -.2$.

c. $V(1 - 2Y_1) = 4V(Y_1)$, $V(3 - 4Y_2) = 16V(Y_2)$, and $\text{Cov}(1 - 2Y_1, 3 - 4Y_2) = 8\text{Cov}(Y_1, Y_2)$.

So, $\dfrac{8\text{Cov}(Y_1, Y_2)}{\sqrt{4V(Y_1)}\sqrt{16V(Y_2)}} = \rho = .2$.

5.111 **a.** $V(a + bY_1) = b^2V(Y_1)$, $V(c + dY_2) = d^2V(Y_2)$, and $\text{Cov}(a + bY_1, c + dY_2) = bd\text{Cov}(Y_1, Y_2)$. So, $\rho_{W_1,W_2} = \dfrac{bd\text{Cov}(Y_1,Y_2)}{\sqrt{b^2V(Y_1)}\sqrt{d^2V(Y_2)}} = \dfrac{bd}{|bd|}\rho_{Y_1,Y_2}$. Provided that the constants b and d are nonzero, $\dfrac{bd}{|bd|}$ is either 1 or –1. Thus, $|\rho_{W_1,W_2}| = |\rho_{Y_1,Y_2}|$.

b. Yes, the answers agree.

5.112 In Ex. 5.61, it was showed that Y_1 and Y_2 are independent. In addition, Y_1 has a gamma distribution with $\alpha = 2$ and $\beta = 2$, and Y_2 has an exponential distribution with $\beta = 2$. So, with $C = 50 + 2Y_1 + 4Y_2$, it is clear that

$E(C) = 50 + 2E(Y_1) + 4E(Y_2) = 50 + (2)(4) + (4)(2) = 66$
$V(C) = 4V(Y_1) + 16V(Y_2) = 4(2)(4) + 16(4) = 96.$

5.113 The net daily gain is given by the random variable $G = X – Y$. Thus, given the distributions for X and Y in the problem,

$E(G) = E(X) – E(Y) = 50 – (4)(2) = 42$
$V(G) = V(G) + V(G) = 3^2 + 4(2^2) = 25.$

The value \$70 is (70 – 42)/5 = 7.2 standard deviations above the mean, an unlikely value.

5.114 Observe that Y_1 has a gamma distribution with $\alpha = 4$ and $\beta = 1$ and Y_2 has an exponential distribution with $\beta = 2$. Thus, with $U = Y_1 – Y_2$,

a. $E(U) = 4(1) – 2 = 2$
b. $V(U) = 4(1^2) + 2^2 = 8$
c. The value 0 has a z–score of $(0 – 2)/\sqrt{8} = –.707$, or it is –.707 standard deviations *below* the mean. This is not extreme so it is likely the profit drops below 0.

5.115 Following Ex. 5.88:
a. Note that for non–negative integers a and b and $i \neq j$,

$$P(Y_i = a, Y_j = b) = P(Y_j = b \mid Y_i = a)P(Y_i = a)$$

But, $P(Y_j = b \mid Y_i = a) = P(Y_j = b)$ since the trials (i.e. die tosses) are independent — the experiments that generate Y_i and Y_j represent independent experiments via the memoryless property. So, Y_i and Y_j are independent and thus $\text{Cov}(Y_i. Y_j) = 0$.

b. $V(Y) = V(Y_1) + \ldots + V(Y_6) = 0 + \frac{1/6}{(5/6)^2} + \frac{2/6}{(4/6)^2} + \frac{3/6}{(3/6)^2} + \frac{4/6}{(2/6)^2} + \frac{5/6}{(1/6)^2} = 38.99.$

c. From Ex. 5.88, E(Y) = 14.7. Using Tchebysheff's theorem with $k = 2$, the interval is $14.7 \pm 2\sqrt{38.99}$ or (0 , 27.188)

5.116 $V(Y_1 + Y_2) = V(Y_1) + V(Y_2) + 2\text{Cov}(Y_1, Y_2)$, $V(Y_1 - Y_2) = V(Y_1) + V(Y_2) - 2\text{Cov}(Y_1, Y_2)$. When Y_1 and Y_2 are independent, $\text{Cov}(Y_1, Y_2) = 0$ so the quantities are the same.

5.117 Refer to Example 5.29 in the text. The situation here is analogous to drawing n balls from an urn containing N balls, r_1 of which are red, r_2 of which are black, and $N - r_1 - r_2$ are neither red nor black. Using the argument given there, we can deduce that:

$E(Y_1) = np_1$ $\quad V(Y_1) = np_1(1 - p_1)\left(\frac{N-n}{N-1}\right)$ $\quad$ where $p_1 = r_1/N$

$E(Y_2) = np_2$ $\quad V(Y_2) = np_2(1 - p_2)\left(\frac{N-n}{N-1}\right)$ $\quad$ where $p_2 = r_2/N$

Now, define new random variables for $i = 1, 2, \ldots, n$:

$$U_i = \begin{cases} 1 & \text{if alligator } i \text{ is a mature female} \\ 0 & \text{otherwise} \end{cases} \qquad V_i = \begin{cases} 1 & \text{if alligator } i \text{ is a mature male} \\ 0 & \text{otherwise} \end{cases}$$

Then, $Y_1 = \sum_{i=1}^{n} U_i$ and $Y_2 = \sum_{i=1}^{n} V_i$. Now, we must find $\text{Cov}(Y_1, Y_2)$. Note that:

$$E(Y_1Y_2) = E\left(\sum_{i=1}^{n} U_i , \sum_{i=1}^{n} V_i\right) = \sum_{i=1}^{n} E(U_iV_i) + \sum_{i \neq j} E(U_iV_j).$$

Now, since for all i, $E(U_i, V_i) = P(U_i = 1, V_i = 1) = 0$ (an alligator can't be both female and male), we have that $E(U_i, V_i) = 0$ for all i. Now, for $i \neq j$,

$$E(U_i, V_j) = P(U_i = 1, V_i = 1) = P(U_i = 1)P(V_i = 1|U_i = 1) = \tfrac{r_1}{N}\left(\tfrac{r_2}{N-1}\right) = \tfrac{N}{N-1} p_1 p_2 .$$

Since there are $n(n - 1)$ terms in $\sum_{i \neq j} E(U_iV_j)$, we have that $E(Y_1Y_2) = n(n - 1)\frac{N}{N-1} p_1 p_2$.

Thus, $\qquad \text{Cov}(Y_1, Y_2) = n(n - 1)\frac{N}{N-1} p_1 p_2 - (np_1)(np_2) = -\frac{n(N-n)}{N-1} p_1 p_2$.

So, $\qquad E\left[\frac{Y_1}{n} - \frac{Y_2}{n}\right] = \frac{1}{n}\left(np_1 - np_2\right) = p_1 - p_2$,

$$V\left[\frac{Y_1}{n} - \frac{Y_2}{n}\right] = \frac{1}{n^2}\left[V(Y_1) + V(Y_2) - 2\text{Cov}(Y_1, Y_2)\right] = \frac{N-n}{n(N-1)}\left(p_1 + p_2 - (p_1 - p_2)^2\right)$$

5.118 Let $Y = X_1 + X_2$, the total sustained load on the footing.

a. Since X_1 and X_2 have gamma distributions and are independent, we have that

$E(Y) = 50(2) + 20(2) = 140$

$V(Y) = 50(2^2) + 20(2^2) = 280$.

b. Consider Tchebysheff's theorem with $k = 4$: the corresponding interval is

$$140 + 4\sqrt{280} \text{ or } (73.07, 206.93).$$

So, we can say that the sustained load will exceed 206.93 kips with probability less than 1/16.

5.119 **a.** Using the multinomial distribution with $p_1 = p_2 = p_3 = 1/3$,

$$P(Y_1 = 3, Y_2 = 1, Y_3 = 2) = \tfrac{6!}{3!1!2!}\left(\tfrac{1}{3}\right)^6 = .0823.$$

b. $E(Y_1) = n/3$, $V(Y_1) = n(1/3)(2/3) = 2n/9$.

c. $\text{Cov}(Y_2, Y_3) = -n(1/3)(1/3) = -n/9$.

d. $E(Y_2 - Y_3) = n/3 - n/3 = 0$, $V(Y_2 - Y_3) = V(Y_2) + V(Y_3) - 2\text{Cov}(Y_2, Y_3) = 2n/3$.

5.120 $E(C) = E(Y_1) + 3E(Y_2) = np_1 + 3np_2$.
$V(C) = V(Y_1) + 9V(Y_2) + 6\text{Cov}(Y_1, Y_2) = np_1q_1 + 9np_2q_2 - 6np_1p_2$.

5.121 If N is large, the multinomial distribution is appropriate:

a. $P(Y_1 = 2, Y_2 = 1) = \tfrac{5!}{2!1!2!}(.3)^2(.1)^1(.6)^2 = .0972$.

b. $E\left[\tfrac{Y_1}{n} - \tfrac{Y_2}{n}\right] = \; p_1 - p_2 = .3 - .1 = .2$

$V\left[\tfrac{Y_1}{n} - \tfrac{Y_2}{n}\right] = \tfrac{1}{n^2}\left[V(Y_1) + V(Y_2) - 2\text{Cov}(Y_1, Y_2)\right] = \tfrac{p_1q_1}{n} + \tfrac{p_2q_2}{n} + 2\tfrac{p_1p_2}{n} = .072.$

5.122 Let Y_1 = # of mice weighing between 80 and 100 grams, and let Y_2 = # weighing over 100 grams. Thus, with X having a normal distribution with $\mu = 100$ g. and $\sigma = 20$ g.,

$p_1 = P(80 \le X \le 100) = P(-1 \le Z \le 0) = .3413$
$p_2 = P(X > 100) = P(Z > 0) = .5$

a. $P(Y_1 = 2, Y_2 = 1) = \tfrac{4!}{2!1!1!}(.3413)^2(.5)^1(.1587)^1 = .1109$.

b. $P(Y_2 = 4) = \tfrac{4!}{0!4!0!}(.5)^4 = .0625$.

5.123 Let Y_1 = # of family home fires, Y_2 = # of apartment fires, and Y_3 = # of fires in other types. Thus, (Y_1, Y_2, Y_3) is multinomial with $n = 4$, $p_1 = .73$, $p_2 = .2$ and $p_3 = .07$. Thus,

$$P(Y_1 = 2, Y_2 = 1, Y_3 = 1) = 6(.73)^2(.2)(.07) = .08953.$$

5.124 Define C = total cost = $20{,}000Y_1 + 10{,}000Y_2 + 2000Y_3$

a. $E(C) = 20{,}000E(Y_1) + 10{,}000E(Y_2) + 2000E(Y_3)$
$= 20{,}000(2.92) + 10{,}000(.8) + 2000(.28) = 66{,}960.$

b. $V(C) = (20{,}000)^2V(Y_1) + (10{,}000)^2V(Y_2) + (2000)^2V(Y_3)$ + covariance terms
$= (20{,}000)^2(4)(.73)(.27) + (10{,}000)^2(4)(.8)(.2) + (2000)^2(4)(.07)(.93)$
$+ 2[20{,}000(10{,}000)(-4)(.73)(.2) + 20{,}000(2000)(-4)(.73)(.07) +$
$10{,}000(2000)(-4)(.2)(.07)] = 380{,}401{,}600 - 252{,}192{,}000 = 128{,}209{,}600.$

5.125 Let Y_1 = # of planes with no wine cracks, Y_2 = # of planes with detectable wing cracks, and Y_3 = # of planes with critical wing cracks. Therefore, (Y_1, Y_2, Y_3) is multinomial with $n = 5$, $p_1 = .7$, $p_2 = .25$ and $p_3 = .05$.

a. $P(Y_1 = 2, Y_2 = 2, Y_3 = 1) = 30(.7)^2(.25)^2(.05) = .046.$

b. The distribution of Y_3 is binomial with $n = 5$, $p_3 = .05$, so

$$P(Y_3 \geq 1) = 1 - P(Y_3 = 0) = 1 - (.95)^5 = .2262.$$

5.126 Using formulas for means, variances, and covariances for the multinomial:

$E(Y_1) = 10(.1) = 1$ $\quad V(Y_1) = 10(.1)(.9) = .9$
$E(Y_2) = 10(.05) = .5$ $\quad V(Y_2) = 10(.05)(.95) = .475$
$\text{Cov}(Y_1, Y_2) = -10(.1)(.05) = -.05$

So,

$E(Y_1 + 3Y_2) = 1 + 3(.5) = 2.5$
$V(Y_1 + 3Y_2) = .9 + 9(.475) + 6(-.05) = 4.875.$

5.127 Y is binomial with $n = 10$, $p = .10 + .05 = .15$.

a. $P(Y = 2) = \binom{10}{2}(.15)^2(.85)^8 = .2759.$

b. $P(Y \geq 1) = 1 - \text{P}(Y = 0) = 1 - (.85)^{10} = .8031.$

5.128 The marginal distribution for Y_1 is found by

$$f_1(y_1) = \int_{-\infty}^{\infty} f(y_1, y_2)dy_2 .$$

Making the change of variables $u = (y_1 - \mu_1)/\sigma_1$ and $v = (y_2 - \mu_2)/\sigma_2$ yields

$$f_1(y_1) = \frac{1}{2\pi\sigma_1\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2(1-\rho^2)}(u^2 + v^2 - 2\rho uv)\right]dv .$$

To evaluate this, note that $u^2 + v^2 - 2\rho uv = (v - \rho u)^2 + u^2(1-\rho^2)$ so that

$$f_1(y_1) = \frac{1}{2\pi\sigma_1\sqrt{1-\rho^2}} e^{-u^2/2} \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2(1-\rho^2)}(v - \rho u)^2\right]dv ,$$

So, the integral is that of a normal density with mean ρu and variance $1 - \rho^2$. Therefore,

$$f_1(y_1) = \frac{1}{2\pi\sigma_1} e^{-(y_1-\mu_1)^2/2\sigma_1^2} , -\infty < y_1 < \infty,$$

which is a normal density with mean μ_1 and standard deviation σ_1. A similar procedure will show that the marginal distribution of Y_2 is normal with mean μ_2 and standard deviation σ_2.

5.129 The result follows from Ex. 5.128 and defining $f(y_1 \mid y_2) = f(y_1, y_2)/f_2(y_2)$, which yields a density function of a normal distribution with mean $\mu_1 + \rho(\sigma_1/\sigma_2)(y_2 - \mu_2)$ and variance $\sigma_1^2(1-\rho^2)$.

5.130 **a.** $\text{Cov}(U_1, U_2) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i b_j \text{Cov}(Y_i, Y_j) = \sum_{i=1}^{n} a_i b_j V(Y_i) = \sigma^2 \sum_{i=1}^{n} a_i b_j$, since the Y_i's are independent. If $\text{Cov}(U_1, U_2) = 0$, it must be true that $\sum_{i=1}^{n} a_i b_j = 0$ since $\sigma^2 > 0$. But, it is trivial to see if $\sum_{i=1}^{n} a_i b_j = 0$, $\text{Cov}(U_1, U_2) = 0$. So, U_1 and U_2 are orthogonal.

b. Given in the problem, (U_1, U_2) has a bivariate normal distribution. Note that $E(U_1) = \mu\sum_{i=1}^{n} a_i$, $E(U_2) = \mu\sum_{i=1}^{n} b_i$, $V(U_1) = \sigma^2\sum_{i=1}^{n} a_i^2$, and $V(U_2) = \sigma^2\sum_{i=1}^{n} b_i^2$. If they are orthogonal, $\text{Cov}(U_1, U_2) = 0$ and then $\rho_{U_1,U_2} = 0$. So, they are also independent.

5.131 **a.** The joint distribution of Y_1 and Y_2 is simply the product of the marginals $f_1(y_1)$ and $f_2(y_2)$ since they are independent. It is trivial to show that this product of density has the form of the bivariate normal density with $\rho = 0$.

b. Following the result of Ex. 5.130, let $a_1 = a_2 = b_1 = 1$ and $b_2 = -1$. Thus, $\sum_{i=1}^{n} a_i b_j = 0$ so U_1 and U_2 are independent.

5.132 Following Ex. 5.130 and 5.131, U_1 is normal with mean $\mu_1 + \mu_2$ and variance $2\sigma^2$ and U_2 is normal with mean $\mu_1 - \mu_2$ and variance $2\sigma^2$.

5.133 From Ex. 5.27, $f(y_1 | y_2) = 1/y_2$, $0 \le y_1 \le y_2$ and $f_2(y_2) = 6y_2(1 - y_2)$, $0 \le y_2 \le 1$.

a. To find $E(Y_1 | Y_2 = y_2)$, note that the conditional distribution of Y_1 given Y_2 is uniform on the interval $(0, y_2)$. So, $E(Y_1 | Y_2 = y_2) = \dfrac{y_2}{2}$.

b. To find $E(E(Y_1 | Y_2))$, note that the marginal distribution is beta with $\alpha = 2$ and $\beta = 2$. So, from part a, $E(E(Y_1 | Y_2)) = E(Y_2/2) = 1/4$. This is the same answer as in Ex. 5.77.

5.134 The z–score is $(6 - 1.25)/\sqrt{1.5625} = 3.8$, so the value 6 is 3.8 standard deviations above the mean. This is not likely.

5.135 Refer to Ex. 5.41:

a. Since Y is binomial, $E(Y|p) = 3p$. Now p has a uniform distribution on (0, 1), thus

$$E(Y) = E[E(Y|p)] = E(3p) = 3(1/2) = 3/2.$$

b. Following part a, $V(Y|p) = 3p(1 - p)$. Therefore,

$$V(p) = E[3p(1 - p)] + V(3p) = 3E(p - p^2) + 9V(p)$$
$$= 3E(p) - 3[V(p) + (E(p))^2] + 9V(p) = 1.25$$

5.136 **a.** For a given value of λ, Y has a Poisson distribution. Thus, $E(Y | \lambda) = \lambda$. Since the marginal distribution of λ is exponential with mean 1, $E(Y) = E[E(Y | \lambda)] = E(\lambda) = 1$.

b. From part a, $E(Y | \lambda) = \lambda$ and so $V(Y | \lambda) = \lambda$. So, $V(Y) = E[V(Y | \lambda)] + E[V(Y | \lambda)] = 2$

c. The value 9 is $(9 - 1)/\sqrt{2} = 5.657$ standard deviations above the mean (unlikely score).

5.137 Refer to Ex. 5.38: $E(Y_2 | Y_1 = y_1) = y_1/2$. For $y_1 = 3/4$, $E(Y_2 | Y_1 = 3/4) = 3/8$.

5.138 If Y = # of bacteria per cubic centimeter,

a. $E(Y) = E(Y) = E[E(Y | \lambda)] = E(\lambda) = \alpha\beta$.

b. $V(Y) = E[V(Y \mid \lambda)] + V[E(Y \mid \lambda)] = \alpha\beta + \alpha\beta^2 = \alpha\beta(1+\beta)$. Thus, $\sigma = \sqrt{\alpha\beta(1+\beta)}$.

5.139 **a.** $E(T \mid N = n) = E\left(\sum_{i=1}^{n} Y_i\right) = \sum_{i=1}^{n} E(Y_i) = n\alpha\beta$.

b. $E(T) = E[E(T \mid N)] = E(N\alpha\beta) = \lambda\alpha\beta$. Note that this is $E(N)E(Y)$.

5.140 Note that $V(Y_1) = E[V(Y_1 \mid Y_2)] + V[E(Y_1 \mid Y_2)]$, so $E[V(Y_1 \mid Y_2)] = V(Y_1) - V[E(Y_1 \mid Y_2)]$. Thus, $E[V(Y_1 \mid Y_2)] \le V(Y_1)$.

5.141 $E(Y_2) = E(E(Y_2 \mid Y_1)) = E(Y_1/2) = \dfrac{\lambda}{2}$

$V(Y_2) = E[V(Y_2 \mid Y_1)] + V[E(Y_2 \mid Y_1)] = E[Y_1^2/12] + V[Y_1/2] = (2\lambda^2)/12 + (\lambda^2)/2 = \dfrac{2\lambda^2}{3}$.

5.142 **a.** $E(Y) = E[E(Y|p)] = E(np) = nE(p) = \dfrac{n\alpha}{\alpha+\beta}$.

b. $V(Y) = E[V(Y \mid p)] + V[E(Y \mid p)] = E[np(1-p)] + V(np) = nE(p - p^2) + n^2V(p)$. Now:

$$nE(p - p^2) = \frac{n\alpha}{\alpha+\beta} - \frac{n\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)}$$

$$n^2V(p) = \frac{n^2\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}.$$

So, $V(Y) = \dfrac{n\alpha}{\alpha+\beta} - \dfrac{n\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)} + \dfrac{n^2\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} = \dfrac{n\alpha\beta(\alpha+\beta+n)}{(\alpha+\beta)^2(\alpha+\beta+1)}$.

5.143 Consider the random variable y_1Y_2 for the fixed value of Y_1. It is clear that y_1Y_2 has a normal distribution with mean 0 and variance y_1^2 and the mgf for this random variable is

$$m(t) = E(e^{ty_1Y_2}) = e^{t^2y_1^2/2}.$$

Thus, $m_U(t) = E(e^{tU}) = E(e^{tY_1Y_2}) = E[E(e^{tY_1Y_2} \mid Y_1)] = E(e^{tY_1^2/2}) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{(-y_1^2/2)(1-t^2)} dy_1$.

Note that this integral is essentially that of a normal density with mean 0 and variance $\frac{1}{1-t^2}$, so the necessary constant that makes the integral equal to 0 is the reciprocal of the standard deviation. Thus, $m_U(t) = (1-t^2)^{-1/2}$. Direct calculations give $m_U'(0) = 0$ and $m_U''(0) = 1$. To compare, note that $E(U) = E(Y_1Y_2) = E(Y_1)E(Y_2) = 0$ and $V(U) = E(U^2) = E(Y_1^2Y_2^2) = E(Y_1^2)E(Y_2^2) = (1)(1) = 1$.

5.144 $E[g(Y_1)h(Y_2)] = \sum_{y_1}\sum_{y_2} g(y_1)h(y_2)p(y_1,y_2) = \sum_{y_1}\sum_{y_2} g(y_1)h(y_2)p_1(y_1)p_2(y_2) =$
$\sum_{y_1} g(y_1)p_1(y_1)\sum_{y_2} h(y_2)p_2(y_2) = E[g(Y_1)]\times E[h(Y_2)]$.

5.145 The probability of interest is $P(Y_1 + Y_2 < 30)$, where Y_1 is uniform on the interval (0, 15) and Y_2 is uniform on the interval (20, 30). Thus, we have

$$P(Y_1 + Y_2 < 30) = \int_{20}^{30}\int_{0}^{30-y2}\left(\frac{1}{15}\right)\left(\frac{1}{10}\right)dy_1 dy_2 = 1/3.$$

5.146 Let (Y_1, Y_2) represent the coordinates of the landing point of the bomb. Since the radius is one mile, we have that $0 \le y_1^2 + y_2^2 \le 1$. Now,

P(target is destroyed) = P(bomb destroys everything within 1/2 of landing point)

This is given by $P(Y_1^2 + Y_2^2 \le (\frac{1}{2})^2)$. Since (Y_1, Y_2) are uniformly distributed over the unit circle, the probability in question is simply the area of a circle with radius 1/2 divided by the area of the unit circle, or simply 1/4.

5.147 Let Y_1 = arrival time for 1st friend, $0 \le y_1 \le 1$, Y_2 = arrival time for 2nd friend, $0 \le y_2 \le 1$. Thus $f(y_1, y_2) = 1$. If friend 2 arrives 1/6 hour (10 minutes) before or after friend 1, they will meet. We can represent this event as $|Y_1 - Y_2| < 1/3$. To find the probability of this event, we must find:

$$P(|Y_1 - Y_2| < 1/3) = \int_0^{1/6}\int_0^{y_1+1/6} 1dy_2dy_1 + \int_{1/6}^{5/6}\int_{y_1-1/6}^{y_1+1/6} 1dy_2dy_1 + \int_{5/6}^{1}\int_{y_1-1/6}^{1} 1dy_2dy_1 = 11/36.$$

5.148 **a.** $p(y_1, y_2) = \dfrac{\binom{4}{y_1}\binom{3}{y_2}\binom{2}{3-y_1-y_2}}{\binom{9}{3}}$, $y_1 = 0, 1, 2, 3$, $y_2 = 0, 1, 2, 3$, $y_1 + y_2 \le 3$.

b. Y_1 is hypergeometric w/ $r = 4$, $N = 9$, $n = 3$; Y_2 is hypergeometric w/ $r = 3$, $N = 9$, $n = 3$

c. $P(Y_1 = 1 \mid Y_2 \ge 1) = [p(1, 1) + p(1, 2)]/[1 - p_2(0)] = 9/16$

5.149 **a.** $f_1(y_1) = \int_0^{y_1} 3y_1 dy_2 = 3y_1^2$, $0 \le y_1 \le 1$, $f_1(y_1) = \int_{y_2}^{1} 3y_1 dy_1 = \frac{3}{2}(1 - y_2^2)$, $0 \le y_2 \le 1$.

b. $P(Y_1 \le 3/4 \mid Y_2 \le 1/2) = 23/44$.

c. $f(y_1 \mid y_2) = 2y_1/(1 - y_2^2)$, $y_2 \le y_1 \le 1$.

d. $P(Y_1 \le 3/4 \mid Y_2 = 1/2) = 5/12$.

5.150 **a.** Note that $f(y_2 \mid y_1) = f(y_1, y_2)/f(y_1) = 1/y_1$, $0 \le y_2 \le y_1$. This is the same conditional density as seen in Ex. 5.38 and Ex. 5.137. So, $E(Y_2 \mid Y_1 = y_1) = y_1/2$.

b. $E(Y_2) = E[E(Y_2 \mid Y_1)] = E(Y_1/2) = \int_0^1 \frac{y_1}{2} 3y_1^2 dy_1 = 3/8$.

c. $E(Y_2) = \int_0^1 y_2 \frac{3}{2}(1 - y_2^2) dy_2 = 3/8$.

5.151 **a.** The joint density is the product of the marginals: $f(y_1, y_2) = \frac{1}{\beta^2} e^{-(y_1+y_2)/\beta}$, $y_1 \geq \infty$, $y_2 \geq \infty$

b. $P(Y_1 + Y_2 \leq a) = \int_0^a \int_0^{a-y_2} \frac{1}{\beta^2} e^{-(y_1+y_2)/\beta} dy_1 dy_2 = 1 - [1 + a/\beta] e^{-a/\beta}$.

5.152 The joint density of (Y_1, Y_2) is $f(y_1, y_2) = 18(y_1 - y_1^2) y_2^2$, $0 \leq y_1 \leq 1$, $0 \leq y_2 \leq 1$. Thus, $P(Y_1 Y_2 \leq .5) = P(Y_1 \leq .5/Y_2) = 1 - P(Y_1 > .5/Y_2) = 1 - \int_{.5}^{1} \int_{.5/y_2}^{1} 18(y_1 - y_1^2) y_2^2 dy_1 dy_2$. Using straightforward integration, this is equal to (5 – 3ln2)/4 = .73014.

5.153 This is similar to Ex. 5.139:

a. Let N = # of eggs laid by the insect and Y = # of eggs that hatch. Given $N = n$, Y has a binomial distribution with n trials and success probability p. Thus, $E(Y \mid N = n) = np$. Since N follows as Poisson with parameter λ, $E(Y) = E[E(Y \mid N)] = E(Np) = \lambda p$.

b. $V(Y) = E[V(Y \mid N)] + V[E(Y \mid N)] = E[Np(1 - p)] + V[Np] = \lambda p$.

5.154 The conditional distribution of Y given p is binomial with parameter p, and note that the marginal distribution of p is beta with $\alpha = 3$ and $\beta = 2$.

a. Note that $f(y) = \int_0^1 f(y, p) = \int_0^1 f(y \mid p) f(p) dp = 12\binom{n}{y} \int_0^1 p^{y+2}(1-p)^{n-y+1} dp$. This integral can be evaluated by relating it to a beta density w/ $\alpha = y + 3$, $\beta = n + y + 2$. Thus,

$$f(y) = 12\binom{n}{y} \frac{\Gamma(n - y + 2)\Gamma(y + 3)}{\Gamma(n + 5)}, \; y = 0, 1, 2, \ldots, n.$$

b. For $n = 2$, $E(Y \mid p) = 2p$. Thus, $E(Y) = E[E(Y|p)] = E(2p) = 2E(p) = 2(3/5) = 6/5$.

5.155 **a.** It is easy to show that

$$\begin{aligned} \text{Cov}(W_1, W_2) &= \text{Cov}(Y_1 + Y_2, Y_1 + Y_3) \\ &= \text{Cov}(Y_1, Y_1) + \text{Cov}(Y_1, Y_3) + \text{Cov}(Y_2, Y_1) + \text{Cov}(Y_2, Y_3) \\ &= \text{Cov}(Y_1, Y_1) = \text{V}(Y_1) = 2\nu_1. \end{aligned}$$

b. It follows from part a above (i.e. the variance is positive).

5.156 **a.** Since $E(Z) = E(W) = 0$, $\text{Cov}(Z, W) = E(ZW) = E(Z^2 Y^{-1/2}) = E(Z^2)E(Y^{-1/2}) = E(Y^{-1/2})$. This expectation can be found by using the result Ex. 4.112 with $a = -1/2$. So,

$$\text{Cov}(Z, W) = E(Y^{-1/2}) = \frac{\Gamma(\frac{\nu}{2} - \frac{1}{2})}{\sqrt{2}\Gamma(\frac{\nu}{2})}, \text{ provided } \nu > 1.$$

b. Similar to part a, $\text{Cov}(Y, W) = E(YW) = E(\sqrt{Y}\,W) = E(\sqrt{Y}\,)E(W) = 0$.

c. This is clear from parts (a) and (b) above.

5.157 $p(y) = \int_0^\infty p(y \mid \lambda) f(\lambda) d\lambda = \int_0^\infty \frac{\lambda^{y+\alpha-1} e^{-\lambda[(\beta+1)/\beta]}}{\Gamma(y+1)\Gamma(\alpha)\beta^\alpha} d\lambda = \frac{\Gamma(y+\alpha)\left(\frac{\beta}{\beta+1}\right)^{y+\alpha}}{\Gamma(y+1)\Gamma(\alpha)\beta^\alpha}$, $y = 0, 1, 2, \ldots$. Since it was assumed that α was an integer, this can be written as

$$p(y) = \binom{y+\alpha-1}{y}\left(\frac{\beta}{\beta+1}\right)^y \left(\frac{1}{\beta+1}\right)^\alpha, \; y = 0, 1, 2, \ldots .$$

5.158 Note that for each X_i, $E(X_i) = p$ and $V(X_i) = pq$. Then, $E(Y) = \Sigma E(X_i) = np$ and $V(Y) = npq$. The second result follows from the fact that the X_i are independent so therefore all covariance expressions are 0.

5.159 For each W_i, $E(W_i) = 1/p$ and $V(W_i) = q/p^2$. Then, $E(Y) = \Sigma E(X_i) = r/p$ and $V(Y) = rq/p^2$. The second result follows from the fact that the W_i are independent so therefore all covariance expressions are 0.

5.160 The marginal probabilities can be written directly:

$P(X_1 = 1) = P(\text{select ball 1 or 2}) = .5$	$P(X_1 = 0) = .5$
$P(X_2 = 1) = P(\text{select ball 1 or 3}) = .5$	$P(X_2 = 0) = .5$
$P(X_3 = 1) = P(\text{select ball 1 or 4}) = .5$	$P(X_3 = 0) = .5$

Now, for $i \neq j$, X_i and X_j are clearly pairwise independent since, for example,

$P(X_1 = 1, X_2 = 1) = P(\text{select ball 1}) = .25 = P(X_1 = 1)P(X_2 = 1)$
$P(X_1 = 0, X_2 = 1) = P(\text{select ball 3}) = .25 = P(X_1 = 0)P(X_2 = 1)$

However, X_1, X_2, and X_3 are <u>not</u> mutually independent since

$P(X_1 = 1, X_2 = 1, X_3 = 1) = P(\text{select ball 1}) = .25 \neq P(X_1 = 1)P(X_2 = 1)P(X_1 = 3)$.

5.161 $E(\overline{Y}-\overline{X}) = E(\overline{Y}) - E(\overline{X}) = \frac{1}{n}\sum E(Y_i) - \frac{1}{m}\sum E(X_i) = \mu_1 - \mu_2$
$V(\overline{Y}-\overline{X}) = V(\overline{Y}) + V(\overline{X}) = \frac{1}{n^2}\sum V(Y_i) + \frac{1}{m^2}\sum V(X_i) = \sigma_1^2/n + \sigma_2^2/m$

5.162 Using the result from Ex. 5.65, choose two different values for α with $-1 \le \alpha \le 1$.

5.163 **a.** The distribution functions with the exponential distribution are:

$$F_1(y_1) = 1 - e^{-y_1},\ y_1 \ge 0; \qquad F_2(y_2) = 1 - e^{-y_2},\ y_2 \ge 0.$$

Then, the joint distribution function is

$$F(y_1, y_2) = [1 - e^{-y_1}][1 - e^{-y_2}][1 - \alpha(e^{-y_1})(e^{-y_2})].$$

Finally, show that $\frac{\partial^2}{\partial y_1 \partial y_2} F(y_1, y_2)$ gives the joint density function seen in Ex. 5.162.

b. The distribution functions with the uniform distribution on (0, 1) are:

$$F_1(y_1) = y_1,\ 0 \le y_1 \le 1; \qquad F_2(y_2) = y_2,\ 0 \le y_2 \le 1.$$

Then, the joint distribution function is

$$F(y_1, y_2) = y_1 y_2 [1 - \alpha(1 - y_1)(1 - y_2)].$$

c. $\frac{\partial^2}{\partial y_1 \partial y_2} F(y_1, y_2) = f(y_1, y_2) = 1 - \alpha[(1 - 2y_1)(1 - 2y_2)]$, $0 \le y_1 \le 1$, $0 \le y_2 \le 1$.

d. Choose two different values for α with $-1 \le \alpha \le 1$.

5.164 **a.** If $t_1 = t_2 = t_3 = t$, then $m(t, t, t) = E\left(e^{t(X_1+X_2+X_3)}\right)$. This, by definition, is the mgf for the random variable $X_1 + X_2 + X_3$.

b. Similarly with $t_1 = t_2 = t$ and $t_3 = 0$, $m(t, t, 0) = E\left(e^{t(X_1+X_2)}\right)$.

c. We prove the continuous case here (the discrete case is similar). Let (X_1, X_2, X_3) be continuous random variables with joint density function $f(x_1, x_2, x_3)$. Then,

$$m(t_1, t_2, t_3) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{t_1x_1} e^{t_2x_2} e^{t_3x_3} f(x_1, x_2, x_3)\,dx_1 dx_2 dx_3 .$$

Then,

$$\left.\frac{\partial^{k_1+k_2+k_3}}{\partial t_1^{k_1} \partial t_2^{k_2} \partial t_3^{k_3}} m(t_1, t_2, t_3)\right|_{t_1=t_2=t_3=0} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1^{k_1} x_2^{k_2} x_3^{k_3} f(x_1, x_2, x_3)\,dx_1 dx_2 dx_3 .$$

This is easily recognized as $E\left(X_1^{k_1} X_2^{k_2} X_3^{k_3}\right)$.

5.165 **a.** $m(t_1, t_2, t_3) = \sum_{x_1}\sum_{x_2}\sum_{x_3} \frac{n!}{x_1!x_2!x_3!} e^{t_1x_1+t_2x_2+t_3x_3} p_1^{x_1} p_2^{x_2} p_3^{x_3}$

$$= \sum_{x_1}\sum_{x_2}\sum_{x_3} \frac{n!}{x_1!x_2!x_3!} (p_1e^{t_1})^{x_1} (p_2e^{t_2})^{x_2} (p_3e^{t_3})^{x_3} = (p_1e^{t_1} + p_2e^{t_2} + p_3e^{t_3})^n.$$ The

final form follows from the multinomial theorem.

b. The mgf for X_1 can be found by evaluating $m(t, 0, 0)$. Note that $q = p_2 + p_3 = 1 - p_1$.

c. Since $\text{Cov}(X_1, X_2) = E(X_1X_2) - E(X_1)E(X_2)$ and $E(X_1) = np_1$ and $E(X_2) = np_2$ since X_1 and X_2 have marginal binomial distributions. To find $E(X_1X_2)$, note that

$$\frac{\partial^2}{\partial t_1 \partial t_2} m(t_1, t_2, 0)\Big|_{t_1=t_2=0} = n(n-1)p_1 p_2 .$$

Thus, $\text{Cov}(X_1, X_2) = n(n-1)p_1p_2 - (np_1)(np_2) = -np_1p_2$.

5.166 The joint probability mass function of (Y_1, Y_2, Y_3) is given by

$$p(y_1, y_2, y_3) = \frac{\binom{N_1}{y_1}\binom{N_2}{y_2}\binom{N_3}{y_3}}{\binom{N}{n}} = \frac{\binom{Np_1}{y_1}\binom{Np_2}{y_2}\binom{Np_3}{y_3}}{\binom{N}{n}},$$

where $y_1 + y_2 + y_3 = n$. The marginal distribution of Y_1 is hypergeometric with $r = Np_1$, so $E(Y_1) = np_1$, $V(Y_1) = np_1(1-p_1)\left(\frac{N-n}{N-1}\right)$. Similarly, $E(Y_2) = np_2$, $V(Y_2) = np_2(1-p_2)\left(\frac{N-n}{N-1}\right)$. It can be shown that (using mathematical expectation and straightforward albeit messy algebra) $E(Y_1Y_2) = n(n-1)p_1 p_2 \frac{N}{N-1}$. Using this, it is seen that

$$\text{Cov}(Y_1, Y_2) = n(n-1)p_1 p_2 \tfrac{N}{N-1} - (np_1)(np_2) = -np_1p2\left(\tfrac{N-n}{N-1}\right).$$

(Note the similar expressions in Ex. 5.165.) Finally, it can be found that

$$\rho = -\sqrt{\frac{p_1 p_2}{(1-p_1)(1-p_2)}} .$$

5.167 **a.** For this exercise, the quadratic form of interest is

$$At^2 + Bt + C = E(Y_1^2)t^2 + [-2E(Y_1Y_2)]t + [E(Y_2^2)]^2 .$$

Since $E[(tY_1 - Y_2)^2] \geq 0$ (it is the integral of a non–negative quantity), so we must have that $At^2 + Bt + C \geq 0$. In order to satisfy this inequality, the two roots of this quadratic must either be imaginary or equal. In terms of the discriminant, we have that

$$B^2 - 4AC \leq 0 \text{, or}$$

$$[-2E(Y_1Y_2)]^2 - 4E(Y_1^2)E(Y_2^2) \leq 0 .$$

Thus, $[E(Y_1Y_2)]^2 \leq E(Y_1^2)E(Y_2^2)$.

b. Let $\mu_1 = E(Y_1)$, $\mu_2 = E(Y_2)$, and define $Z_1 = Y_1 - \mu_1$, $Z_2 = Y_2 - \mu_2$. Then,

$$\rho^2 = \frac{[E(Y_1 - \mu_1)(Y_2 - \mu_2)]^2}{[E(Y_1 - \mu_1)^2]E[(Y_2 - \mu_2)^2]} = \frac{[E(Z_1Z_2)]^2}{E(Z_1^{\ 2})E(Z_2^{\ 2})} \leq 1$$

by the result in part **a.**

Chapter 6: Functions of Random Variables

6.1 The distribution function of Y is $F_Y(y) = \int_0^y 2(1-t)dt = 2y - y^2$, $0 \le y \le 1$.

a. $F_{U_1}(u) = P(U_1 \le u) = P(2Y - 1 \le u) = P(Y \le \frac{u+1}{2}) = F_Y(\frac{u+1}{2}) = 2(\frac{u+1}{2}) - (\frac{u+1}{2})^2$. Thus, $f_{U_1}(u) = F'_{U_1}(u) = \frac{1-u}{2}$, $-1 \le u \le 1$.

b. $F_{U_2}(u) = P(U_2 \le u) = P(1 - 2Y \le u) = P(Y \le \frac{1-u}{2}) = F_Y(\frac{1-u1}{2}) = 1 - 2(\frac{u+1}{2}) = (\frac{u+1}{2})^2$. Thus, $f_{U_2}(u) = F'_{U_2}(u) = \frac{u+1}{2}$, $-1 \le u \le 1$.

c. $F_{U_3}(u) = P(U_3 \le u) = P(Y^2 \le u) = P(Y \le \sqrt{u}) = F_Y(\sqrt{u}) = 2\sqrt{u} - u$ Thus, $f_{U_3}(u) = F'_{U_3}(u) = \frac{1}{\sqrt{u}} - 1$, $0 \le u \le 1$.

d. $E(U_1) = -1/3$, $E(U_2) = 1/3$, $E(U_3) = 1/6$.

e. $E(2Y - 1) = -1/3$, $E(1 - 2Y) = 1/3$, $E(Y^2) = 1/6$.

6.2 The distribution function of Y is $F_Y(y) = \int_{-1}^y (3/2)t^2 dt = (1/2)(y^3 - 1)$, $-1 \le y \le 1$.

a. $F_{U_1}(u) = P(U_1 \le u) = P(3Y \le u) = P(Y \le u/3) = F_Y(u/3) = \frac{1}{2}(u^3/18 - 1)$. Thus, $f_{U_1}(u) = F'_{U_1}(u) = u^2/18$, $-3 \le u \le 3$.

b. $F_{U_2}(u) = P(U_2 \le u) = P(3 - Y \le u) = P(Y \ge 3 - u) = 1 - F_Y(3 - u) = \frac{1}{2}[1 - (3-u)^3]$. Thus, $f_{U_2}(u) = F'_{U_2}(u) = \frac{3}{2}(3-u)^2$, $2 \le u \le 4$.

c. $F_{U_3}(u) = P(U_3 \le u) = P(Y^2 \le u) = P(-\sqrt{u} \le Y \le \sqrt{u}) = F_Y(\sqrt{u}) - F_Y(-\sqrt{u}) = u^{3/2}$. Thus, $f_{U_3}(u) = F'_{U_3}(u) = \frac{3}{2}\sqrt{u}$, $0 \le u \le 1$.

6.3 The distribution function for Y is $F_Y(y) = \begin{cases} y^2/2 & 0 \le y \le 1 \\ y - 1/2 & 1 < y \le 1.5 \\ 1 & y > 1.5 \end{cases}$.

a. $F_U(u) = P(U \le u) = P(10Y - 4 \le u) = P(Y \le \frac{u+4}{10}) = F_Y(\frac{u+4}{10})$. So,

$$F_U(u) = \begin{cases} \frac{(u+4)^2}{200} & -4 \le u \le 6 \\ \frac{u-1}{10} & 6 < u \le 11 \\ 1 & u > 11 \end{cases}, \text{ and } f_U(u) = F'_U(u) = \begin{cases} \frac{u+4}{100} & -4 \le u \le 6 \\ \frac{1}{10} & 6 < u \le 11 \\ 0 & \text{elsewhere} \end{cases}.$$

b. $E(U) = 5.583$.

c. $E(10Y - 4) = 10(23/24) - 4 = 5.583$.

6.4 The distribution function of Y is $F_Y(y) = 1 - e^{-y/4}$, $0 \le y$.

a. $F_U(u) = P(U \le u) = P(3Y + 1 \le u) = P(Y \le \frac{u-1}{3}) = F_Y(\frac{u-1}{3}) = 1 - e^{-(u-1)/12}$. Thus, $f_U(u) = F'_U(u) = \frac{1}{12}e^{-(u-1)/12}$, $u \ge 1$.

b. $E(U) = 13$.

6.5 The distribution function of Y is $F_Y(y) = y/4$, $1 \le y \le 5$.
$F_U(u) = P(U \le u) = P(2Y^2 + 3 \le u) = P(Y \le \sqrt{\tfrac{u-3}{2}}) = F_Y(\sqrt{\tfrac{u-3}{2}}) = \tfrac{1}{4}\sqrt{\tfrac{u-3}{2}}$. Differentiating,
$f_U(u) = F_U'(u) = \tfrac{1}{16}\left(\tfrac{u-3}{2}\right)^{-1/2}$, $5 \le u \le 53$.

6.6 Refer to Ex. 5.10 ad 5.78. Define $F_U(u) = P(U \le u) = P(Y_1 - Y_2 \le u) = P(Y_1 \le Y_2 + u)$.

a. For $u \le 0$, $F_U(u) = P(U \le u) = P(Y_1 - Y_2 \le u) = 0$.

For $0 \le u < 1$, $F_U(u) = P(U \le u) = P(Y_1 - Y_2 \le u) = \int_0^u \int_{2y_2}^{y_2+u} 1 dy_1 dy_2 = u^2/2$.

For $1 \le u \le 2$, $F_U(u) = P(U \le u) = P(Y_1 - Y_2 \le u) = 1 - \int_0^{2-u} \int_{y_2+u}^{2} 1 dy_1 dy_2 = 1 - (2-u)^2/2$.

Thus, $f_U(u) = F_U'(u) = \begin{cases} u & 0 \le u < 1 \\ 2-u & 1 \le y \le 2 \\ 0 & \text{elsewhere} \end{cases}$.

b. $E(U) = 1$.

6.7 Let $F_Z(z)$ and $f_Z(z)$ denote the standard normal distribution and density functions respectively.

a. $F_U(u) = P(U \le u) = P(Z^2 \le u) = P(-\sqrt{u} \le Z \le \sqrt{u}) = F_Z(\sqrt{u}) - F_Z(-\sqrt{u})$. The density function for U is then

$$f_U(u) = F_U'(u) = \tfrac{1}{2\sqrt{u}} f_Z(\sqrt{u}) + \tfrac{1}{2\sqrt{u}} f_Z(-\sqrt{u}) = \tfrac{1}{\sqrt{u}} f_Z(\sqrt{u}),\ u \ge 0.$$

Evaluating, we find $f_U(u) = \tfrac{1}{\sqrt{\pi}\sqrt{2}} u^{-1/2} e^{-u/2}$ $u \ge 0$.

b. U has a gamma distribution with $\alpha = 1/2$ and $\beta = 2$ (recall that $\Gamma(1/2) = \sqrt{\pi}$).

c. This is the chi–square distribution with one degree of freedom.

6.8 Let $F_Y(y)$ and $f_Y(y)$ denote the beta distribution and density functions respectively.

a. $F_U(u) = P(U \le u) = P(1 - Y \le u) = P(Y \ge 1 - u) = 1 - F_Y(1-u)$. The density function for U is then $f_U(u) = F_U'(u) = f_Y(1-u) = \tfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} u^{\beta-1}(1-u)^{\alpha-1}$, $0 \le u \le 1$.

b. $E(U) = 1 - E(Y) = \tfrac{\beta}{\alpha+\beta}$.

c. $V(U) = V(Y)$.

6.9 Note that this is the same density from Ex. 5.12: $f(y_1, y_2) = 2$, $0 \le y_1 \le 1$, $0 \le y_2 \le 1$, $0 \le y_1 + y_2 \le 1$.

a. $F_U(u) = P(U \le u) = P(Y_1 + Y_2 \le u) = P(Y_1 \le u - Y_2) = \int_0^u \int_0^{u-y_2} 2 dy_1 dy_2 = u^2$. Thus,

$f_U(u) = F_U'(u) = 2u$, $0 \le u \le 1$.

b. $E(U) = 2/3$.

c. (found in an earlier exercise in Chapter 5) $E(Y_1 + Y_2) = 2/3$.

6.10 Refer to Ex. 5.15 and Ex. 5.108.

a. $F_U(u) = P(U \le u) = P(Y_1 - Y_2 \le u) = P(Y_1 \le u + Y_2) = \int_0^{\infty}\int_{y_2}^{u+y_2} e^{-y_1}dy_1dy_2 = 1 - e^{-u}$, so that

$f_U(u) = F_U'(u) = e^{-u}$, $u \ge 0$, so that U has an exponential distribution with β = 1.

b. From part a above, $E(U) = 1$.

6.11 It is given that $f_i(y_i) = e^{-y_i}$, $y_i \ge 0$ for $i = 1, 2$. Let $U = (Y_1 + Y_2)/2$.

a. $F_U(u) = P(U \le u) = P(\frac{Y_1+Y_2}{2} \le u) = P(Y_1 \le 2u - Y_2) = \int_0^{2u}\int_{y_2}^{2u-y_2} e^{-y_1-y_2}dy_1dy_2 = 1 - e^{-2u} - 2ue^{-2u}$,

so that $f_U(u) = F_U'(u) = 4ue^{-2u}$, $u \ge 0$, a gamma density with α = 2 and β = 1/2.

b. From part (a), $E(U) = 1$, $V(U) = 1/2$.

6.12 Let $F_Y(y)$ and $f_Y(y)$ denote the gamma distribution and density functions respectively.

a. $F_U(u) = P(U \le u) = P(cY \le u) = P(Y \le u/c)$. The density function for U is then $f_U(u) = F_U'(u) = \frac{1}{c} f_Y(u/c) = \frac{1}{\Gamma(\alpha)(c\beta)^{\alpha}} u^{\alpha-1}e^{-u/c\beta}$, $u \ge 0$. Note that this is another gamma distribution.

b. The shape parameter is the same (α), but the scale parameter is cβ.

6.13 Refer to Ex. 5.8;

$$F_U(u) = P(U \le u) = P(Y_1 + Y_2 \le u) = P(Y_1 \le u - Y_2) = \int_0^{u}\int_0^{u-y_2} e^{-y_1-y_2}dy_1dy_2 = 1 - e^{-u} - ue^{-u}.$$

Thus, $f_U(u) = F_U'(u) = ue^{-u}$, $u \ge 0$.

6.14 Since Y_1 and Y_2 are independent, so $f(y_1, y_2) = 18(y_1 - y_1^2)y_2^2$, for $0 \le y_1 \le 1$, $0 \le y_2 \le 1$. Let $U = Y_1Y_2$. Then,

$$F_U(u) = P(U \le u) = P(Y_1Y_2 \le u) = P(Y_1 \le u/Y_2) = P(Y_1 > u/Y_2) = 1 - \int_u^1\int_{u/y_2}^1 18(y_1 - y_1^2)y_2^2dy_1dy_2$$

$= 9u^2 - 8u^3 + 6u^3\ln u$.

$f_U(u) = F_U'(u) = 18u(1 - u + u\ln u)$, $0 \le u \le 1$.

6.15 Let U have a uniform distribution on (0, 1). The distribution function for U is $F_U(u) = P(U \le u) = u$, $0 \le u \le 1$. For a function G, we require $G(U) = Y$ where Y has distribution function $F_Y(y) = 1 - e^{-y^2}$, $y \ge 0$. Note that

$$F_Y(y) = P(Y \le y) = P(G(U) \le y) = P[U \le G^{-1}(y)] = F_U[G^{-1}(y)] = u.$$

So it must be true that $G^{-1}(y) = 1 - e^{-y^2} = u$ so that $G(u) = [-\ln(1-u)]^{-1/2}$. Therefore, the random variable $Y = [-\ln(U-1)]^{-1/2}$ has distribution function $F_Y(y)$.

6.16 Similar to Ex. 6.15. The distribution function for *Y* is $F_Y(y) = b\int_b^y t^{-2}dt = 1 - \frac{b}{y},\ y \ge b.$

$F_Y(y) = P(Y \le y) = P(G(U) \le y) = P[U \le G^{-1}(y)] = F_U[G^{-1}(y)] = u.$
So it must be true that $G^{-1}(y) = 1 - \frac{b}{y} = u$ so that $G(u) = \frac{b}{1-u}$. Therefore, the random variable $Y = b/(1 - U)$ has distribution function $F_Y(y)$.

6.17 **a.** Taking the derivative of $F(y)$, $f(y) = \frac{\alpha y^{\alpha-1}}{\theta^\alpha}$, $0 \le y \le \theta$.

b. Following Ex. 6.15 and 6.16, let $u = \left(\frac{y}{\theta}\right)^\alpha$ so that $y = \theta u^{1/\alpha}$. Thus, the random variable $Y = \theta U^{1/\alpha}$ has distribution function $F_Y(y)$.
c. From part (b), the transformation is $y = 4\sqrt{u}$. The values are 2.0785, 3.229, 1.5036, 1.5610, 2.403.

6.18 **a.** Taking the derivative of the distribution function yields $f(y) = \alpha\beta^\alpha y^{-\alpha-1}$, $y \ge \beta$.
b. Following Ex. 6.15, let $u = 1 - \left(\frac{\beta}{y}\right)^\alpha$ so that $y = \frac{\beta}{(1-u)^{1/\alpha}}$. Thus, $Y = \beta(1-U)^{-1/\alpha}$.
c. From part (b), $y = 3/\sqrt{1-u}$. The values are 3.0087, 3.3642, 6.2446, 3.4583, 4.7904.

6.19 The distribution function for *X* is:
$F_X(x) = P(X \le x) = P(1/Y \le x) = P(Y \ge 1/x) = 1 - F_Y(1/x)$
$= 1 - \left[1 - (\beta x)^\alpha\right] = (\beta x)^\alpha$, $0 < x < \beta^{-1}$, which is a power distribution with $\theta = \beta^{-1}$.

6.20 **a.** $F_W(w) = P(W \le w) + P(Y^2 \le w) = P(Y \le \sqrt{w}) = F_Y(\sqrt{w}) = \sqrt{w},\ 0 \le w \le 1.$
b. $F_W(w) = P(W \le w) + P(\sqrt{Y} \le w) = P(Y \le w^2) = F_Y(w^2) = w^2, 0 \le w \le 1.$

6.21 By definition, $P(X = i) = P[F(i - 1) < U \le F(i)] = F(i) - F(i - 1)$, for $i = 1, 2, \ldots$, since for any $0 \le a \le 1$, $P(U \le a) = a$ for any $0 \le a \le 1$. From Ex. 4.5, $P(Y = i) = F(i) - F(i - 1)$, for $i = 1, 2, \ldots$. Thus, *X* and *Y* have the same distribution.

6.22 Let *U* have a uniform distribution on the interval (0, 1). For a geometric distribution with parameter *p* and distribution function *F*, define the random variable *X* as:
$X = k$ if and only if $F(k - 1) < U \le F(k)$, $k = 1, 2, \ldots$.
Or since $F(k) = 1 - q^k$, we have that:
$X = k$ if and only if $1 - q^{k-1} < U \le 1 - q^k$, OR
$X = k$ if and only if $q^k, < 1-U \le q^{k-1}$, OR
$X = k$ if and only if $k\ln q \le \ln(1-U) \le (k-1)\ln q$, OR
$X = k$ if and only if $k-1 < [\ln(1-U)]/\ln q \le k$.

6.23 **a.** If $U = 2Y - 1$, then $Y = \frac{U+1}{2}$. Thus, $\frac{dy}{du} = \frac{1}{2}$ and $f_U(u) = \frac{1}{2}2(1 - \frac{u+1}{2}) = \frac{1-u}{2}, -1 \le u \le 1.$
b. If $U = 1 - 2Y$, then $Y = \frac{1-U}{2}$. Thus, $\frac{dy}{du} = \frac{1}{2}$ and $f_U(u) = \frac{1}{2}2(1 - \frac{1-u}{2}) = \frac{1+u}{2}, -1 \le u \le 1.$
c. If $U = Y^2$, then $Y = \sqrt{U}$. Thus, $\frac{dy}{du} = \frac{1}{2\sqrt{u}}$ and $f_U(u) = \frac{1}{2\sqrt{u}}2(1 - \sqrt{u}) = \frac{1-\sqrt{u}}{\sqrt{u}}, 0 \le u \le 1.$

6.24 If $U = 3Y + 1$, then $Y = \frac{U-1}{3}$. Thus, $\frac{dy}{du} = \frac{1}{3}$. With $f_Y(y) = \frac{1}{4}e^{-y/4}$, we have that $f_U(u) = \frac{1}{3}\left[\frac{1}{4}e^{-(u-1)/12}\right] = \frac{1}{12}e^{-(u-1)/12}$, $1 \le u$.

6.25 Refer to Ex. 6.11. The variable of interest is $U = \frac{Y_1+Y_2}{2}$. Fix $Y_2 = y_2$. Then, $Y_1 = 2u - y_2$ and $\frac{dy_1}{du} = 2$. The joint density of U and Y_2 is $g(u, y_2) = 2e^{-2u}$, $u \ge 0$, $y_2 \ge 0$, and $y_2 < 2u$.

Thus, $f_U(u) = \int_0^{2u} 2e^{-2u}dy_2 = 4ue^{-2u}$ for $u \ge 0$.

6.26 **a.** Using the transformation approach, $Y = U^{1/m}$ so that $\frac{dy}{du} = \frac{1}{m}u^{-(m-1)/m}$ so that the density function for U is $f_U(u) = \frac{1}{\alpha}e^{-u/\alpha}$, $u \ge 0$. Note that this is the exponential distribution with mean α.

b. $E(Y^k) = E(U^{k/m}) = \int_0^\infty u^{k/m}\frac{1}{\alpha}e^{-u/\alpha}du = \Gamma\left(\frac{k}{m}+1\right)\alpha^{k/m}$, using the result from Ex. 4.111.

6.27 **a.** Let $W = \sqrt{Y}$. The random variable Y is exponential so $f_Y(y) = \frac{1}{\beta}e^{-y/\beta}$. Then, $Y = W^2$ and $\frac{dy}{dw} = 2w$. Then, $f_Y(y) = \frac{2}{\beta}we^{-w^2/\beta}$, $w \ge 0$, which is Weibull with $m = 2$.

b. It follows from Ex. 6.26 that $E(Y^{k/2}) = \Gamma\left(\frac{k}{2}+1\right)\beta^{k/2}$

6.28 If Y is uniform on the interval (0, 1), $f_U(u) = 1$. Then, $Y = e^{-U/2}$ and $\frac{dy}{du} = -\frac{1}{2}e^{-u/2}$. Then, $f_Y(y) = 1\,|-\frac{1}{2}e^{-u/2}| = \frac{1}{2}e^{-u/2}$, $u \ge 0$ which is exponential with mean 2.

6.29 **a.** With $W = \frac{mV^2}{2}$, $V = \sqrt{\frac{2W}{m}}$ and $|\frac{dv}{dw}| = \frac{1}{\sqrt{2mw}}$. Then,

$$f_W(w) = \frac{a(2w/m)}{\sqrt{2mw}}e^{-2bw/m} = \frac{a\sqrt{2}}{m^{3/2}}w^{1/2}e^{-w/kT},\ w \ge 0.$$

The above expression is in the form of a gamma density, so the constant a must be chosen so that the density integrate to 1, or simply

$$\frac{a\sqrt{2}}{m^{3/2}} = \frac{1}{\Gamma(\frac{3}{2})(kT)^{3/2}}.$$

So, the density function for W is

$$f_W(w) = \frac{1}{\Gamma(\frac{3}{2})(kT)^{3/2}}w^{1/2}e^{-w/kT}.$$

b. For a gamma random variable, $E(W) = \frac{3}{2}kT$.

6.30 The density function for I is $f_I(i) = 1/2$, $9 \le i \le 11$. For $P = 2I^2$, $I = \sqrt{P/2}$ and $\frac{di}{dp} = (1/2)^{3/2}p^{-1/2}$. Then, $f_p(p) = \frac{1}{4\sqrt{2p}}$, $162 \le p \le 242$.

6.31 Similar to Ex. 6.25. Fix $Y_1 = y_1$. Then, $U = Y_2/y_1$, $Y_2 = y_1U$ and $|\frac{dy_2}{du}| = y_1$. The joint density of Y_1 and U is $f(y_1, u) = \frac{1}{8}y_1^2 e^{-y_1(1+u)/2}$, $y_1 \ge 0$, $u \ge 0$. So, the marginal density for U is $f_U(u) = \int_0^\infty \frac{1}{8}y_1^2 e^{-y_1(1+u)/2} dy_1 = \frac{2}{(1+u)^3}$, $u \ge 0$.

6.32 Now $f_Y(y) = 1/4$, $1 \le y \le 5$. If $U = 2Y^2 + 3$, then $Y = \left(\frac{U-3}{2}\right)^{1/2}$ and $|\frac{dy}{du}| = \frac{1}{4}\left(\frac{\sqrt{2}}{\sqrt{u-3}}\right)$. Thus, $f_U(u) = \frac{1}{8\sqrt{2(u-3)}}$, $5 \le u \le 53$.

6.33 If $U = 5 - (Y/2)$, $Y = 2(5 - U)$. Thus, $|\frac{dy}{du}| = 2$ and $f_U(u) = 4(80 - 31u + 3u^2)$, $4.5 \le u \le 5$.

6.34 **a.** If $U = Y^2$, $Y = \sqrt{U}$. Thus, $|\frac{dy}{du}| = \frac{1}{2\sqrt{u}}$ and $f_U(u) = \frac{1}{\theta}e^{-u/\theta}$, $u \ge 0$. This is the exponential density with mean θ.

b. From part a, $E(Y) = E(U^{1/2}) = \frac{\sqrt{\pi\theta}}{2}$. Also, $E(Y^2) = E(U) = \theta$, so $V(Y) = \theta[1 - \frac{\pi}{4}]$.

6.35 By independence, $f(y_1, y_2) = 1$, $0 \le y_1 \le 0$, $0 \le y_2 \le 1$. Let $U = Y_1Y_2$. For a fixed value of Y_1 at y_1, then $y_2 = u/y_1$. So that $\frac{dy_2}{du} = \frac{1}{y_1}$. So, the joint density of Y_1 and U is

$$g(y_1, u) = 1/y_1,\ 0 \le y_1 \le 0,\ 0 \le u \le y_1.$$

Thus, $f_U(u) = \int_u^1 (1/y_1)dy_1 = -\ln(u)$, $0 \le u \le 1$.

6.36 By independence, $f(y_1, y_2) = \frac{4y_1y_2}{\theta^2}e^{-(y_1^2+y_2^2)}$, $y_1 > 0$, $y_2 > 0$. Let $U = Y_1^2 + Y_2^2$. For a fixed value of Y_1 at y_1, then $U = y_1^2 + Y_2^2$ so we can write $y_2 = \sqrt{u - y_1^2}$. Then, $\frac{dy_2}{du} = \frac{1}{2\sqrt{u-y_1^2}}$ so that the joint density of Y_1 and U is

$$g(y_1, u) = \frac{4y_1\sqrt{u-y_1^2}}{\theta^2}e^{-u/\theta}\frac{1}{2\sqrt{u-y_1^2}} = \frac{2}{\theta^2}y_1e^{-u/\theta}, \text{ for } 0 < y_1 < \sqrt{u}.$$

Then, $f_U(u) = \int_0^{\sqrt{u}} \frac{2}{\theta^2}y_1e^{-u/\theta}dy_1 = \frac{1}{\theta^2}ue^{-u/\theta}$. Thus, U has a gamma distribution with α = 2.

6.37 The mass function for the Bernoulli distribution is $p(y) = p^y(1-p)^{1-y}$, $y = 0, 1$.

a. $m_{Y_1}(t) = E(e^{tY_1}) = \sum_{x=0}^{1} e^{ty}p(y) = 1 - p + pe^t$.

b. $m_W(t) = E(e^{tW}) = \prod_{i=1}^{n} m_{Y_i}(t) = [1 - p + pe^t]^n$

c. Since the mgf for W is in the form of a binomial mgf with n trials and success probability p, this is the distribution for W.

6.38 Let Y_1 and Y_2 have mgfs as given, and let $U = a_1Y_1 + a_2Y_2$. The mdf for U is

$$m_U(t) = E(e^{Ut}) = E(e^{(a_1Y_1+a_2Y_2)t}) = E(e^{(a_1t)Y_1})E(e^{(a_2t)Y_2}) = m_{Y_1}(a_1t)m_{Y_2}(a_2t).$$

6.39 The mgf for the exponential distribution with β = 1 is $m(t) = (1-t)^{-1}$, $t < 1$. Thus, with Y_1 and Y_2 each having this distribution and $U = (Y_1 + Y_2)/2$. Using the result from Ex. 6.38, let $a_1 = a_2 = 1/2$ so the mgf for U is $m_U(t) = m(t/2)m(t/2) = (1-t/2)^{-2}$. Note that this is the mgf for a gamma random variable with α = 2, β = 1/2, so the density function for U is $f_U(u) = 4ue^{-2u}$, $u \geq 0$.

6.40 It has been shown that the distribution of both Y_1^2 and Y_2^2 is chi–square with ν = 1. Thus, both have mgf $m(t) = (1-2t)^{-1/2}$, $t < 1/2$. With $U = Y_1^2 + Y_2^2$, use the result from Ex. 6.38 with $a_1 = a_2 = 1$ so that $m_U(t) = m(t)m(t) = (1-2t)^{-1}$. Note that this is the mgf for a exponential random variable with β = 2, so the density function for U is $f_U(u) = \frac{1}{2}e^{-u/2}$, $u \geq 0$ (this is also the chi–square distribution with ν = 2.)

6.41 (Special case of Theorem 6.3) The mgf for the normal distribution with parameters μ and σ is $m(t) = e^{\mu t + \sigma^2 t^2/2}$. Since the Y_i's are independent, the mgf for U is given by

$$m_U(t) = E(e^{Ut}) = \prod_{i=1}^{n} E(e^{a_i t Y_i}) = \prod_{i=1}^{n} m(a_i t) = \exp\left[\mu t \sum_i a_i + (t^2\sigma^2/2)\sum_i a_i^2\right].$$

This is the mgf for a normal variable with mean $\mu\sum_i a_i$ and variance $\sigma^2\sum_i a_i^2$.

6.42 The probability of interest is $P(Y_2 > Y_1) = P(Y_2 - Y_1 > 0)$. By Theorem 6.3, the distribution of $Y_2 - Y_1$ is normal with μ = 4000 – 5000 = –1000 and $\sigma^2 = 400^2 + 300^2 = 250{,}000$. Thus, $P(Y_2 - Y_1 > 0) = P(Z > \frac{0-(-1000)}{\sqrt{250{,}000}}) = P(Z > 2) = .0228$.

6.43 **a.** From Ex. 6.41, $\bar{Y}$ has a normal distribution with mean μ and variance σ^2/n.

b. For the given values, $\bar{Y}$ has a normal distribution with variance $\sigma^2/n = 16/25$. Thus, the standard deviation is 4/5 so that

$$P(|\bar{Y} - \mu| \leq 1) = P(-1 \leq \bar{Y} - \mu \leq 1) = P(-1.25 \leq Z \leq 1.25) = .7888.$$

c. Similar to the above, the probabilities are .8664, .9544, .9756. So, as the sample size increases, so does the probability that $P(|\bar{Y} - \mu| \leq 1)$.

6.44 The total weight of the watermelons in the packing container is given by $U = \sum_{i=1}^{n} Y_i$, so by Theorem 6.3 U has a normal distribution with mean $15n$ and variance $4n$. We require that $.05 = P(U > 140) = P(Z > \frac{140-15n}{\sqrt{4n}})$. Thus, $\frac{140-15n}{\sqrt{4n}} = z_{.05} = 1.645$. Solving this nonlinear expression for n, we see that $n \approx 8.687$. Therefore, the maximum number of watermelons that should be put in the container is 8 (note that with this value n, we have $P(U > 140) = .0002$).

6.45 By Theorem 6.3 we have that $U = 100 + 7Y_1 + 3Y_2$ is a normal random variable with mean $\mu = 100 + 7(10) + 3(4) = 182$ and variance $\sigma^2 = 49(.5)^2 + 9(.2)^2 = 12.61$. We require a value c such that $P(U > c) = P(Z > \frac{c-182}{\sqrt{12.61}})$. So, $\frac{c-182}{\sqrt{12.61}} = 2.33$ and $c = \$190.27$.

6.46 The mgf for W is $m_W(t) = E(e^{Wt}) = E(e^{(2Y/\beta)t}) = m_Y(2t/\beta) = (1-2t)^{-n/2}$. This is the mgf for a chi–square variable with n degrees of freedom.

6.47 By Ex. 6.46, $U = 2Y/4.2$ has a chi–square distribution with $\nu = 7$. So, by Table III, $P(Y > 33.627) = P(U > 2(33.627)/4.2) = P(U > 16.0128) = .025$.

6.48 From Ex. 6.40, we know that $V = Y_1^2 + Y_2^2$ has a chi–square distribution with $\nu = 2$. The density function for V is $f_V(v) = \frac{1}{2}e^{-v/2}$, $v \geq 0$. The distribution function of $U = \sqrt{V}$ is $F_U(u) = P(U \leq u) = P(V \leq u^2) = F_V(u^2)$, so that $f_U(u) = F_U'(u) = ue^{-u^2/2}$, $u \geq 0$. A sharp observer would note that this is a Weibull density with shape parameter 2 and scale 2.

6.49 The mgfs for Y_1 and Y_2 are, respectively, $m_{Y_1}(t) = [1-p+pe^t]^{n_1}$, $m_{Y_2}(t) = [1-p+pe^t]^{n_2}$. Since Y_1 and Y_2 are independent, the mgf for $Y_1 + Y_2$ is $m_{Y_1}(t) \times m_{Y_2}(t) = [1-p+pe^t]^{n_1+n_2}$. This is the mgf of a binomial with $n_1 + n_2$ trials and success probability p.

6.50 The mgf for Y is $m_Y(t) = [1-p+pe^t]^n$. Now, define $X = n - Y$. The mgf for X is

$$m_X(t) = E(e^{tX}) = E(e^{t(n-Y)}) = e^{tn}m_Y(-t) = [p + (1-p)e^t]^n.$$

This is an mgf for a binomial with n trials and "success" probability $(1 - p)$. Note that the random variable X = # of *failures* observed in the experiment.

6.51 From Ex. 6.50, the distribution of $n_2 - Y_2$ is binomial with n_2 trials and "success" probability $1 - .8 = .2$. Thus, by Ex. 6.49, the distribution of $Y_1 + (n_2 - Y_2)$ is binomial with $n_1 + n_2$ trials and success probability $p = .2$.

6.52 The mgfs for Y_1 and Y_2 are, respectively, $m_{Y_1}(t) = e^{\lambda_1(e^t-1)}$, $m_{Y_2}(t) = e^{\lambda_2(e^t-1)}$.

a. Since Y_1 and Y_2 are independent, the mgf for $Y_1 + Y_2$ is $m_{Y_1}(t) \times m_{Y_2}(t) = e^{(\lambda_1+\lambda_2)(e^t-1)}$. This is the mgf of a Poisson with mean $\lambda_1 + \lambda_2$.

b. From Ex. 5.39, the distribution is binomial with m trials and $p = \frac{\lambda_1}{\lambda_1+\lambda_2}$.

6.53 The mgf for a binomial variable Y_i with n_i trials and success probability p_i is given by $m_{Y_i}(t) = [1-p_i+p_ie^t]^{n_i}$. Thus, the mgf for $U = \sum_{i=1}^n Y_i$ is $m_U(t) = \prod_i [1-p_i+p_ie^t]^{n_i}$.

a. Let $p_i = p$ and $n_i = m$ for all i. Here, U is binomial with $m(n)$ trials and success probability p.

b. Let $p_i = p$. Here, U is binomial with $\sum_{i=1}^n n_i$ trials and success probability p.

c. (Similar to Ex. 5.40) The cond. distribution is hypergeometric w/ $r = n_i$, $N = \sum n_i$.

d. By definition,

$$P(Y_1 + Y_2 = k \mid \textstyle\sum_{i=1}^{n} Y_i) = \frac{P(Y_1+Y_2=k, \sum Y_i = m)}{P(\sum Y_i = m)} = \frac{P(Y_1+Y_2=k, \sum_{i=3}^{n} Y_i = m-k)}{P(\sum Y_i = m)} = \frac{P(Y_1+Y_2=k)P(\sum_{i=3}^{n} Y_i = m-k)}{P(\sum Y_i = m)}$$

$$= \frac{\binom{n_1+n_2}{k}\binom{\sum_{i=3}^{n} n_i}{m-k}}{\binom{\sum_{i=1}^{n} n_i}{m}}$$, which is hypergeometric with $r = n_1 + n_2$.

e. No, the mgf for U does not simplify into a recognizable form.

6.54 **a.** The mgf for $U = \sum_{i=1}^{n} Y_i$ is $m_U(t) = e^{(e^t-1)\sum_i \lambda_i}$, which is recognized as the mgf for a Poisson w/ mean $\sum_i \lambda_i$.

b. This is similar to 6.52. The distribution is binomial with m trials and $p = \frac{\lambda_1}{\sum \lambda_i}$.

c. Following the same steps as in part d of Ex. 6.53, it is easily shown that the conditional distribution is binomial with m trials and success probability $\frac{\lambda_1+\lambda_2}{\sum \lambda_i}$.

6.55 Let $Y = Y_1 + Y_2$. Then, by Ex. 6.52, Y is Poisson with mean 7 + 7 = 14. Thus, $P(Y \geq 20) = 1 - P(Y \leq 19) = .077$.

6.56 Let U = total service time for two cars. Similar to Ex. 6.13, U has a gamma distribution with $\alpha = 2$, $\beta = 1/2$. Then, $P(U > 1.5) = \int_{1.5}^{\infty} 4ue^{-2u}du = .1991$.

6.57 For each Y_i, the mgf is $m_{Y_i}(t) = (1-\beta t)^{-\alpha_i}$, $t < 1/\beta$. Since the Y_i are independent, the mgf for $U = \sum_{i=1}^{n} Y_i$ is $m_U(t) = \prod (1-\beta t)^{-\alpha_i} = (1-\beta t)^{-\sum_{i=1}^{n} \alpha_i}$.

This is the mgf for the gamma with shape parameter $\sum_{i=1}^{n} \alpha_i$ and scale parameter β.

6.58 **a.** The mgf for each W_i is $m(t) = \frac{pe^t}{(1-qe^t)}$. The mgf for Y is $[m(t)]^r = \left(\frac{pe^t}{1-qe^t}\right)^r$, which is the mgf for the negative binomial distribution.

b. Differentiating with respect to t, we have

$$m'(t)\big|_{t=0} = r\left(\frac{pe^t}{1-qe^t}\right)^{r-1} \times \frac{pe^t}{(1-qe^t)^2}\bigg|_{t=0} = \frac{r}{p} = E(Y).$$

Taking another derivative with respect to t yields

$$m''(t)\big|_{t=0} = \frac{(1-qe^t)^{r+1} r^2 pe^t (pe^t)^{r-1} - r(pe^t)^r (r+1)(-qe^t)(1-qe^t)^r}{(1-qe^t)^{2(r+1)}}\bigg|_{t=0} = \frac{pr^2 + r(r+1)q}{p^2} = E(Y^2).$$

Thus, $V(Y) = E(Y^2) - [E(Y)]^2 = rq/p^2$.

c. This is similar to Ex. 6.53. By definition,

$$P(W_1 = k \mid \Sigma W_i) = \frac{P(W_1=k, \Sigma W_i = m)}{P(\Sigma W_i = m)} = \frac{P(W_1=k, \sum_{i=2}^{n} W_i = m-k)}{P(\Sigma W_i = m)} = \frac{P(W_1=k)P(\sum_{i=2}^{n} W_i = m-k)}{P(\Sigma W_i = m)} = \frac{\binom{m-k-1}{r-2}}{\binom{m-1}{r-1}}.$$

6.59 The mgfs for Y_1 and Y_2 are, respectively, $m_{Y_1}(t) = (1-2t)^{-\nu_1/2}$, $m_{Y_2}(t) = (1-2t)^{-\nu_2/2}$. Thus the mgf for $U = Y_1 + Y_2 = m_U(t) = m_{Y_1}(t) \times m_{Y_2}(t) = (1-2t)^{-(\nu_1+\nu_2)/2}$, which is the mgf for a chi–square variable with $\nu_1 + \nu_2$ degrees of freedom.

6.60 Note that since Y_1 and Y_2 are independent, $m_W(t) = m_{Y_1}(t) \times m_{Y_2}(t)$. Therefore, it must be so that $m_W(t)/m_{Y_1}(t) = m_{Y_2}(t)$. Given the mgfs for W and Y_1, we can solve for $m_{Y_2}(t)$:

$$m_{Y_2}(t) = \frac{(1-2t)^{-\nu}}{(1-2t)^{-\nu_1}} = (1-2t)^{-(\nu-\nu_1)/2}.$$

This is the mgf for a chi–squared variable with $\nu - \nu_1$ degrees of freedom.

6.61 Similar to Ex. 6.60. Since Y_1 and Y_2 are independent, $m_W(t) = m_{Y_1}(t) \times m_{Y_2}(t)$. Therefore, it must be so that $m_W(t)/m_{Y_1}(t) = m_{Y_2}(t)$. Given the mgfs for W and Y_1,

$$m_{Y_2}(t) = \frac{e^{\lambda(e^t-1)}}{e^{\lambda_1(e^t-1)}} = e^{(\lambda-\lambda_1)(e^t-1)}.$$

This is the mgf for a Poisson variable with mean $\lambda - \lambda_1$.

6.62
$$\begin{aligned} E\{\exp[t_1(Y_1+Y_2)+t_2(Y_1-Y_2)]\} &= E\{\exp[(t_1+t_2)Y_1+(t_1+t_2)Y_2]\} = m_{Y_1}(t_1+t_2)m_{Y_2}(t_1+t_2) \\ &= \exp[\tfrac{\sigma^2}{2}(t_1+t_2)^2]\exp[\tfrac{\sigma^2}{2}(t_1-t_2)^2] = \exp[\tfrac{\sigma^2}{2}t_1^{\,2}]\exp[\tfrac{\sigma^2}{2}t_2]^2 \\ &= m_{U_1}(t_1)m_{U_1}(t_2). \end{aligned}$$

Since the joint mgf factors, U_1 and U_2 are independent.

6.63 **a.** The marginal distribution for U_1 is $f_{U_1}(u_1) = \int_0^\infty \frac{1}{\beta^2} u_2 e^{-u_2/\beta} du_2 = 1$, $0 < u_1 < 1$.

b. The marginal distribution for U_2 is $f_{U_2}(u_2) = \int_0^1 \frac{1}{\beta^2} u_2 e^{-u_2/\beta} du_1 = \frac{1}{\beta^2} u_2 e^{-u_2/\beta}$, $u_2 > 0$. This is a gamma density with $\alpha = 2$ and scale parameter β.

c. Since the joint distribution factors into the product of the two marginal densities, they are independent.

6.64 **a.** By independence, the joint distribution of Y_1 and Y_2 is the product of the two marginal densities:

$$f(y_1, y_2) = \frac{1}{\Gamma(\alpha_1)\Gamma(\alpha_a)\beta^{\alpha_1+\alpha_2}} y_1^{\alpha_1-1} y_2^{\alpha_2-1} e^{-(y_1+y_2)/\beta}, \ y_1 \ge 0, y_2 \ge 0.$$

With U and V as defined, we have that $y_1 = u_1 u_2$ and $y_2 = u_2(1-u_1)$. Thus, the Jacobian of transformation $J = u_2$ (see Example 6.14). Thus, the joint density of U_1 and U_2 is

$$f(u_1,u_2) = \frac{1}{\Gamma(\alpha_1)\Gamma(\alpha_a)\beta^{\alpha_1+\alpha_2}}(u_1u_2)^{\alpha_1-1}[u_2(1-u_1)]^{\alpha_2-1}e^{-u_2/\beta}u_2$$
$$= \frac{1}{\Gamma(\alpha_1)\Gamma(\alpha_a)\beta^{\alpha_1+\alpha_2}}u_1^{\alpha_1-1}(1-u_1)^{\alpha_2-1}u_2^{\alpha_1+\alpha_2-1}e^{-u_2/\beta}, \text{ with } 0 < u_1 < 1, \text{ and } u_2 > 0.$$

b. $f_{U_1}(u_1) = \frac{1}{\Gamma(\alpha_1)\Gamma(\alpha_a)}u_1^{\alpha_1-1}(1-u_1)^{\alpha_2-1}\int_0^\infty \frac{1}{\beta^{\alpha_1+\alpha_2}}v^{\alpha_1+\alpha_2-1}e^{-v/\beta}dv = \frac{\Gamma(\alpha_1+\alpha_a)}{\Gamma(\alpha_1)\Gamma(\alpha_a)}u_1^{\alpha_1-1}(1-u_1)^{\alpha_2-1}$, with $0 < u_1 < 1$. This is the beta density as defined.

c. $f_{U_2}(u_2) = \frac{1}{\beta^{\alpha_1+\alpha_2}}u_2^{\alpha_1+\alpha_2-1}e^{-u_2/\beta}\int_0^1 \frac{1}{\Gamma(\alpha_1)\Gamma(\alpha_a)}u_1^{\alpha_1-1}(1-u_1)^{\alpha_2-1}du_1 = \frac{1}{\beta^{\alpha_1+\alpha_2}\Gamma(\alpha_1+\alpha_2)}u_2^{\alpha_1+\alpha_2-1}e^{-u_2/\beta}$, with $u_2 > 0$. This is the gamma density as defined.

d. Since the joint distribution factors into the product of the two marginal densities, they are independent.

6.65 **a.** By independence, the joint distribution of Z_1 and Z_2 is the product of the two marginal densities:

$$f(z_1,z_2) = \frac{1}{2\pi}e^{-(z_1^2+z_2^2)/2}.$$

With $U_1 = Z_1$ and $U_2 = Z_1 + Z_2$, we have that $z_1 = u_1$ and $z_2 = u_2 - u_1$. Thus, the Jacobian of transformation is

$$J = \begin{vmatrix} 1 & 0 \\ -1 & 1 \end{vmatrix} = 1.$$

Thus, the joint density of U_1 and U_2 is

$$f(u_1,u_2) = \frac{1}{2\pi}e^{-[u_1^2+(u_2-u_1)^2]/2} = \frac{1}{2\pi}e^{-(2u_1^2-2u_1u_2+u_2^2)/2}.$$

b. $E(U_1) = E(Z_1) = 0$, $E(U_2) = E(Z_1+Z_2) = 0$, $V(U_1) = V(Z_1) = 1$,
$V(U_2) = V(Z_1+Z_2) = V(Z_1)+V(Z_2) = 2$, $Cov(U_1,U_2) = E(Z_1^2) = 1$

c. Not independent since $\rho \neq 0$.

d. This is the bivariate normal distribution with $\mu_1 = \mu_2 = 0$, $\sigma_1^2 = 1$, $\sigma_2^2 = 2$, and $\rho = \frac{1}{\sqrt{2}}$.

6.66 **a.** Similar to Ex. 6.65, we have that $y_1 = u_1 - u_2$ and $y_2 = u_2$. So, the Jacobian of transformation is

$$J = \begin{vmatrix} 1 & -1 \\ 0 & 1 \end{vmatrix} = 1.$$

Thus, by definition the joint density is as given.

b. By definition of a marginal density, the marginal density for U_1 is as given.

c. If Y_1 and Y_2 are independent, their joint density factors into the product of the marginal densities, so we have the given form.

6.67 **a.** We have that $y_1 = u_1u_2$ and $y_2 = u_2$. So, the Jacobian of transformation is

$$J = \begin{vmatrix} u_2 & u_1 \\ 0 & 1 \end{vmatrix} = |u_2|.$$

Thus, by definition the joint density is as given.

b. By definition of a marginal density, the marginal density for U_1 is as given.

c. If Y_1 and Y_2 are independent, their joint density factors into the product of the marginal densities, so we have the given form.

6.68 **a.** Using the result from Ex. 6.67,

$$f(u_1,u_2) = 8(u_1u_2)u_2u_2 = 8u_1u_2^3,\ 0 \le u_1 \le 1,\ 0 \le u_2 \le 1.$$

b. The marginal density for U_1 is

$$f_{U_1}(u_1) = \int_0^1 8u_1u_2^3 du_2 = 2u_1,\ 0 \le u_1 \le 1.$$

The marginal density for U_1 is

$$f_{U_2}(u_2) = \int_0^1 8u_1u_2^3 du_1 = 4u_2^3,\ 0 \le u_2 \le 1.$$

The joint density factors into the product of the marginal densities, thus independence.

6.69 **a.** The joint density is $f(y_1,y_2) = \frac{1}{y_1^2y_2^2}$, $y_1 > 1$, $y_2 > 1$.

b. We have that $y_1 = u_1u_2$ and $y_2 = u_2(1 - u_1)$. The Jacobian of transformation is u_2. So,

$$f(u_1,u_2) = \frac{1}{u_1^2u_2^3(1-u_1)^2},$$

with limits as specified in the problem.

c. The limits may be simplified to: $1/u_1 < u_2$, $0 < u_1 < 1/2$, or $1/(1-u_1) < u_2$, $1/2 \le u_1 \le 1$.

d. If $0 < u_1 < 1/2$, then $f_{U_1}(u_1) = \int_{1/u_1}^{\infty} \frac{1}{u_1^2u_2^3(1-u_1)^2} du_2 = \frac{1}{2(1-u_1)^2}$.

If $1/2 \le u_1 \le 1$, then $f_{U_1}(u_1) = \int_{1/(1-u_1)}^{\infty} \frac{1}{u_1^2u_2^3(1-u_1)^2} du_2 = \frac{1}{2u_1^2}$.

e. Not independent since the joint density does not factor. Also note that the support is not rectangular.

6.70 **a.** Since Y_1 and Y_2 are independent, their joint density is $f(y_1, y_2) = 1$. The inverse transformations are $y_1 = \frac{u_1+u_2}{2}$ and $y_2 = \frac{u_1-u_2}{2}$. Thus the Jacobian is

$$J = \begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{vmatrix} = \frac{1}{2} \text{, so that}$$

$$f(u_1, u_2) = \tfrac{1}{2}, \text{ with limits as specified in the problem.}$$

b. The support is in the shape of a square with corners located (0, 0), (1, 1), (2, 0), (1, –1).

c. If $0 < u_1 < 1$, then $f_{U_1}(u_1) = \int_{-u_1}^{u_1} \frac{1}{2} du_2 = u_1$.

If $1 \le u_1 < 2$, then $f_{U_1}(u_1) = \int_{u_1-2}^{2-u_1} \frac{1}{2} du_2 = 2 - u_1$.

d. If $-1 < u_2 < 0$, then $f_{U_2}(u_2) = \int_{-u_2}^{2+u_2} \frac{1}{2} du_2 = 1 + u_2$.

If $0 \le u_2 < 1$, then $f_{U_2}(u_2) = \int_{u_2}^{2-u_2} \frac{1}{2} du_2 = 1 - u_2$.

6.71 **a.** The joint density of Y_1 and Y_2 is $f(y_1, y_2) = \frac{1}{\beta^2} e^{-(y_1+y_2)/\beta}$. The inverse transformations are $y_1 = \frac{u_1 u_2}{1+u_2}$ and $y_2 = \frac{u_1}{1+u_2}$ and the Jacobian is

$$J = \begin{vmatrix} \frac{u_2}{1+u_2} & \frac{u_1}{(1+u_2)^2} \\ \frac{1}{1+u_2} & \frac{-u_1}{(1+u_2)^2} \end{vmatrix} = \left| \frac{-u_1}{(1+u_2)^2} \right|$$

So, the joint density of U_1 and U_2 is

$$f(u_1, u_2) = \frac{1}{\beta^2} e^{-u_1/\beta} \frac{u_1}{(1+u_2)^2}, \; u_1 > 0, \; u_2 > 0.$$

b. Yes, U_1 and U_2 are independent since the joint density factors and the support is rectangular (Theorem 5.5).

6.72 Since the distribution function is $F(y) = y$ for $0 \le y \le 1$,

a. $g_{(1)}(u) = 2(1-u)$, $0 \le u \le 1$.

b. Since the above is a beta density with $\alpha = 1$ and $\beta = 2$, $E(U_1) = 1/3$, $V(U_1) = 1/18$.

6.73 Following Ex. 6.72,

a. $g_{(2)}(u) = 2u$, $0 \le u \le 1$.

b. Since the above is a beta density with $\alpha = 2$ and $\beta = 1$, $E(U_2) = 2/3$, $V(U_2) = 1/18$.

6.74 Since the distribution function is $F(y) = y/\theta$ for $0 \le y \le \theta$,

a. $G_{(n)}(y) = (y/\theta)^n$, $0 \le y \le \theta$.

b. $g_{(n)}(y) = G'_{(n)}(y) = ny^{n-1}/\theta^n$, $0 \le y \le \theta$.

c. It is easily shown that $E(Y_{(n)}) = \frac{n}{n+1}\theta$, $V(Y_{(n)}) = \frac{n\theta^2}{(n+1)^2(n+2)}$.

6.75 Following Ex. 6.74, the required probability is $P(Y_{(n)} < 10) = (10/15)^5 = .1317$.

6.76 Following Ex. 6.74 with $f(y) = 1/\theta$ for $0 \le y \le \theta$,

a. By Theorem 6.5, $g_{(k)}(y) = \frac{n!}{(k-1)!(n-k)!}\left(\frac{y}{\theta}\right)^{k-1}\left(\frac{\theta-y}{\theta}\right)^{n-k}\frac{1}{\theta} = \frac{n!}{(k-1)!(n-k)!}\frac{y^{k-1}(\theta-y)^{n-k}}{\theta^n}$, $0 \le y \le \theta$.

b. $E(Y_{(k)}) = \frac{n!}{(k-1)!(n-k)!}\int_0^\theta \frac{y^k(\theta-y)^{n-k}}{\theta^n}dy = \frac{k}{n+1}\frac{\Gamma(n+2)}{\Gamma(k+1)\Gamma(n-k+1)}\int_0^\theta \left(\frac{y}{\theta}\right)^k\left(1-\frac{y}{\theta}\right)^{n-k}dy$. To evaluate this integral, apply the transformation $z = \frac{y}{\theta}$ and relate the resulting integral to that of a beta density with $\alpha = k + 1$ and $\beta = n - k + 1$. Thus, $E(Y_{(k)}) = \frac{k}{n+1}\theta$.

c. Using the same techniques in part b above, it can be shown that $E(Y_{(k)}^2) = \frac{k(k+1)}{(n+1)(n+2)}\theta^2$ so that $V(Y_{(k)}) = \frac{(n-k+1)k}{(n+1)^2(n+2)}\theta^2$.

d. $E(Y_{(k)} - Y_{(k-1)}) = E(Y_{(k)}) - E(Y_{(k-1)}) = \frac{k}{n+1}\theta - \frac{k-1}{n+1}\theta = \frac{1}{n+1}\theta$. Note that this is constant for all k, so that the expected order statistics are equally spaced.

6.77 **a.** Using Theorem 6.5, the joint density of $Y_{(j)}$ and $Y_{(k)}$ is given by

$$g_{(j)(k)}(y_j, y_k) = \frac{n!}{(j-1)!(k-1-j)!(n-k)!}\left(\frac{y_j}{\theta}\right)^{j-1}\left(\frac{y_k}{\theta}-\frac{y_j}{\theta}\right)^{k-1-j}\left(1-\frac{y_k}{\theta}\right)^{n-k}\left(\frac{1}{\theta}\right)^2,\ 0 \le y_j \le y_k \le \theta.$$

b. $\text{Cov}(Y_{(j)}, Y_{(k)}) = E(Y_{(j)}Y_{(k)}) - E(Y_{(j)})E(Y_{(k)})$. The expectations $E(Y_{(j)})$ and $E(Y_{(k)})$ were derived in Ex. 6.76. To find $E(Y_{(j)}Y_{(k)})$, let $u = y_j/\theta$ and $v = y_k/\theta$ and write

$$E(Y_{(j)}Y_{(k)}) = c\theta\int_0^1\int_0^v u^j(v-u)^{k-1-j}v(1-v)^{n-k}\,du\,dv,$$

where $c = \frac{n!}{(j-1)!(k-1-j)!(n-k)!}$. Now, let $w = u/v$ so $u = wv$ and $du = vdw$. Then, the integral is

$$c\theta^2\left[\int_0^1 u^{k+1}(1-u)^{n-k}du\right]\left[\int_0^1 w^j(1-w)^{k-1-j}dw\right] = c\theta^2\left[B(k+2, n-k+1)\right]\left[B(j+1, k-j)\right].$$

Simplifying, this is $\frac{(k+1)j}{(n+1)(n+2)}\theta^2$. Thus, $\text{Cov}(Y_{(j)}, Y_{(k)}) = \frac{(k+1)j}{(n+1)(n+2)}\theta^2 - \frac{jk}{(n+1)^2}\theta^2 = \frac{n-k+1}{(n+1)^2(n+2)}\theta^2$.

c. $V(Y_{(k)} - Y_{(j)}) = V(Y_{(k)}) + V(Y_{(j)}) - 2\text{Cov}(Y_{(j)}, Y_{(k)})$

$$= \frac{(n-k+1)k}{(n+1)^2(n+2)}\theta^2 + \frac{(n-j+1)j}{(n+1)^2(n+2)}\theta^2 - \frac{2(n-k+1)}{(n+1)^2(n+2)}\theta^2 = \frac{(k-j)(n-k+k+1)}{(n+1)^2(n+2)}\theta^2.$$

6.78 From Ex. 6.76 with $\theta = 1$, $g_{(k)}(y) = \frac{n!}{(k-1)!(n-k)!}y^{k-1}(1-y)^{n-k} = \frac{\Gamma(n+1)}{\Gamma(k)\Gamma(n-k+1)}y^{k-1}(1-y)^{n-k}$. Since $0 \le y \le 1$, this is the beta density as described.

6.79 The joint density of $Y_{(1)}$ and $Y_{(n)}$ is given by (see Ex. 6.77 with $j = 1$, $k = n$),

$$g_{(1)(n)}(y_1, y_n) = n(n-1)\left(\frac{y_n}{\theta}-\frac{y_1}{\theta}\right)^n\left(\frac{1}{\theta}\right)^2 = n(n-1)\left(\frac{1}{\theta}\right)^n(y_n - y_1)^{n-2},\ 0 \le y_1 \le y_n \le \theta.$$

Applying the transformation $U = Y_{(1)}/Y_{(n)}$ and $V = Y_{(n)}$, we have that $y_1 = uv$, $y_n = v$ and the Jacobian of transformation is v. Thus,

$$f(u, v) = n(n-1)\left(\frac{1}{\theta}\right)^n(v-uv)^{n-2}v = n(n-1)\left(\frac{1}{\theta}\right)^n(1-u)^{n-2}v^{n-1},\ 0 \le u \le 1,\ 0 \le v \le \theta.$$

Since this joint density factors into separate functions of u and v and the support is rectangular, thus $Y_{(1)}/Y_{(n)}$ and $V = Y_{(n)}$ are independent.

6.80 The density and distribution function for Y are $f(y) = 6y(1-y)$ and $F(y) = 3y^2 - 2y^3$, respectively, for $0 \le y \le 1$.

a. $G_{(n)}(y) = \left(3y^2 - 2y^3\right)^n$, $0 \le y \le 1$.

b. $g_{(n)}(y) = G'_{(n)}(y) = n\left(3y^2 - 2y^3\right)^{n-1}(6y - 6y^2) = 6ny(1-y)\left(3y^2 - 2y^3\right)^{n-1}$, $0 \le y \le 1$.

c. Using the above density with $n = 2$, it is found that $E(Y_{(2)})$=.6286.

6.81 **a.** With $f(y) = \frac{1}{\beta}e^{-y/\beta}$ and $F(y) = 1 - e^{-y/\beta}$, $y \ge 0$:

$$g_{(1)}(y) = n\left[e^{-y/\beta}\right]^{n-1}\tfrac{1}{\beta}e^{-y/\beta} = \tfrac{n}{\beta}e^{-ny/\beta},\ y \ge 0.$$

This is the exponential density with mean β/n.

b. With $n = 5$, $\beta = 2$, $Y_{(1)}$ has and exponential distribution with mean .4. Thus

$$P(Y_{(1)} \le 3.6) = 1 - e^{-9} = .99988.$$

6.82 Note that the distribution function for the largest order statistic is

$$G_{(n)}(y) = \left[F(y)\right]^n = \left[1 - e^{-y/\beta}\right]^n,\ y \ge 0.$$

It is easily shown that the median m is given by $m = \phi_{.5} = \beta\ln 2$. Now,

$$P(Y_{(m)} > m) = 1 - P(Y_{(m)} \le m) = 1 - \left[F(\beta\ln 2)\right]^n = 1 - (.5)^n.$$

6.83 Since $F(m) = P(Y \le m) = .5$, $P(Y_{(m)} > m) = 1 - P(Y_{(n)} \le m) = 1 - G_{(n)}(m) = 1 - (.5)^n$. So, the answer holds regardless of the continuous distribution.

6.84 The distribution function for the Weibull is $F(y) = 1 - e^{-y^m/\alpha}$, $y > 0$. Thus, the distribution function for $Y_{(1)}$, the smallest order statistic, is given by

$$G_{(1)}(y) = 1 - \left[1 - F(y)\right]^n = 1 - \left[e^{-y^m/\alpha}\right]^n = 1 - e^{-ny^m/\alpha},\ y > 0.$$

This is the Weibull distribution function with shape parameter m and scale parameter α/n.

6.85 Using Theorem 6.5, the joint density of $Y_{(1)}$ and $Y_{(2)}$ is given by

$$g_{(1)(2)}(y_1, y_2) = 2,\ 0 \le y_1 \le y_2 \le 1.$$

$$\text{Thus, } P(2Y_{(1)} < Y_{(2)}) = \int_0^{1/2}\int_{2y_1}^{1} 2\,dy_2\,dy_1 = .5.$$

6.86 Using Theorem 6.5 with $f(y) = \frac{1}{\beta}e^{-y/\beta}$ and $F(y) = 1 - e^{-y/\beta}$, $y \ge 0$:

a. $g_{(k)}(y) = \frac{n!}{(k-1)!(n-k)!}\left(1 - e^{-y/\beta}\right)^{k-1}\left(e^{-y/\beta}\right)^{n-k}\frac{e^{-y/\beta}}{\beta} = \frac{n!}{(k-1)!(n-k)!}\left(1 - e^{-y/\beta}\right)^{k-1}\left(e^{-y/\beta}\right)^{n-k+1}\frac{1}{\beta}$, $y \ge 0$.

b. $g_{(j)(k)}(y_j, y_k) = \frac{n!}{(j-1)!(k-1-j)!(n-k)!}\left(1 - e^{-y_j/\beta}\right)^{j-1}\left(e^{-y_j/\beta} - e^{-y_k/\beta}\right)^{k-1-j}\left(e^{-y_k/\beta}\right)^{n-k+1}\frac{1}{\beta^2}e^{-y_j/\beta}$,
$0 \le y_j \le y_k < \infty$.

6.87 For this problem, we need the distribution of $Y_{(1)}$ (similar to Ex. 6.72). The distribution function of Y is

$$F(y) = P(Y \le y) = \int_4^y (1/2)e^{-(1/2)(t-4)}dy = 1 - e^{-(1/2)(y-4)} \text{ , } y \ge 4.$$

a. $g_{(1)}(y) = 2\left[e^{-(1/2)(y-4)}\right]^1 \frac{1}{2}e^{-(1/2)(y-4)} = e^{-(y-4)}, y \ge 4.$

b. $E(Y_{(1)}) = 5$.

6.88 This is somewhat of a generalization of Ex. 6.87. The distribution function of Y is

$$F(y) = P(Y \le y) = \int_\theta^y e^{-(t-\theta)}dy = 1 - e^{-(y-\theta)} \text{ , } y > \theta.$$

a. $g_{(1)}(y) = n\left[e^{-(y-\theta)}\right]^{n-1} e^{-(y-\theta)} = ne^{-n(y-4)}, y > \theta.$

b. $E(Y_{(1)}) = \frac{1}{n} + \theta$.

6.89 Theorem 6.5 gives the joint density of $Y_{(1)}$ and $Y_{(n)}$ is given by (also see Ex. 6.79)

$$g_{(1)(n)}(y_1, y_n) = n(n-1)(y_n - y_1)^{n-2}, \ 0 \le y_1 \le y_n \le 1.$$

Using the method of transformations, let $R = Y_{(n)} - Y_{(1)}$ and $S = Y_{(1)}$. The inverse transformations are $y_1 = s$ and $y_n = r + s$ and Jacobian of transformation is 1. Thus, the joint density of R and S is given by

$$f(r,s) = n(n-1)(r+s-s)^{n-2} = n(n-1)r^{n-2}, \ 0 \le s \le 1 - r \le 1.$$

(Note that since $r = y_n - y_1$, $r \le 1 - y_1$ or equivalently $r \le 1 - s$ and then $s \le 1 - r$).
The marginal density of R is then

$$f_R(r) = \int_0^{1-r} n(n-1)r^{n-2}ds = n(n-1)r^{n-2}(1-r), \ 0 \le r \le 1.$$

FYI, this is a beta density with $\alpha = n - 1$ and $\beta = 2$.

6.90 Since the points on the interval $(0, t)$ at which the calls occur are uniformly distributed, we have that $F(w) = w/t$, $0 \le w \le t$.

a. The distribution of $W_{(4)}$ is $G_{(4)}(w) = [F(w)]^4 = w^4/t^4$, $0 \le w \le t$. Thus $P(W_{(4)} \le 1) = G_{(4)}(1) = 1/16$.

b. With $t = 2$, $E(W_{(4)}) = \int_0^2 4w^4/2^4 dw = \int_0^2 w^4/4dw = 1.6$.

6.91 With the exponential distribution with mean θ, we have $f(y) = \frac{1}{\theta}e^{-y/\theta}$, $F(y) = 1 - e^{-y/\theta}$, for $y \ge 0$.

a. Using Theorem 6.5, the joint distribution of order statistics $W_{(j)}$ and $W_{(j-1)}$ is given by

$$g_{(j-1)(j)}(w_{j-1}, w_j) = \frac{n!}{(j-2)!(n-j)!}\left(1 - e^{-w_{j-1}/\theta}\right)^{j-2}\left(e^{-w_j/\theta}\right)^{n-j}\frac{1}{\theta^2}\left(e^{-(w_{j-1}+w_j)/\theta}\right), \ 0 \le w_{j-1} \le w_j < \infty.$$

Define the random variables $S = W_{(j-1)}$, $T_j = W_{(j)} - W_{(j-1)}$. The inverse transformations are $w_{j-1} = s$ and $w_j = t_j + s$ and Jacobian of transformation is 1. Thus, the joint density of S and T_j is given by

$$f(s,t_j) = \tfrac{n!}{(j-2)!(n-j)!}\left(1-e^{-s/\theta}\right)^{j-2}\left(e^{-(t_j+s)/\theta}\right)^{n-j}\tfrac{1}{\theta^2}\left(e^{-(2s+t_j)/\theta}\right)$$
$$= \tfrac{n!}{(j-2)!(n-j)!}e^{-(n-j+1)t_j/\theta}\tfrac{1}{\theta^2}\left(1-e^{-s/\theta}\right)^{j-2}\left(e^{-(n-j+2)s/\theta}\right),\ s \ge 0,\ t_j \ge 0.$$

The marginal density of T_j is then

$$f_{T_j}(t_j) = \tfrac{n!}{(j-2)!(n-j)!}e^{-(n-j+1)t_j/\theta}\tfrac{1}{\theta^2}\int_0^\infty \left(1-e^{-s/\theta}\right)^{j-2}\left(e^{-(n-j+2)s/\theta}\right)ds.$$

Employ the change of variables $u = e^{-s/\theta}$ and the above integral becomes the integral of a scaled beta density. Evaluating this, the marginal density becomes

$$f_{T_j}(t_j) = \tfrac{n-j+1}{\theta}e^{-(n-j+1)t_j/\theta},\ t_j \ge 0.$$

This is the density of an exponential distribution with mean $\theta/(n-j+1)$.

b. Observe that

$$\sum_{j=1}^{r}(n-j+1)T_j = nW_1 + (n-1)(W_2-W_1) + (n-2)(W_3-W_2) + \ldots + (n-r+1)(W_r - W_{r-1})$$
$$= W_1 + W_2 + \ldots + W_{r-1} + (n-r+1)W_r = \sum\nolimits_{j=1}^{r} W_j + (n-r)W_r = U_r.$$

Hence, $E(U_r) = \sum_{j=1}^{r}(n-r+1)E(T_j) = r\theta$.

6.92 By Theorem 6.3, U will have a normal distribution with mean $(1/2)(\mu - 3\mu) = -\mu$ and variance $(1/4)(\sigma^2 + 9\sigma^2) = 2.5\sigma^2$.

6.93 By independence, the joint distribution of I and R is $f(i,r) = 2r$, $0 \le i \le 1$ and $0 \le r \le 1$. To find the density for W, fix $R = r$. Then, $W = I^2r$ so $I = \sqrt{W/r}$ and $\left|\frac{di}{dw}\right| = \frac{1}{2r}\left(\frac{w}{r}\right)^{-1/2}$ for the range $0 \le w \le r \le 1$. Thus, $f(w,r) = \sqrt{r/w}$ and

$$f(w) = \int_w^1 \sqrt{r/w}\,dr = \tfrac{2}{3}\left(\tfrac{1}{\sqrt{w}} - w\right),\ 0 \le w \le 1.$$

6.94 Note that Y_1 and Y_2 have identical gamma distributions with $\alpha = 2$, $\beta = 2$. The mgf is

$$m(t) = (1-2t)^{-2},\ t < 1/2.$$

The mgf for $U = (Y_1 + Y_2)/2$ is

$$m_U(t) = E(e^{tU}) = E(e^{t(Y_1+Y_2)/2}) = m(t/2)m(t/2) = (1-t)^{-4}.$$

This is the mgf for a gamma distribution with $\alpha = 4$ and $\beta = 1$, so that is the distribution of U.

6.95 By independence, $f(y_1, y_2) = 1$, $0 \le y_1 \le 0$, $0 \le y_2 \le 1$.

a. Consider the joint distribution of $U_1 = Y_1/Y_2$ and $V = Y_2$. Fixing V at v, we can write $U_1 = Y_1/v$. Then, $Y_1 = vU_1$ and $\frac{dy_1}{du} = v$. The joint density of U_1 and V is $g(u,v) = v$. The ranges of u and v are as follows:

- if $y_1 \leq y_2$, then $0 \leq u \leq 1$ and $0 \leq v \leq 1$
- if $y_1 > y_2$, then u has a minimum value of 1 and a maximum at $1/y_2 = 1/v$. Similarly, $0 \leq v \leq 1$

So, the marginal distribution of U_1 is given by

$$f_{U_1}(u) = \begin{cases} \int_0^1 v\,dv = \frac{1}{2} & 0 \leq u \leq 1 \\ \int_0^{1/u} v\,dv = \frac{1}{2u^2} & u > 1 \end{cases}.$$

b. Consider the joint distribution of $U_2 = -\ln(Y_1Y_2)$ and $V = Y_1$. Fixing V at v, we can write $U_2 = -\ln(vY_2)$. Then, $Y_2 = e^{-U_2}/v$ and $\frac{dy_2}{du} = -e^{-u}/v$. The joint density of U_2 and V is $g(u,v) = -e^{-u}/v$, with $-\ln v \leq u < \infty$ and $0 \leq v \leq 1$. Or, written another way, $e^{-u} \leq v \leq 1$.

So, the marginal distribution of U_2 is given by

$$f_{U_2}(u) = \int_{e^{-u}}^{1} -e^{-u}/v\,dv = ue^{-u},\ 0 \leq u.$$

c. Same as Ex. 6.35.

6.96 Note that $P(Y_1 > Y_2) = P(Y_1 - Y_2 > 0)$. By Theorem 6.3, $Y_1 - Y_2$ has a normal distribution with mean $5 - 4 = 1$ and variance $1 + 3 = 4$. Thus,

$$P(Y_1 - Y_2 > 0) = P(Z > -1/2) = .6915.$$

6.97 The probability mass functions for Y_1 and Y_2 are:

y_1	0	1	2	3	4
$p_1(y_1)$	.4096	.4096	.1536	.0256	.0016

y_2	0	1	2	3
$p_2(y_2)$	.125	.375	.375	.125

Note that $W = Y_1 + Y_2$ is a random variable with support (0, 1, 2, 3, 4, 5, 6, 7). Using the hint given in the problem, the mass function for W is given by

w	$p(w)$
0	$p_1(0)p_2(0) = .4096(.125) =$ **.0512**
1	$p_1(0)p_2(1) + p_1(1)p_2(0) = .4096(.375) + .4096(.125) =$ **.2048**
2	$p_1(0)p_2(2) + p_1(2)p_2(0) + p_1(1)p_2(1) = .4096(.375) + .1536(.125) + .4096(.375) =$ **.3264**
3	$p_1(0)p_2(3) + p_1(3)p_2(0) + p_1(1)p_2(2) + p_1(2)p_2(1) = .4096(.125) + .0256(.125) + .4096(.375) + .1536(.375) =$ **.2656**
4	$p_1(1)p_2(3) + p_1(3)p_2(1) + p_1(2)p_2(2) + p_1(4)p_2(0) = .4096(.125) + .0256(.375) + .1536(.375) + .0016(.125) =$ **.1186**
5	$p_1(2)p_2(3) + p_1(3)p_2(2) + p_1(4)p_2(1) = .1536(.125) + .0256(.375) + .0016(.375) =$ **.0294**
6	$p_1(4)p_2(2) + p_1(3)p_2(3) = .0016(.375) + .0256(.125) =$ **.0038**
7	$p_1(4)p_2(3) = .0016(.125) =$ **.0002**

Check: .0512 + .2048 + .3264 + .2656 + .1186 + .0294 + .0038 + .0002 = 1.

6.98 The joint distribution of Y_1 and Y_2 is $f(y_1, y_2) = e^{-(y_1+y_2)}$, $y_1 > 0$, $y_2 > 0$. Let $U_1 = \frac{Y_1}{Y_1+Y_2}$, $U_2 = Y_2$. The inverse transformations are $y_1 = u_1u_2/(1 - u_1)$ and $y_2 = u_2$ so the Jacobian of transformation is

$$J = \begin{vmatrix} \frac{u_2}{(1-u_1)^2} & \frac{u_1}{1-u_1} \\ 0 & 1 \end{vmatrix} = \frac{u_2}{(1-u_1)^2}.$$

Thus, the joint distribution of U_1 and U_2 is

$$f(u_1, u_2) = e^{-[u_1u_2/(1-u_1)+u_2]}\frac{u_2}{(1-u_1)^2} = e^{-[u_2/(1-u_1)}\frac{u_2}{(1-u_1)^2},\ 0 \le u_1 \le 1,\ u_2 > 0.$$

Therefore, the marginal distribution for U_1 is

$$f_{U_1}(u_1) = \int_0^\infty e^{-[u_2/(1-u_1)}\frac{u_2}{(1-u_1)^2}du_2 = 1,\ 0 \le u_1 \le 1.$$

Note that the integrand is a gamma density function with $\alpha = 1$, $\beta = 1 - u_1$.

6.99 This is a special case of Example 6.14 and Ex. 6.63.

6.100 Recall that by Ex. 6.81, $Y_{(1)}$ is exponential with mean 15/5 = 3.
a. $P(Y_{(1)} > 9) = e^{-3}$.
b. $P(Y_{(1)} < 12) = 1 - e^{-4}$.

6.101 If we let $(A, B) = (-1, 1)$ and $T = 0$, the density function for X, the landing point is

$$f(x) = 1/2,\ -1 < x < 1.$$

We must find the distribution of $U = |X|$. Therefore,

$$F_U(u) = P(U \le u) = P(|X| \le u) = P(-u \le X \le u) = [u - (-u)]/2 = u.$$

So, $f_U(u) = F'_U(u) = 1$, $0 \le u \le 1$. Therefore, U has a uniform distribution on (0, 1).

6.102 Define Y_1 = point chosen for sentry 1 and Y_2 = point chosen for sentry 2. Both points are chosen along a one–mile stretch of highway, so assuming independent uniform distributions on (0, 1), the joint distribution for Y_1 and Y_2 is

$$f(y_1, y_2) = 1,\ 0 \le y_1 \le 1,\ 0 \le y_2 \le 1.$$

The probability of interest is $P(|Y_1 - Y_2| < \frac{1}{2})$. This is most easily solved using geometric considerations (similar to material in Chapter 5): $P(|Y_1 - Y_2| < \frac{1}{2}) = .75$ (this can easily be found by considering the complement of the event).

6.103 The joint distribution of Y_1 and Y_2 is $f(y_1, y_2) = \frac{1}{2\pi}e^{-(y_1^2+y_2^2)/2}$. Considering the transformations $U_1 = Y_1/Y_2$ and $U_2 = Y_2$. With $y_1 = u_1u_2$ and $y_2 = |u_2|$, the Jacobian of transformation is u_2 so that the joint density of U_1 and U_2 is

$$f(u_1, u_2) = \frac{1}{2\pi}|u_2|e^{-[(u_1u_2)^2+u_2^2]/2} = \frac{1}{2\pi}|u_2|e^{-[u_2^2(1+u_1^2)]/2}.$$

The marginal density of U_1 is

$$f_{U_1}(u_1) = \int_{-\infty}^{\infty}\frac{1}{2\pi}|u_2|e^{-[u_2^2(1+u_1^2)]/2}du_2 = \int_0^\infty \frac{1}{\pi}u_2e^{-[u_2^2(1+u_1^2)]/2}du_2.$$

Using the change of variables $v = u_2^2$ so that $du_2 = \frac{1}{2\sqrt{v}}dv$ gives the integral

$$f_{U_1}(u_1) = \int_0^\infty \tfrac{1}{2\pi} e^{-[v(1+u_1^2)]/2} dv = \tfrac{1}{\pi(1+u_1^2)}, \infty < u_1 < \infty.$$

The last expression above comes from noting the integrand is related an exponential density with mean $2/(1+u_1^2)$. The distribution of U_1 is called the Cauchy distribution.

6.104 **a.** The event $\{Y_1 = Y_2\}$ occurs if

$\{(Y_1 = 1, Y_2 = 1), (Y_1 = 2, Y_2 = 2), (Y_1 = 3, Y_2 = 3), \ldots\}$

So, since the probability mass function for the geometric is given by $p(y) = p(1-p)^{y-1}$, we can find the probability of this event by

$$P(Y_1 = Y_2) = p(1)^2 + p(2)^2 + p(3)^2 \ldots = p^2 + p^2(1-p)^2 + p^2(1-p)^4 + \ldots$$

$$= p^2 \sum_{j=0}^{\infty} (1-p)^{2j} = \frac{p^2}{1-(1-p)^2} = \frac{p}{2-p}.$$

b. Similar to part a, the event $\{Y_1 - Y_2 = 1\} = \{Y_1 = Y_2 + 1\}$ occurs if

$\{(Y_1 = 2, Y_2 = 1), (Y_1 = 3, Y_2 = 2), (Y_1 = 4, Y_2 = 3), \ldots\}$

Thus,

$$P(Y_1 - Y_2 = 1) = p(2)\,p(1) + p(3)\,p(2) + p(4)\,p(3) + \ldots$$

$$= p^2(1-p) + p^2(1-p)^3 + p^2(1-p)^5 + \ldots = \frac{p(1-p)}{2-p}.$$

c. Define $U = Y_1 - Y_2$. To find $p_U(u) = P(U = u)$, assume first that $u > 0$. Thus,

$$P(U = u) = P(Y_1 - Y_2 = u) = \sum_{y_2=1}^{\infty} P(Y_1 = u + y_2) P(Y_2 = y_2) = \sum_{y_2=1}^{\infty} p(1-p)^{u+y_2-1} p(1-p)^{y_2-1}$$

$$= p^2(1-p)^u \sum_{y_2=1}^{\infty} (1-p)^{2(y_2-1)} = p^2(1-p)^u \sum_{x=1}^{\infty} (1-p)^{2x} = \frac{p(1-p)^u}{2-p}.$$

If $u < 0$, proceed similarly with $y_2 = y_1 - u$ to obtain $P(U = u) = \dfrac{p(1-p)^{-u}}{2-u}$. These two results can be combined to yield $p_U(u) = P(U = u) = \dfrac{p(1-p)^{|u|}}{2-u}$, $u = 0, \pm 1, \pm 2, \ldots$.

6.105 The inverse transformation is $y = 1/u - 1$. Then,

$$f_U(u) = \tfrac{1}{B(\alpha,\beta)} \left(\tfrac{1-u}{u}\right)^{\alpha-1} u^{\alpha+\beta} \tfrac{1}{u^2} = \tfrac{1}{B(\alpha,\beta)} u^{\beta-1}(1-u)^{\alpha-1}, 0 < u < 1.$$

This is the beta distribution with parameters β and α.

6.106 Recall that the distribution function for a continuous random variable is monotonic increasing and returns values on [0, 1]. Thus, the random variable $U = F(Y)$ has support on (0, 1) and has distribution function

$$F_U(u) = P(U \le u) = P(F(Y) \le u) = P(Y \le F^{-1}(u)) = F[F^{-1}(u)] = u, \ 0 \le u \le 1.$$

The density function is $f_U(u) = F_U'(u) = 1$, $0 \le u \le 1$, which is the density for the uniform distribution on (0, 1).

6.107 The density function for Y is $f(y) = \frac{1}{4}$, $-1 \le y \le 3$. For $U = Y^2$, the density function for U is given by

$$f_U(u) = \tfrac{1}{2\sqrt{u}}\left[f(\sqrt{u}) + f(-\sqrt{u})\right],$$

as with Example 6.4. If $-1 \le y \le 3$, then $0 \le u \le 9$. However, if $1 \le u \le 9$, $f(-\sqrt{u})$ is not positive. Therefore,

$$f_U(u) = \begin{cases} \tfrac{1}{2\sqrt{u}}\left(\tfrac{1}{4} + \tfrac{1}{4}\right) = \tfrac{1}{4\sqrt{u}} & 0 \le u < 1 \\ \tfrac{1}{2\sqrt{u}}\left(\tfrac{1}{4} + 0\right) = \tfrac{1}{8\sqrt{u}} & 1 \le u \le 9 \end{cases}.$$

6.108 The system will operate provided that C_1 *and* C_2 function *and* C_3 *or* C_4 function. That is, defining the system as S and using set notation, we have

$$S = (C_1 \cap C_2) \cap (C_3 \cup C_4) = (C_1 \cap C_2 \cap C_3) \cup (C_1 \cap C_2 \cap C_4).$$

At some y, the probability that a component is operational is given by $1 - F(y)$. Since the components are independent, we have

$$P(S) = P(C_1 \cap C_2 \cap C_3) + P(C_1 \cap C_2 \cap C_4) - P(C_1 \cap C_2 \cap C_3 \cap C_4).$$

Therefore, the reliability of the system is given by

$$[1 - F(y)]^3 + [1 - F(y)]^3 - [1 - F(y)]^4 = [1 - F(y)]^3[1 + F(y)].$$

6.109 Let C_3 be the production cost. Then U, the profit function (per gallon), is

$$U = \begin{cases} C_1 - C_3 & \tfrac{1}{3} < Y < \tfrac{2}{3} \\ C_2 - C_3 & \text{otherwise} \end{cases}.$$

So, U is a discrete random variable with probability mass function

$$P(U = C_1 - C_3) = \int_{1/3}^{2/3} 20y^3(1-y)dy = .4156.$$

$$P(U = C_2 - C_3) = 1 - ,4156 = .5844.$$

6.110 **a.** Let X = next gap time. Then, $P(X \le 60) = F_X(60) = 1 - e^{-6}$.

b. If the next four gap times are assumed to be independent, then $Y = X_1 + X_2 + X_3 + X_4$ has a gamma distribution with $\alpha = 4$ and $\beta = 10$. Thus,

$$f(y) = \tfrac{1}{\Gamma(4)10^4} y^3 e^{-y/10},\ y \ge 0.$$

6.111 **a.** Let $U = \ln Y$. So, $\frac{du}{dy} = \frac{1}{y}$ and with $f_U(u)$ denoting the normal density function,

$$f_Y(y) = \tfrac{1}{y} f_U(\ln y) = \tfrac{1}{y\sigma\sqrt{2\pi}} \exp\left[-\tfrac{(\ln y - \mu)^2}{2\sigma^2}\right],\ y > 0.$$

b. Note that $E(Y) = E(e^U) = m_U(1) = e^{\mu + \sigma^2/2}$, where $m_U(t)$ denotes the mgf for U. Also, $E(Y^2) = E(e^{2U}) = m_U(2) = e^{2\mu + 2\sigma^2}$ so $V(Y) = e^{2\mu+2\sigma^2} - \left(e^{\mu+\sigma^2/2}\right)^2 = e^{2\mu+\sigma^2}\left(e^{\sigma^2} - 1\right)$.

6.112 **a.** Let $U = \ln Y$. So, $\frac{du}{dy} = \frac{1}{y}$ and with $f_U(u)$ denoting the gamma density function,

$$f_Y(y) = \tfrac{1}{y} f_U(\ln y) = \tfrac{1}{y\Gamma(\alpha)\beta^\alpha}(\ln y)^{\alpha-1} e^{-(\ln y)/\beta} = \tfrac{1}{\Gamma(\alpha)\beta^\alpha}(\ln y)^{\alpha-1} y^{-(1+\beta)/\beta}, y > 1.$$

b. Similar to Ex. 6.111: $E(Y) = E(e^U) = m_U(1) = (1-\beta)^{-\alpha}$, $\beta < 1$, where $m_U(t)$ denotes the mgf for U.

c. $E(Y^2) = E(e^{2U}) = m_U(2) = (1-2\beta)^{-\alpha}$, $\beta < .5$, so that $V(Y) = (1-2\beta)^{-\alpha} - (1-\beta)^{-2\alpha}$.

6.113 **a.** The inverse transformations are $y_1 = u_1/u_2$ and $y_2 = u_2$ so that the Jacobian of transformation is $1/|u_2|$. Thus, the joint density of U_1 and U_2 is given by

$$f_{U_1,U_2}(u_1,u_2) = f_{Y_1,Y_2}(u_1/u_2,u_2)\frac{1}{|u_2|}.$$

b. The marginal density is found using standard techniques.

c. If Y_1 and Y_2 are independent, the joint density will factor into the product of the marginals, and this is applied to part b above.

6.114 The volume of the sphere is $V = \frac{4}{3}\pi R^3$, or $R = \left(\frac{3}{4\pi}V\right)^{1/3}$, so that $\frac{dr}{dv} = \frac{1}{3}\left(\frac{3}{4\pi}\right)^{1/3} v^{-2/3}$. Thus, $f_V(v) = \frac{2}{3}\left(\frac{3}{4\pi}\right)^{2/3} v^{-1/3}$, $0 \le v \le \frac{4}{3}\pi$.

6.115 **a.** Let R = distance from a randomly chosen point to the nearest particle. Therefore,

$P(R > r) = P(\text{no particles in the sphere of radius } r) = P(Y = 0 \text{ for volume } \frac{4}{3}\pi r^3)$.

Since Y = # of particles in a volume v has a Poisson distribution with mean λv, we have

$$P(R > r) = P(Y = 0) = e^{-(4/3)\pi r^3 \lambda}, \ r > 0.$$

Therefore, the distribution function for R is $F(r) = 1 - P(R > r) = 1 - e^{-(4/3)\pi r^3 \lambda}$ and the density function is

$$f(r) = F'(r) = 4\lambda\pi r^2 e^{-(4/3)\lambda\pi r^3}, \ r > 0.$$

b. Let $U = R^3$. Then, $R = U^{1/3}$ and $\frac{dr}{du} = \frac{1}{3}u^{-2/3}$. Thus,

$$f_U(u) = \tfrac{4\lambda\pi}{3} e^{-(4\lambda\pi/3)u}, \ u > 0.$$

This is the exponential density with mean $\frac{3}{4\lambda\pi}$.

6.116 **a.** The inverse transformations are $y_1 = u_1 + u_2$ and $y_2 = u_2$. The Jacobian of transformation is 1 so that the joint density of U_1 and U_2 is

$$f_{U_1,U_2}(u_1,u_2) = f_{Y_1,Y_2}(u_1+u_2,u_2).$$

b. The marginal density is found using standard techniques.

c. If Y_1 and Y_2 are independent, the joint density will factor into the product of the marginals, and this is applied to part b above.

Chapter 7: Sampling Distributions and the Central Limit Theorem

7.1 **a.** – **c.** Answers vary.
d. The histogram exhibits a mound shape. The sample mean should be close to 3.5 = μ
e. The standard deviation should be close to $\sigma/\sqrt{3} = 1.708/\sqrt{3} = .9860$.
f. Very similar pictures.

7.2 **a.** $P(\overline{Y} = 2) = P(W = 6) = p(4, 1, 1) + p(1, 4, 1) + p(1, 1, 4) + p(3, 2, 1) + p(3, 1, 2)$
$= p(2, 3, 1) + p(2, 1, 3) + p(1, 3, 2) + p(1, 2, 3) + p(2, 2, 2) = \frac{10}{216}$.
b. Answers vary, but the relative frequency should be fairly close.
c. The relative frequency should be even closer than what was observed in part b.

7.3 **a.** The histogram should be similar in shape, but this histogram has a smaller spread.
b. Answers vary.
c. The normal curve should approximate the histogram fairly well.

7.4 **a.** The histogram has a right–skewed shape. It appears to follow $p(y) = y/21$, $y = 1, \ldots, 6$.
b. From the Stat Report window, $\mu = 2.667$, $\sigma = 1.491$.
c. Answers vary.
d. i. It has a right–skewed shape. ii. The mean is larger, but the std. dev. is smaller.
e. i. sample mean = 2.667, sample std. dev = $1.491/\sqrt{12} = .4304$.
ii. The histogram is closely mound shaped.
iii. Very close indeed.

7.5 **a.** Answers vary.
b. Answers vary, but the means are probably not equal.
c. The sample mean values cluster around the population mean.
d. The theoretical standard deviation for the sample mean is $6.03/\sqrt{5} = 2.6967$.
e. The histogram has a mound shape.
f. Yes.

7.6 The larger the sample size, the smaller the spread of the histogram. The normal curves approximate the histograms equally well.

7.7 **a.** – **b.** Answers vary.
c. The mean should be close to the population variance
d. The sampling distribution is not mound–shaped for this case.
e. The theoretical density should fit well.
f. Yes, because the chi–square density is right–skewed.

7.8 **a.** $\sigma^2 = (6.03)^2 = 36.3609$.
b. The two histograms have similar shapes, but the histogram generated from the smaller sample size exhibits a greater spread. The means are similar (and close to the value found in part a). The theoretical density should fit well in both cases.
c. The histogram generated with n = 50 exhibits a mound shape. Here, the theoretical density is chi–square with $\nu = 50 - 1 = 49$ degrees of freedom (a large value).

7.9 **a.** $P(|\overline{Y} - \mu| \leq .3) = P(-1.2 \leq Z \leq 1.2) = .7698$.
b. $P(|\overline{Y} - \mu| \leq .3) = P(-.3\sqrt{n} \leq Z \leq .3\sqrt{n}) = 1 - 2P(Z > .3\sqrt{n})$. For n = 25, 36, 69, and 64, the probabilities are (respectively) .8664, .9284, .9642, and .9836.
c. The probabilities increase with n, which is intuitive since the variance of $\overline{Y}$ decreases with n.
d. Yes, these results are consistent since the probability was less than .95 for values of n less than 43.

7.10 **a.** $P(|\overline{Y} - \mu| \leq .3) = P(-.15\sqrt{n} \leq Z \leq .15\sqrt{n}) = 1 - 2P(Z > .15\sqrt{n})$. For $n = 9$, the probability is .3472 (a smaller value).
b.
For $n = 25$: $P(|\overline{Y} - \mu| \leq .3) = 1 - 2P(Z > .75) = .5468$
For $n = 36$: $P(|\overline{Y} - \mu| \leq .3) = 1 - 2P(Z > .9) = .6318$
For $n = 49$: $P(|\overline{Y} - \mu| \leq .3) = 1 - 2P(Z > 1.05) = .7062$
For $n = 64$: $P(|\overline{Y} - \mu| \leq .3) = 1 - 2P(Z > 1.2) = .7698$
c. The probabilities increase with n.
d. The probabilities are smaller with a larger standard deviation (more diffuse density).

7.11 $P(|\overline{Y} - \mu| \leq 2) = P(-1.5 \leq Z \leq 1.5) = 1 - 2P(Z > 1.5) = 1 - 2(.0668) = .8664$.

7.12 From Ex. 7.11, we require $P(|\overline{Y} - \mu| \leq 1) = P(-.25\sqrt{n} \leq Z \leq .25\sqrt{n}) = .90$. This will be solved by taking $.25\sqrt{n} = 1.645$, so $n = 43.296$. Hence, sample 44 trees.

7.13 Similar to Ex. 7.11: $P(|\overline{Y} - \mu| \leq .5) = P(-2.5 \leq Z \leq 2.5) = .9876$.

7.14 Similar to Ex. 7.12: we require $P(|\overline{Y} - \mu| \leq .5) = P(-\frac{.5}{\sqrt{.4}}\sqrt{n} \leq Z \leq \frac{.5}{\sqrt{.4}}\sqrt{n}) = .95$. Thus, $\frac{.5}{\sqrt{.4}}\sqrt{n} = 1.96$ so that $n = 6.15$. Hence, run 7 tests.

7.15 Using Theorems 6.3 and 7.1:
a. $E(\overline{X} - \overline{Y}) = \mu_1 - \mu_2$.
b. $V(\overline{X} - \overline{Y}) = \sigma_1^2/m + \sigma_2^2/n$.
c. It is required that $P(|\overline{X} - \overline{Y} - (\mu_1 - \mu_2)| \leq 1) = .95$. Using the result in part b for standardization with $n = m$, $\sigma_1^2 = 2$, and $\sigma_2^2 = 2.5$, we obtain $n = 17.29$. Thus, the two sample sizes should be at least 18.

7.16 Following the result in Ex. 7.15 and since the two population means are equal, we find
$$P(\overline{X}_A - \overline{Y}_B \geq 1) = P\left(\frac{\overline{X}_A - \overline{Y}_B}{\sqrt{\frac{.4}{10} + \frac{.8}{10}}} \geq \frac{1}{\sqrt{\frac{.4}{10} + \frac{.8}{10}}}\right) = P(Z \geq 2.89) = .0019.$$

7.17 $P(\sum_{i=1}^{6} Z_i^2 \leq 6) = .57681$.

7.18 $P(S^2 \geq 3) = P(9S^2 \geq 27) = .0014$.

7.19 Given that $s^2 = .065$ and $n = 10$, suppose $\sigma^2 = .04$. The probability of observing a value of s^2 that is as extreme or more so is given by

$$P(S^2 \geq .065) = P(9S^2/.04 \geq 9(.065)/.04) = P(9S^2/.04 \geq 14.925) = .10$$

Thus, it is fairly unlikely, so this casts some doubt that $\sigma^2 = .04$.

7.20 **a.** Using the fact that the chi–square distribution is a special case of the gamma distribution, $E(U) = \nu$, $V(U) = 2\nu$.
b. Using Theorem 7.3 and the result from part a:

$$E(S^2) = \tfrac{\sigma^2}{n-1}E(\tfrac{n-1}{\sigma^2}S^2) = \tfrac{\sigma^2}{n-1}(n-1) = \sigma^2.$$

$$V(S^2) = \left(\tfrac{\sigma^2}{n-1}\right)^2 V(\tfrac{n-1}{\sigma^2}S^2) = \left(\tfrac{\sigma^2}{n-1}\right)^2\left[2(n-1)\right] = 2\sigma^4/(n-1).$$

7.21 These values can be found by using percentiles from the chi–square distribution. With $\sigma^2 = 1.4$ and $n = 20$, $\frac{19}{1.4}S^2$ has a chi–square distribution with 19 degrees of freedom.

a. $P(S^2 \leq b) = P(\frac{n-1}{\sigma^2}S^2 \leq \frac{n-1}{\sigma^2}b) = P(\frac{19}{1.4}S^2 \leq \frac{19}{1.4}b) = .975$. It must be true that $\frac{19}{1.4}b = 32.8523$, the 97.5%-tile of this chi–square distribution, and so $b = 2.42$.

b. Similarly, $P(S^2 \geq a) = P(\frac{n-1}{\sigma^2}S^2 \geq \frac{n-1}{\sigma^2}a) = .974$. Thus, $\frac{19}{1.4}a = 8.96055$, the 2.5%-tile of this chi–square distribution, and so $a = .656$.

c. $P(a \leq S^2 \leq b) = .95$.

7.22 **a.** The corresponding gamma densities with parameters (α, β) are (5, 2), (20, 2), (40, 2), respectively.
b. The chi–square densities become more symmetric with larger values of ν.
c. They are the same.
d. Not surprising, given the answer to part b.

7.23 **a.** The three probabilities are found to be .44049, .47026, and .47898, respectively.
b. As the degrees of freedom increase, so do the probabilities.
c. Since the density is becoming more symmetric, the probability is approaching .5.

7.24 **a.** .05097
b. .05097
c. 1 – 2(.05097) = .8806.
d. The t–distribution with 5 degrees of freedom exhibits greater variability.

7.25 **a.** Using Table 5, $t_{.10} = 1.476$. Using the applet, $t_{.10} = 1.47588$.
b. The value $t_{.10}$ is the 90th percentile/quantile.
c. The values are 1.31042, 1.29582, 1.28865, respectively.
d. The t–distribution exhibits greater variability than the standard normal, so the percentiles are more extreme than $z_{.10}$.
e. As the degrees of freedom increase, the t–distribution approaches the standard normal distribution.

7.26 From Definition 7.2,
$P(g_1 \le \overline{Y} - \mu \le g_2) = P(\frac{\sqrt{n}g_1}{S} \le T \le \frac{\sqrt{n}g_2}{S})$ = .90. Thus, it must be true that $\frac{\sqrt{n}g_1}{S} = -t_{.05}$ and $\frac{\sqrt{n}g_2}{S} = t_{.05}$. Thus, with n = 9 and $t_{.05}$ = 1.86, $g_1 = -\frac{1.86}{3}S$, $g_2 = \frac{1.86}{3}S$.

7.27 By Definition 7.3, S_1^2 / S_2^2 has an F–distribution with 5 numerator and 9 denominator degrees of freedom. Then,
a. $P(S_1^2 / S_2^2 > 2) = .17271$.
b. $P(S_1^2 / S_2^2 < .5) = .23041$.
c. $P(S_1^2 / S_2^2 > 2) + P(S_1^2 / S_2^2 < .5)$ = .17271 + .23041 = .40312.

7.28 **a.** Using Table 7, $F_{.025}$ = 6.23.
b. The value $F_{.025}$ is the 97.5%-tile/quantile.
c. Using the applet, $F_{.975}$ = .10873.
d. Using the applet, $F_{.025}$ = 9.19731.
e. The relationship is 1/.10873 ≈ 9.19731.

7.29 By Definition 7.3, $Y = (W_1/\nu_1) \div (W_2/\nu_2)$ has an F distribution with ν_1 numerator and ν_2 denominator degrees of freedom. Therefore, $U = 1/Y = (W_2/\nu_2) \div (W_1/\nu_1)$ has an F distribution with ν_2 numerator and ν_1 denominator degrees of freedom.

7.30 **a.** $E(Z) = 0$, $E(Z^2) = V(Z) + [E(Z)]^2 = 1$.
b. This is very similar to Ex. 5.86, part a. Using that result, it is clear that
i. $E(T) = 0$
ii. $V(T) = E(T^2) = \nu E(Z^2/Y) = \nu/(\nu-2)$, $\nu > 2$.

7.31 **a.** The values for $F_{.01}$ are 5.99, 4.89, 4.02, 3.65, 3.48, and 3.32, respectively.
b. The values for $F_{.01}$ are decreasing as the denominator degrees of freedom increase.
c. From Table 6, $\chi^2_{.01} = 13.2767$.
d. 13.2767/3.32 ≈ 4. This follows from the fact that the F ratio as given in Definition 7.3 converges to W_1/ν_1 as ν_2 increases without bound.

7.32 **a.** Using the applet, $t_{.05}$ = 2.01505.
b. $P(T^2 > t^2_{.05}) = P(T > t_{.05}) + P(T < -t_{.05}) = .10$.
c. Using the applet, $F_{.10}$ = 4.06042.
d. $F_{.10} = 4.06042 = (2.01505)^2 = t^2_{.05}$.
e. Let $F = T^2$. Then, $.10 = P(F > F_{.10}) = P(T^2 > F_{.10}) = P(T < -\sqrt{F_{.10}}) + P(T > \sqrt{F_{.10}})$.
This must be equal to the expression given in part b.

7.33 Define $T = Z/\sqrt{W/\nu}$ as in Definition 7.2. Then, $T^2 = Z^2/(W/\nu)$. Since Z^2 has a chi–square distribution with 1 degree of freedom, and Z and W are independent, T^2 has an F distribution with 1 numerator and ν denominator degrees of freedom.

7.34 This exercise is very similar to Ex. 5.86, part b. Using that result, is can be shown that

a. $E(F) = \frac{\nu_2}{\nu_1} E(W_1)E(W_2^{-1}) = \frac{\nu_2}{\nu_1} \times \left(\frac{\nu_1}{\nu_2 - 2}\right) = \nu_2 / (\nu_2 - 2)$, $\nu_2 > 2$.

b. $V(F) = E(F^2) - [E(F)]^2 = \left(\frac{\nu_2}{\nu_1}\right)^2 E(W_1^2)E(W_2^{-2}) - \left(\frac{\nu_2}{\nu_2 - 2}\right)^2$

$= \left(\frac{\nu_2}{\nu_1}\right)^2 \nu_1(\nu_1 + 2)\frac{1}{(\nu_2 - 2)(\nu_2 - 4)} - \left(\frac{\nu_2}{\nu_2 - 2}\right)^2$

$= \left[2\nu_2^2(\nu_1 + \nu_2 - 2)\right] / \left[\nu_1(\nu_2 - 2)^2(\nu_2 - 4)\right]$, $\nu_2 > 4$.

7.35 Using the result from Ex. 7.34,

a. $E(F)$ = 70/(70–2) = 1.029.

b. $V(F) = [2(70)^2(118)]/[50(68)^2(66)] = .076$

c. Note that the value 3 is $(3 - 1.029)/\sqrt{.076}$ = 7.15 standard deviations above this mean. This represents and unlikely value.

7.36 We are given that $\sigma_1^2 = 2\sigma_2^2$. Thus, $\sigma_1^2/\sigma_2^2 = 2$ and $S_1^2/(2S_2^2)$ has an F distribution with 10 – 1 = 9 numerator and 8 – 1 = 7 denominator degrees of freedom.

a. We have $P(S_1^2/S_2^2 \le b) = P(S_1^2/(2S_2^2) \le b/2) = .95$. It must be that $b/2 = F_{.05} = 3.68$, so $b = 7.36$.

b. Similarly, $a/2 = F_{.95}$, but we must use the relation $a/2 = 1/F_{.05}$, where $F_{.05}$ is the 95th percentile of the F distribution with 7 numerator and 9 denominator degrees of freedom (see Ex. 7.29). Thus, with $F_{.05} = 3.29 = .304$, $a/2 = 2/3.29 = .608$.

c. $P(.608 \le S_1^2/S_2^2 \le 7.36) = .90$.

7.37 **a.** By Theorem 7.2, χ^2 with 5 degrees of freedom.

b. By Theorem 7.3, χ^2 with 4 degrees of freedom (recall that $\sigma^2 = 1$).

c. Since Y_6^2 is distributed as χ^2 with 1 degrees of freedom, and $\sum_{i=1}^{5}(Y_i - \overline{Y})^2$ and Y_6^2 are independent, the distribution of $W + U$ is χ^2 with 4 + 1 = 5 degrees of freedom.

7.38 **a.** By Definition 7.2, t–distribution with 5 degrees of freedom.

b. By Definition 7.2, t–distribution with 4 degrees of freedom.

c. $\overline{Y}$ follows a normal distribution with $\mu = 0$, $\sigma^2 = 1/5$. So, $\sqrt{5}\overline{Y}$ is standard normal and $\left(\sqrt{5}\overline{Y}\right)^2$ is chi–square with 1 degree of freedom. Therefore, $5\overline{Y}^2 + Y_6^2$ has a chi–square distribution with 2 degrees of freedom (the two random variables are independent). Now, the quotient

$$2(5\overline{Y}^2 + Y_6^2)/U = [(5\overline{Y}^2 + Y_6^2)/2] \div [U/4]$$

has an F-distribution with 2 numerator and 4 denominator degrees of freedom.

Note: we have assumed that $\overline{Y}$ and U are independent (as in Theorem 7.3).

7.39 **a.** Note that for $i = 1, 2, \ldots, k$, the $\overline{X}_i$ have independent a normal distributions with mean μ_i and variance σ/n_i. Since $\hat{\theta}$, a linear combination of independent normal random variables, by Theorem 6.3 $\hat{\theta}$ has a normal distribution with mean given by

$$E(\hat{\theta}) = E(c_1\overline{X}_1 + \ldots + c_k\overline{X}_k) = \sum_{i=1}^{k} c_i\mu_i$$

and variance

$$V(\hat{\theta}) = V(c_1\overline{X}_1 + \ldots + c_k\overline{X}_k) = \sigma^2\sum_{i=1}^{k} c_i^2 / n_i^2 .$$

b. For $i = 1, 2, \ldots, k$, $(n_i - 1)S_i^2 / \sigma^2$ follows a chi–square distribution with $n_i - 1$ degrees of freedom. In addition, since the S_i^2 are independent,

$$\frac{SSE}{\sigma^2} = \sum_{i=1}^{k}(n_i - 1)S_i^2 / \sigma^2$$

is a sum of independent chi–square variables. Thus, the above quantity is also distributed as chi–square with degrees of freedom $\sum_{i=1}^{k}(n_i - 1) = \sum_{i=1}^{k} n_i - k$.

c. From part a, we have that $\dfrac{\hat{\theta} - \theta}{\sigma\sqrt{\sum_{i=1}^{k} c_i^2 / n_i^2}}$

has a standard normal distribution. Therefore, by Definition 7.2, a random variable constructed as

$$\frac{\hat{\theta} - \theta}{\sigma\sqrt{\sum_{i=1}^{k} c_i^2 / n_i^2}} \Bigg/ \sqrt{\frac{\sum_{i=1}^{k}(n_i - 1)S_i^2 / \sigma^2}{\sum_{i=1}^{k} n_i - k}} = \frac{\hat{\theta} - \theta}{\sqrt{\text{MSE}\sum_{i=1}^{k} c_i^2 / n_i^2}}$$

has the t–distribution with $\sum_{i=1}^{k} n_i - k$ degrees of freedom. Here, we are assuming that $\hat{\theta}$ and SSE are independent (similar to $\overline{Y}$ and S^2 as in Theorem 7.3).

7.40 **a.** Both histograms are centered about the mean M = 16.50, but the variation is larger for sample means of size 1.
b. For sample means of size 1, the histogram closely resembles the population. For sample means of size 3, the histogram resembles the shape of the population but the variability is smaller.
c. Yes, the means are very close and the standard deviations are related by a scale of $\sqrt{3}$.
d. The normal densities approximate the histograms fairly well.
e. The normal density has the best approximation for the sample size of 25.

7.41 **a.** For sample means of size 1, the histogram closely resembles the population. For sample means of size 3, the histogram resembles that of a multi–modal population. The means and standard deviations follow the result of Ex. 7.40 (c), but the normal densities are not appropriate for either case. The normal density is better with $n = 10$, but it is best with $n = 25$.
b. For the "U–shaped population," the probability is greatest in the two extremes in the distribution.

7.42 Let $\overline{Y}$ denote the sample mean strength of 100 random selected pieces of glass. Thus, the quantity ($\overline{Y}$ – 14.5)/.2 has an approximate standard normal distribution.

a. $P(\overline{Y} > 14) \approx P(Z > 2.5) = .0062$.
b. We have that $P(–1.96 < Z < 1.96) = .95$. So, denoting the required interval as (a, b) such that $P(a < \overline{Y} < b) = .95$, we have that $–1.96 = (a – 14)/.2$ and $1.96 = (b – 14)/.2$. Thus, $a = 13.608$, $b = 14.392$.

7.43 Let $\overline{Y}$ denote the mean height and σ = 2.5 inches. By the Central Limit Theorem,
$P(|\overline{Y}-\mu| \le .5) = P(-.5 \le \overline{Y}-\mu \le .5) \approx P(\frac{-.5(10)}{2.5} \le Z \le \frac{-5(10)}{2.5}) = P(-2 \le Z \le 2) = .9544$.

7.44 Following Ex. 7.43, we now require

$$P(|\overline{Y}-\mu| \le .4) = P(-.4 \le \overline{Y}-\mu \le .4) \approx P(\tfrac{-.5\sqrt{n}}{2.5} \le Z \le \tfrac{5\sqrt{n}}{2.5}) = .95.$$

Thus, it must be true that $\frac{5\sqrt{n}}{2.5}$ = 1.96, or n = 150.0625. So, 151 men should be sampled.

7.45 Let $\overline{Y}$ denote the mean wage calculated from a sample of 64 workers. Then,

$$P(\overline{Y} \le 6.90) \approx P(Z \le \tfrac{\sqrt{8}(6.90-7.00)}{.5}) = P(Z \le -1.60) = .0548\,.$$

7.46 With n = 40 and σ ≈ (range)/4 = (8 – 5)/4 = .75, the approximation is

$$P(|\overline{Y}-\mu| \le .2) \approx P(|Z| \le \tfrac{\sqrt{40}(.2)}{.75}) = P(-1.69 \le Z \le 1.69) = .9090.$$

7.47 (Similar to Ex. 7.44). Following Ex. 7.47, we require

$$P(|\overline{Y}-\mu| \le .1) \approx P(|Z| \le \tfrac{\sqrt{n}(.1)}{.75}) = .90.$$

Thus, we have that $\frac{\sqrt{n}(.1)}{.75}$ = 1.645, so n = 152.21. Therefore, 153 core samples should be taken.

7.48 **a.** Although the population is not normally distributed, with n = 35 the sampling distribution of $\overline{Y}$ will be approximately normal. The probability of interest is

$$P(|\overline{Y}-\mu| \le 1) = P(-1 \le \overline{Y}-\mu \le 1)\,.$$

In order to evaluate this probability, the population standard deviation σ is needed. Since it is unknown, we will estimate its value by using the sample standard deviation s = 12 so that the estimated standard deviation of $\overline{Y}$ is $12/\sqrt{35}$ = 2.028. Thus,
$P(|\overline{Y}-\mu| \le 1) = P(-1 \le \overline{Y}-\mu \le 1) \approx P(-\frac{1}{2.028} \le Z \le \frac{1}{2.028}) = P(-.49 \le Z \le .49) = .3758$.

b. No, the measurements are still only *estimates*.

7.49 With μ = 1.4 hours, σ = .7 hour, let $\overline{Y}$ = mean service time for n = 50 cars. Then,

$$P(\overline{Y} > 1.6) \approx P(Z > \tfrac{\sqrt{50}(16-14)}{.7}) = P(Z > 2.02) = .0217.$$

7.50 We have $P(|\overline{Y}-\mu| < 1) = P(|Z| < \frac{1}{\sigma/\sqrt{n}}) = P(-1 < Z < 1) = .6826$.

7.51 We require $P(|\overline{Y}-\mu|<1) = P(|Z| < \frac{1}{\sigma/\sqrt{n}}) = P(-\frac{1}{10/\sqrt{n}} < Z < \frac{1}{10/\sqrt{n}})$ = .99. Thus it must be true that $\frac{1}{10/\sqrt{n}}$ = $z_{.005}$ = 2.576. So, n = 663.57, or 664 measurements should be taken.

7.52 Let $\overline{Y}$ denote the average resistance for the 25 resistors. With μ = 200 and σ = 10 ohms,
a. $P(199 \le \overline{Y} \le 202) \approx P(-.5 \le Z \le 1) = .5328.$
b. Let X = *total resistance* of the 25 resistors. Then,

$$P(X \le 5100) = P(\overline{Y} \le 204) \approx P(Z \le 2) = .9772.$$

7.53 **a.** With these given values for μ and σ, note that the value 0 has a z–score of (0 – 12)/9 = 1.33. This is not considered extreme, and yet this is the smallest possible value for CO concentration in air. So, a normal distribution is not possible for these measurements.
b. $\overline{Y}$ is approximately normal: $P(\overline{Y} > 14) \approx P(Z > \frac{\sqrt{100}(14-12)}{9}) = P(Z > 2.22)$ = .0132.

7.54 $P(\overline{Y} < 1.3) \approx P(Z < \frac{\sqrt{25}(1.3-1.4)}{.05}) = P(Z < -10) \approx 0$, so it is very unlikely.

7.55 **a.** i. We assume that we have a random sample
ii. Note that the standard deviation for the sample mean is $.8/\sqrt{30}$ = .146. The endpoints of the interval (1, 5) are substantially beyond 3 standard deviations from the mean. Thus, the probability is approximately 1.

b. Let Y_i denote the downtime for day i, i = 1, 2, …, 30. Then,

$$P(\sum_{i=1}^{30} Y_i < 115) = P(\overline{Y} < 3.833) \approx P(Z < \frac{\sqrt{30}(3.833-4)}{.8}) = P(Z < -1.14) = .1271.$$

7.56 Let Y_i denote the volume for sample i, i = 1, 2, …, 30. We require

$$P(\sum_{i=1}^{50} Y_i > 200) = P(\overline{Y} - \mu < \tfrac{200}{50} - \mu) \approx P(Z < \frac{\sqrt{50}(4-\mu)}{2}) = .95.$$

Thus, $\frac{\sqrt{50}(4-\mu)}{2} = -z_{.05}$ = –1.645, and then μ = 4.47.

7.57 Let Y_i denote the lifetime of the i^{th} lamp, i = 1, 2, …, 25, and the mean and standard deviation are given as 50 and 4, respectively. The random variable of interest is $\sum_{i=1}^{25} Y_i$, which is the lifetime of the lamp system. So,

$$P(\sum_{i=1}^{25} Y_i \ge 1300) = P(\overline{Y} \ge 52) \approx P(Z \ge \frac{\sqrt{25}(52-50)}{4}) = P(Z \ge 2.5) = .0062.$$

7.58 For $W_i = X_i - Y_i$, we have that $E(W_i) = E(X_i) - E(Y_i) = \mu_1 - \mu_2$ and $V(W_i) = V(X_i) - V(Y_i) = \sigma_1^2 + \sigma_2^2$ since X_i and Y_i are independent. Thus, $\overline{W} = \frac{1}{n}\sum_{i=1}^{n} W_i = \frac{1}{n}\sum_{i=1}^{n}(X_i - Y_i) = \overline{X} - \overline{Y}$ so $E(\overline{W})$ = μ₁ – μ₂, and $V(\overline{W}) = (\sigma_1^2 + \sigma_2^2)/n$. Thus, since the W_i are independent,

$$U_n = \frac{(\overline{X}-\overline{Y}) - (\mu_1 - \mu_2)}{\sqrt{(\sigma_1^2+\sigma_2^2)/n}} = \frac{\overline{W} - E(\overline{W})}{\sqrt{V(\overline{W})}}$$

satisfies the conditions of Theorem 7.4 and has a limiting standard normal distribution.

7.59 Using the result of Ex. 7.58, we have that $n = 50$, $\sigma_1 = \sigma_2 = 2$ and $\mu_1 = \mu_2$. Let $\overline{X}$ denote the mean time for operator *A* and let $\overline{Y}$ denote the mean time for operator *B* (both measured in seconds) Then, operator A will get the job if $\overline{X} - \overline{Y} < -1$. This probability is

$$P(\overline{X} - \overline{Y} < -1) \approx P\left(Z < \tfrac{-1}{\sqrt{(4+4)/50}}\right) = P(Z < -2.5) = .0062.$$

7.60 Extending the result from Ex. 7.58, let $\overline{X}$ denote the mean measurement for soil *A* and $\overline{Y}$ the mean measurement for soil *B*. Then, we require

$$P\left[\left|\overline{X} - \overline{Y} - (\mu_1 - \mu_2)\right| \le .05\right] \approx P\left[|Z| \le \tfrac{.05}{\sqrt{\frac{.01}{50} + \frac{.02}{100}}}\right] = P\left[|Z| \le 2.5\right] = .9876.$$

7.61 It is necessary to have

$$P\left[\left|\overline{X} - \overline{Y} - (\mu_1 - \mu_2)\right| \le .04\right] \approx P\left[|Z| \le \tfrac{.05}{\sqrt{\frac{.01}{n} + \frac{.02}{n}}}\right] = P\left[|Z| \le \tfrac{.05\sqrt{n}}{\sqrt{.01+.02}}\right] = .90.$$

Thus, $\frac{.05\sqrt{n}}{\sqrt{.01+.02}} = z_{.05} = 1.645$, so $n = 50.74$. Each sample size must be at least $n = 51$.

7.62 Let Y_i represent the time required to process the i^{th} person's order, $i = 1, 2, \ldots, 100$. We have that $\mu = 2.5$ minutes and $\sigma = 2$ minutes. So, since 4 hours = 240 minutes,

$$P(\textstyle\sum_{i=1}^{100} Y_i > 240) = P(\overline{Y} > 2.4) \approx P(Z > \frac{\sqrt{100}(2.4-2.5)}{2}) = P(Z > -.5) = .6915.$$

7.63 Following Ex. 7.62, consider the relationship $P(\sum_{i=1}^{n} Y_i < 120) = .1$ as a function of n:

Then, $P(\sum_{i=1}^{n} Y_i < 120) = P(\overline{Y} < 120/n) \approx P(Z < \frac{\sqrt{n}(120/n-2.5)}{2}) = .1$. So, we have that

$$\frac{\sqrt{n}(120/n-2.5)}{2} = -z_{.10} = -1.282.$$

Solving this nonlinear relationship (for example, this can be expressed as a quadratic relation in $\sqrt{n}$), we find that $n = 55.65$ so we should take a sample of 56 customers.

7.64 **a.** two.
b. exact: .27353, normal approximation: .27014
c. this is the continuity correction

7.65 **a.** exact: .91854, normal approximation: .86396.
b. the mass function does not resemble a mound–shaped distribution (*n* is not large here).

7.66 Since $P(|Y - E(Y)| \le 1) = P(E(Y) - 1 \le Y \le E(Y) + 1) = P(np - 1 \le Y \le np + 1)$, if $n = 20$ and $p = .1$, $P(1 \le Y \le 3) = .74547$. Normal Approximation: .73645.

7.67 **a.** $n = 5$ (exact: ..99968, approximate: .95319), $n = 10$ (exact: ..99363, approximate: .97312), $n = 15$ (exact: .98194, approximate: .97613), $n = 20$ (exact: .96786, approximate: .96886).
b. The binomial histograms appear more mound shaped with increasing values of *n*. The exact and approximate probabilities are closer for larger *n* values.
c. rule of thumb: $n > 9(.8/.2) = 36$, which is conservative since $n = 20$ is quite good.

7.68 **a.** The probability of interest is $P(Y \geq 29)$, where Y has a binomial distribution with n = 50 and p = .48. Exact: .10135, approximate: .10137.
b. The two probabilities are close. With n = 50 and p = .48, the binomial histogram is mound shaped.

7.69 **a.** Probably not, since current residents would have learned their lesson.
b. (Answers vary). With b = 32, we have exact: ..03268, approximate: .03289.

7.70 **a.** $p + 3\sqrt{pq/n} < 1 \Leftrightarrow 3\sqrt{pq/n} < q \Leftrightarrow 9pq/n < q^2 \Leftrightarrow 9p/q < n$.
b. $p - 3\sqrt{pq/n} < 1 \Leftrightarrow 3\sqrt{pq/n} < p \Leftrightarrow 9pq/n < p^2 \Leftrightarrow 9q/p < n$.
c. Parts **a** and **b** imply that $n > 9\max\left(\frac{p}{q}, \frac{q}{p}\right)$, and it is trivial to show that $\max\left(\frac{p}{q}, \frac{q}{p}\right) = \frac{\max(p,q)}{\min(p,q)}$ (consider the three cases where $p = q,\ p > q,\ p < q$.

7.71 **a.** $n > 9$.
b. $n > 14$, $n > 14$, $n > 36$, $n > 36$, $n > 891$, $n > 8991$.

7.72 Using the normal approximation, $P(Y \geq 15) \approx P(Z \geq \frac{14.5-10}{\sqrt{100(.1)(.9)}})$ = $P(Z \geq 1.5)$ = .0668.

7.73 Let Y = # the show up for a flight. Then, Y is binomial with n = 160 and p = .95. The probability of interest is $P(Y \leq 155)$, which gives the probability that the airline will be able to accommodate all passengers. Using the normal approximation, this is

$$P(Y \leq 155) \approx P(Z \leq \tfrac{155.5-160(.95)}{\sqrt{160(.95)(.05)}}) = P(Z \leq 1.27) = .8980.$$

7.74 **a.** Note that calculating the exact probability is easier: with n = 1500, p = 1/410,

$$P(Y \geq 1) = 1 - P(Y = 0) = 1 - (409/410)^{1500} = .9504.$$

b. Here, n = 1500, p = 1/64. So,

$$P(Y > 30) \approx P(Z > \tfrac{30.5-23.4375}{\sqrt{23.0713}}) = P(Z > 1.47) = .0708.$$

c. The value y = 30 is $(30 - 23.4375)/\sqrt{23.0713}$ = 1.37 standard deviations above the mean. This does not represent an unlikely value.

7.75 Let Y = # the favor the bond issue. Then, the probability of interest is

$$P\left(\left|\tfrac{Y}{n} - p\right| \leq .06\right) = P\left(-.06 \leq \tfrac{Y}{n} - p \leq .06\right) \approx P\left(\frac{-.06}{\sqrt{\frac{.2(.8)}{64}}} \leq Z \leq \frac{.06}{\sqrt{\frac{.2(.8)}{64}}}\right) = P(-1.2 \leq Z \leq 1.2) = .7698.$$

7.76 **a.** We know that $V(Y/n) = p(1-p)/n$. Consider n fixed and let $g(p) = p(1-p)/n$. This function is maximized at p = 1/2 (verify using standard calculus techniques).

b. It is necessary to have $P\left(\left|\frac{Y}{n} - p\right| \leq .1\right) = .95$, or approximately $P\left(|Z| \leq \frac{.1}{\sqrt{pq/n}}\right) = .95$. Thus, it must be true that $\frac{.1}{\sqrt{pq/n}}$ = 1.96. Since p is unknown, replace it with the value 1/2 found in part a (this represents the "worse case scenario") and solve for n. In so doing, it is found that n = 96.04, so that 97 items should be sampled.

7.77 (Similar to Ex. 7.76). Here, we must solve $\frac{.15}{\sqrt{pq/n}} = z_{.01} = 2.33$. Using $p = 1/2$, we find that $n = 60.32$, so 61 customers should be sampled.

7.78 Following Ex. 7.77: if $p = .9$, then

$$P\left(\left|\tfrac{Y}{n} - p\right| \le .15\right) \approx P\left(|Z| \le \tfrac{.15}{\sqrt{.9(.1)/50}}\right) = P(|Z| \le 3.54) \approx 1.$$

7.79 **a.** Using the normal approximation:
$P(Y \ge 2) = P(Y \ge 1.5) = P(Z \ge \frac{1.5-2.5}{\sqrt{25(.1)(.9)}}) = P(Z \ge -.67) = .7486.$

b. Using the exact binomial probability:
$P(Y \ge 2) = 1 - P(Y \le 1) = 1 - .271 = .729$.

7.80 Let Y = # in the sample that are younger than 31 years of age. Since 31 is the median age, Y will have a binomial distribution with $n = 100$ and $p = 1/2$ (here, we are being rather lax about the specific age of 31 in the population). Then,

$$P(Y \ge 60) = P(Y \ge 59.5) \approx P(Z \ge \tfrac{59.5-50}{\sqrt{100(.5)(.5)}}) = P(Z \ge 1.9) = .0287.$$

7.81 Let Y = # of non–conforming items in our lot. Thus, with $n = 50$:

a. With $p = .1$, P(lot is accepted) = $P(Y \le 5) = P(Y \le 5.5) = P(Z \le \frac{5.5-50(.1)}{\sqrt{50(.1)(.9)}}) =$
$P(Z \le .24) = .5948.$

b. With $p = .2$ and .3, the probabilities are .0559 and .0017 respectively.

7.82 Let Y = # of disks with missing pulses. Then, Y is binomial with $n = 100$ and $p = .2$. Thus, $P(Y \ge 15) = P(Y \ge 14.5) \approx P(Z \ge \frac{14.5-100(.2)}{\sqrt{100(.2)(.8)}}) = P(Z \ge -1.38) = .9162.$

7.83 **a.** Let Y = # that turn right. Then, Y is binomial with $n = 50$ and $p = 1/3$. Using the applet, $P(Y \le 15) = .36897$.
b. Let Y = # that turn (left or right). Then, Y is binomial with $n = 50$ and $p = 2/3$. Using the applet, $P(Y \ge (2/3)50) = P(Y \ge 33.333) = P(Y \ge 34) = .48679$.

7.84 **a.** $E\left(\frac{Y_1}{n_1} - \frac{Y_2}{n_2}\right) = \frac{E(Y_1)}{n_1} - \frac{E(Y_2)}{n_2} = \frac{n_1 p_1}{n_1} - \frac{n_2 p_2}{n_2} = p_1 - p_2$.
b. $V\left(\frac{Y_1}{n_1} - \frac{Y_2}{n_2}\right) = \frac{V(Y_1)}{n_1^2} + \frac{V(Y_2)}{n_2^2} = \frac{n_1 p_1 q_1}{n_1^2} + \frac{n_2 p_2 q_2}{n_2^2} = \frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}$.

7.85 It is given that $p_1 = .1$ and $p_2 = .2$. Using the result of Ex. 7.58, we obtain

$$P\left(\left|\tfrac{Y_1}{n_1} - \tfrac{Y_2}{n_2} - (p_1 - p_2)\right| \le .1\right) \approx P\left(|Z| \le \frac{.1}{\sqrt{\frac{.1(.9)}{50} + \frac{.2(.8)}{50}}}\right) = P(|Z| \le 1.4) = .8414.$$

7.86 Let Y = # of travel vouchers that are improperly documented. Then, Y has a binomial distribution with $n = 100$, $p = .20$. Then, the probability of observing more than 30 is

$$P(Y > 30) = P(Y > 30.5) \approx P(Z > \tfrac{30.5-100(.2)}{\sqrt{100(.2)(.8)}}) = P(Z > 2.63) = .0043.$$

We conclude that the claim is probably incorrect since this probability is very small.

7.87 Let X = waiting time over a 2–day period. Then, X is exponential with β = 10 minutes. Let Y = # of customers whose waiting times is greater than 10 minutes. Then, Y is binomial with n = 100 and p is given by

$$p = \int_{10}^{\infty} \tfrac{1}{10} e^{-y/10} dy = e^{-1} = .3679.$$

Thus, $P(Y \geq 50) \approx P(Z \geq \frac{50-100(.3697)}{\sqrt{1000(.3679)(.6321)}} = P(Z \geq 2.636) = .0041.$

7.88 Since the efficiency measurements follow a normal distribution with mean μ = 9.5 lumens and σ = .5 lumens, then

$$\overline{Y} = \text{mean efficiency of eight bulbs}$$

follows a normal distribution with mean 9.5 lumens and standard deviation $.5/\sqrt{8}$.
Thus, $P(\overline{Y} > 10) = P(Z > \frac{10-9.5}{.5/\sqrt{8}}) = P(Z > 2.83) = .0023.$

7.89 Following Ex. 7.88, it is necessary that $P(\overline{Y} > 10) = P(Z > \frac{10-\mu}{.5/\sqrt{8}}) = .80$, where μ denotes the mean efficiency. Thus, $\frac{10-\mu}{.5/\sqrt{8}} = z_{.2} = -.84$ so μ = 10.15.

7.90 Denote Y = # of successful transplants. Then, Y has a binomial distribution with n = 100 and p = .65. Then, using the normal approximation to the binomial,

$$P(Y > 70) \approx P(Z > \frac{70-100(.65)}{\sqrt{100(.65)(.35)}}) = P(Z > 1.15) = .1251.$$

7.91 Since X, Y, and W are normally distributed, so are $\overline{X}$, $\overline{Y}$, and $\overline{W}$. In addition, by Theorem 6.3 U follows a normal distribution such that

$$\mu_U = E(U) = .4\mu_1 + .2\mu_2 + .4\mu_3$$
$$\sigma_U^2 = V(U) = .16\left(\tfrac{\sigma_1^2}{n_1}\right) + .04\left(\tfrac{\sigma_2^2}{n_2}\right) + .16\left(\tfrac{\sigma_3^2}{n_3}\right).$$

7.92 The desired probability is

$$P\left[\left|\overline{X} - \overline{Y}\right| > .6\right] = P\left[|Z| \leq \frac{.06}{\sqrt{[(6.4)^2 + (6.4)^2]/64}}\right] = P\left[|Z| \leq .50\right] = .6170.$$

7.93 Using the mgf approach, the mgf for the exponential distribution with mean θ is

$$m_Y(t) = (1-\theta t)^{-1},\ t < 1/\theta.$$

The mgf for $U = 2Y/\theta$ is

$$m_U(t) = E(e^{tU}) = E(e^{(t2/\theta)Y}) = m_Y(2t/\theta) = (1-2t)^{-1},\ t < 1/2.$$

This is the mgf for the chi–square distribution with 2 degrees of freedom.

7.94 Using the result from Ex. 7.93, the quantity $2Y_i/20$ is chi–square with 2 degrees of freedom. Further, since the Y_i are independent, $U = \sum_{i=1}^{5} 2Y_i/20$ is chi–square with 10 degrees of freedom. Thus, $P\left(\sum_{i=1}^{5} Y_i > c\right) = P(U > \frac{c}{10}) = .05$. So, it must be true that $\frac{c}{10} = \chi^2_{.05}$ = 18.307, or c = 183.07.

7.95 **a.** Since μ = 0 and by Definition 2, $T = \dfrac{\overline{Y}}{S/\sqrt{10}}$ has a *t*–distribution with 9 degrees of freedom. Also, $T^2 = \dfrac{\overline{Y}^2}{S^2/10} = \dfrac{10\overline{Y}^2}{S^2}$ has an *F*–distribution with 1 numerator and 9 denominator degrees of freedom (see Ex. 7.33).

b. By Definition 3, $T^{-2} = \dfrac{S^2}{10\overline{Y}^2}$ has an *F*–distribution with 9 numerator and 1 denominator degrees of freedom (see Ex. 7.29).

c. With 9 numerator and 1 denominator degrees of freedom, $F_{.05}$ = 240.5. Thus,

$$.95 = P\left(\frac{S^2}{10\overline{Y}^2} < 240.5\right) = P\left(\frac{S^2}{\overline{Y}^2} < 2405\right) = P\left(-49.04 < \frac{S}{\overline{Y}} < 49.04\right),$$

so c = 49.04.

7.96 Note that Y has a beta distribution with α = 3 and β = 1. So, μ = 3/4 and σ^2 = 3/80. By the Central Limit Theorem, $P(\overline{Y} > .7) \approx P(Z > \frac{.7-.75}{\sqrt{.0375/40}}) = P(Z > -1.63) = .9484$.

7.97 **a.** Since the X_i are independent and identically distributed chi–square random variables with 1 degree of freedom, if $Y = \sum_{i=1}^{n} X_i$, then $E(Y) = n$ and $V(Y) = 2n$. Thus, the conditions of the Central Limit Theorem are satisfied and

$$Z = \frac{Y-n}{\sqrt{2n}} = \frac{\overline{X}-1}{\sqrt{2/n}}.$$

b. Since each Y_i is normal with mean 6 and variance .2, we have that

$$U = \sum_{i=1}^{50} \frac{(Y_i - 6)^2}{.2}$$

is chi–square with 50 degrees of freedom. For i = 1, 2, …, 50, let C_i be the cost for a single rod, Then, $C_i = 4(Y_i - 6)^2$ and the total cost is $T = \sum_{i=1}^{50} C_i = .8U$. By Ex. 7.97,

$$P(T > 48) = P(.8U > 48) = P(U > 60) \approx P\left(Z > \frac{60-50}{\sqrt{100}}\right) = P(Z > 1) = .1587.$$

7.98 **a.** Note that since Z has a standard normal distribution, the random variable Z/c also has a normal distribution with mean 0 and variance $1/c^2 = \nu/w$. Thus, we can write the conditional density of T given $W = w$ as

$$f(t \mid w) = \tfrac{1}{\sqrt{2\pi}}\sqrt{\tfrac{w}{\nu}}\, e^{-wt^2/(2\nu)}, \quad -\infty < t < \infty.$$

b. Since W has a chi–square distribution with ν degrees of freedom,

$$f(t,w) = f(t \mid w) f(w) = \tfrac{1}{\sqrt{2\pi}}\sqrt{\tfrac{w}{\nu}}\, e^{-wt^2/(2\nu)} \left(\tfrac{1}{\Gamma(\nu/2)2^{\nu/2}} w^{\nu/2-1} e^{-w/2}\right).$$

c. Integrating over w, we obtain

$$f(t)=\int_0^\infty \tfrac{1}{\sqrt{2\pi}}\sqrt{\tfrac{w}{\nu}}e^{-wt^2/(2\nu)}\left(\tfrac{1}{\Gamma(\nu/2)2^{\nu/2}}w^{\nu/2-1}e^{-w/2}\right)dw=\int_0^\infty \tfrac{1}{\sqrt{\pi\nu}}\tfrac{1}{\Gamma(\nu/2)2^{(\nu+1)/2}}\exp\left[-\tfrac{w}{2}\left(1+\tfrac{t^2}{\nu}\right)\right]w^{[(\nu+1)/2]-1}dw.$$

Writing another way this is,

$$f(t)=\tfrac{\left(1+t^2/\nu\right)^{-(\nu+1)/2}}{\sqrt{\pi\nu}}\tfrac{\Gamma[(\nu+1)/2]}{\Gamma(\nu/2)}\int_0^\infty \tfrac{1}{\Gamma[(\nu+1)/2]}\tfrac{1}{2^{(\nu+1)/2}}\tfrac{1}{\left(1+t^2/\nu\right)^{-(\nu+1)/2}}\exp\left[-\tfrac{w}{2}\left(1+\tfrac{t^2}{\nu}\right)\right]w^{[(\nu+1)/2]-1}dw.$$

The integrand is that of a gamma density with shape parameter $(\nu+1)/2$ and scale parameter $2/\left[1+t^2/\nu\right]$, so it must integrate to one. Thus, the given form for $f(t)$ is correct.

7.99 **a.** Similar to Ex. 7.98. For fixed $W_2 = w_2$, $F = W_1/c$, where $c = w_2\nu_1/\nu_2$. To find this conditional density of F, note that the mgf for W_1 is

$$m_{W_1}(t)=(1-2t)^{-\nu_1/2}.$$

The mgf for $F = W_1/c$ is

$$m_F(t)=m_{W_1}(t/c)=(1-2t/c)^{-\nu_1/2}.$$

Since this mgf is in the form of a gamma mgf, the conditional density of F, conditioned that $W_2 = w_2$, is gamma with shape parameter ν_1 and scale parameter $2\nu_2/(w_2\nu_1)$.

b. Since W_2 has a chi–square distribution with ν_2 degrees of freedom, the joint density is

$$g(f,w_2)=g(f\mid w_2)f(w_2)=\frac{f^{(\nu_1/2)-1}e^{-fw_2\nu_1/(2\nu_2)}w_2^{(\nu_2/2)-1}e^{-w_2/2}}{\Gamma\left(\frac{\nu_1}{2}\right)\left(\frac{2\nu_2}{w_2\nu_1}\right)^{\nu_1/2}\Gamma\left(\frac{\nu_2}{2}\right)2^{\nu_2/2}}$$

$$=\frac{f^{(\nu_1/2)-1}w_2^{[(\nu_1+\nu_2)/2]-1}e^{-(w_2/2)[f\nu_1/\nu_2+1]}}{\Gamma\left(\frac{\nu_1}{2}\right)\left(\frac{\nu_2}{\nu_1}\right)^{\nu_1/2}\Gamma\left(\frac{\nu_2}{2}\right)2^{(\nu_1+\nu_2)/2}}.$$

c. Integrating over w_2, we obtain,

$$g(f)=\frac{f^{(\nu_1/2)-1}}{\Gamma\left(\frac{\nu_1}{2}\right)\left(\frac{\nu_2}{\nu_1}\right)^{\nu_1/2}\Gamma\left(\frac{\nu_2}{2}\right)2^{(\nu_1+\nu_2)/2}}\int_0^\infty w_2^{[(\nu_1+\nu_2)/2]-1}e^{-(w_2/2)[f\nu_1/\nu_2+1]}dw_2 .$$

The integrand can be related to a gamma density with shape parameter $(\nu_1 + \nu_2)/2$ and scale parameter $\frac{1}{2}\left(1+f\nu_1/\nu_2\right)^{-1}$ in order to evaluate the integral. Thus:

$$g(f)=\frac{\Gamma\left(\frac{\nu_1+\nu_2}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)}\frac{f^{(\nu_1/2)-1}\left(1+f\nu_1/\nu_2\right)^{-(\nu_1+\nu_2)/2}}{\left(\frac{\nu_2}{\nu_1}\right)^{\nu_1/2}2^{(\nu_1+\nu_2)/2}},\ f\ge 0.$$

7.100 The mgf for X is $m_X(t)=\exp\left(\lambda e^t-1\right)$.

a. The mgf for $Y=\left(X-\lambda\right)/\sqrt{\lambda}$ is given by

$$m_Y(t)=E(e^{tY})=e^{-t\sqrt{\lambda}}m_X(t/\sqrt{\lambda})=\exp\left(\lambda e^{t/\sqrt{\lambda}}-t\sqrt{\lambda}-\lambda\right).$$

b. Using the expansion as given, we have

$$m_Y(t)=\exp\left[-t\sqrt{\lambda}+\lambda\left(\tfrac{t}{\sqrt{\lambda}}+\tfrac{t^2}{2\lambda}+\tfrac{t^3}{6\lambda^{3/2}}+\cdots\right)\right]=\exp\left(\tfrac{t^2}{2}+\tfrac{t^3}{6\lambda^{1/2}}+\cdots\right).$$

As $\lambda \to \infty$, all terms after the first in the series will go to zero so that the limiting form of the mgf is $m_Y(t) = \exp\left(\frac{t^2}{2}\right)$

c. Since the limiting mgf is the mgf of the standard normal distribution, by Theorem 7.5 the result is proven.

7.101 Using the result in Ex. 7.100,

$$P(X \le 110) \approx P(Z \le \tfrac{110-100}{\sqrt{100}}) = P(Z \le 1) = .8413.$$

7.102 Again use the result in Ex. 7.101,

$$P(Y \ge 45) \approx P(Z \ge \tfrac{45-36}{\sqrt{36}}) = P(Z \ge 1.5) = .0668.$$

7.103 Following the result in Ex. 7.101, and that X and Y are independent, the quantity

$$\frac{X - Y - (\lambda_1 - \lambda_2)}{\sqrt{\lambda_1 + \lambda_2}}$$

has a limiting standard normal distribution (see Ex. 7.58 as applied to the Poisson). Therefore, the approximation is

$$P(X - Y > 10) \approx P(Z > 1) = .1587.$$

7.104 The mgf for Y_n is given by

$$m_{Y_n}(t) = \left[1 - p + pe^t\right]^n.$$

Let $p = \lambda/n$ and this becomes

$$m_{Y_n}(t) = \left[1 - \tfrac{\lambda}{n} + \tfrac{\lambda}{n}e^t\right]^n = \left[1 + \tfrac{1}{n}(\lambda e^t - 1)\right]^n.$$

As $n \to \infty$, this is $\exp(\lambda e^t - 1)$, the mgf for the Poisson with mean λ.

7.105 Let Y = # of people that suffer an adverse reaction. Then, Y is binomial with n = 1000 and p = .001. Using the result in Ex. 7.104, we let λ = 1000(.001) = 1 and evaluate

$$P(Y \ge 2) = 1 - P(Y \le 1) \approx 1 - .736 = .264,$$

using the Poisson table in Appendix 3.

Chapter 8: Estimation

8.1 Let $B = B(\hat{\theta})$. Then,

$$MSE(\hat{\theta}) = E\left[(\hat{\theta}-\theta)^2\right] = E\left[(\hat{\theta}-E(\hat{\theta})+B)^2\right] = E\left[\left(\hat{\theta}-E(\hat{\theta})\right)^2\right] + E(B^2) + 2B\times E\left[\hat{\theta}-E(\hat{\theta})\right]$$
$$= V(\hat{\theta}) + B^2.$$

8.2 **a.** The estimator $\hat{\theta}$ is unbiased if $E(\hat{\theta}) = \theta$. Thus, $B(\hat{\theta}) = 0$.
b. $E(\hat{\theta}) = \theta + 5$.

8.3 **a.** Using Definition 8.3, $B(\hat{\theta}) = a\theta + b - \theta = (a-1)\theta + b$.
b. Let $\hat{\theta}^* = (\hat{\theta} - b)/a$.

8.4 **a.** They are equal.
b. $MSE(\hat{\theta}) > V(\hat{\theta})$.

8.5 **a.** Note that $E(\hat{\theta}^*) = \theta$ and $V(\hat{\theta}^*) = V[(\hat{\theta}-b)/a] = V(\hat{\theta})/a^2$. Then,

$$\text{MSE}(\hat{\theta}^*) = V(\hat{\theta}^*) = V(\hat{\theta})/a^2.$$

b. Note that $\text{MSE}(\hat{\theta}) = V(\hat{\theta}) + B(\hat{\theta}) = V(\hat{\theta}) + [(a-1)\theta + b]^2$. A sufficiently <u>large</u> value of a will force $\text{MSE}(\hat{\theta}^*) < \text{MSE}(\hat{\theta})$. Example: $a = 10$.

c. A amply <u>small</u> value of a will make $\text{MSE}(\hat{\theta}^*) > \text{MSE}(\hat{\theta})$. Example: $a = .5$, $b = 0$.

8.6 **a.** $E(\hat{\theta}_3) = aE(\hat{\theta}_1) + (1-a)E(\hat{\theta}_2) = a\theta + (1-a)\theta = \theta$.

b. $V(\hat{\theta}_3) = a^2V(\hat{\theta}_1) + (1-a)^2V(\hat{\theta}_2) = a^2\sigma_1^2 + (1-a)\sigma_2^2$, since it was assumed that $\hat{\theta}_1$ and $\hat{\theta}_2$ are independent. To minimize $V(\hat{\theta}_3)$, we can take the first derivative (with respect to a), set it equal to zero, to find

$$a = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}.$$

(One should verify that the second derivative test shows that this is indeed a minimum.)

8.7 Following Ex. 8.6 but with the condition that $\hat{\theta}_1$ and $\hat{\theta}_2$ are not independent, we find

$$V(\hat{\theta}_3) = a^2\sigma_1^2 + (1-a)\sigma_2^2 + 2a(1-a)c.$$

Using the same method w/ derivatives, the minimum is found to be

$$a = \frac{\sigma_2^2 - c}{\sigma_1^2 + \sigma_2^2 - 2c}.$$

8.8 **a.** Note that $\hat{\theta}_1$, $\hat{\theta}_2$, $\hat{\theta}_3$ and $\hat{\theta}_5$ are simple linear combinations of Y_1, Y_2, and Y_3. So, it is easily shown that all four of these estimators are *unbiased*. From Ex. 6.81 it was shown that $\hat{\theta}_4$ has an exponential distribution with mean θ/3, so this estimator is biased.

b. It is easily shown that $V(\hat{\theta}_1) = \theta^2$, $V(\hat{\theta}_2) = \theta^2/2$, $V(\hat{\theta}_3) = 5\theta^2/9$, and $V(\hat{\theta}_5) = \theta^2/9$, so the estimator $\hat{\theta}_5$ is unbiased and has the smallest variance.

8.9 The density is in the form of the exponential with mean θ + 1. We know that $\overline{Y}$ is unbiased for the mean θ + 1, so an unbiased estimator for θ is simply $\overline{Y} - 1$.

8.10 **a.** For the Poisson distribution, $E(Y) = \lambda$ and so for the random sample, $E(\overline{Y}) = \lambda$. Thus, the estimator $\hat{\lambda} = \overline{Y}$ is unbiased.

b. The result follows from $E(Y) = \lambda$ and $E(Y^2) = V(Y) + \lambda^2 = 2\lambda^2$, so $E(C) = 4\lambda + \lambda^2$.

c. Since $E(\overline{Y}) = \lambda$ and $E(\overline{Y}^2) = V(\overline{Y}) + [E(\overline{Y})]^2 = \lambda^2/n + \lambda^2 = \lambda^2(1 + 1/n)$. Then, we can construct an unbiased estimator $\hat{\theta} = \overline{Y}^2 + \overline{Y}(4 - 1/n)$.

8.11 The third central moment is defined as

$$E[(Y-\mu)^3] = E[(Y-3)^3] = E(Y^3) - 9E(Y^2) + 54\,.$$

Using the unbiased estimates $\hat{\theta}_2$ and $\hat{\theta}_3$, it can easily be shown that $\hat{\theta}_3 - 9\hat{\theta}_2 + 54$ is an unbiased estimator.

8.12 **a.** For the uniform distribution given here, $E(Y_i) = \theta + .5$. Hence, $E(\overline{Y}) = \theta + .5$ so that $B(\overline{Y}) = .5$.

b. Based on $\overline{Y}$, the unbiased estimator is $\overline{Y} - .5$.

c. Note that $V(\overline{Y}) = 1/(12n)$ so $\text{MSE}(\overline{Y}) = 1/(12n) + .25$.

8.13 **a.** For a random variable Y with the binomial distribution, $E(Y) = np$ and $V(Y) = npq$, so $E(Y^2) = npq + (np)^2$. Thus,

$$E\left\{n\left(\tfrac{Y}{n}\right)\left[1-\tfrac{Y}{n}\right]\right\} = E(Y) - \tfrac{1}{n}E(Y^2) = np - pq - np^2 = (n-1)pq\,.$$

b. The unbiased estimator should have expected value npq, so consider the estimator

$$\hat{\theta} = \left(\tfrac{n}{n-1}\right)n\left(\tfrac{Y}{n}\right)\left[1-\tfrac{Y}{n}\right].$$

8.14 Using standard techniques, it can be shown that $E(Y) = \left(\frac{\alpha}{\alpha+1}\right)\theta$, $E(Y^2) = \left(\frac{\alpha}{\alpha+2}\right)\theta^2$. Also, it is easily shown that $Y_{(n)}$ follows the power family with parameters $n\alpha$ and θ.

a. From the above, $E(\hat{\theta}) = E(Y_{(n)}) = \left(\frac{n\alpha}{n\alpha+1}\right)\theta$, so that the estimator is biased.

b. Since α is known, the unbiased estimator is $\left(\frac{n\alpha+1}{n\alpha}\right)\hat{\theta} = \left(\frac{n\alpha+1}{n\alpha}\right)Y_{(n)}$.

c. $MSE(Y_{(n)}) = E[(Y_{(n)} - \theta)^2] = E(Y_{(n)}^2) - 2\theta E(Y_{(n)}) + \theta^2 = \frac{2}{(n\alpha+1)(n\alpha+2)}\theta^2$.

8.15 Using standard techniques, it can be shown that $E(Y) = (3/2)\beta$, $E(Y^2) = 3\beta^2$. Also it is easiliy shown that $Y_{(1)}$ follows the Pareto family with density function

$$g_{(1)}(y) = 3n\beta^{3n}y^{-(3n+1)},\ y \geq \beta.$$

Thus, $E(Y_{(1)}) = \left(\frac{3n}{3n-1}\right)\beta$ and $E(Y_{(1)}^2) = \frac{3n}{3n-2}\beta^2$.

a. With $\hat{\beta} = Y_{(1)}$, $B(\hat{\beta}) = \left(\frac{3n}{3n-1}\right)\beta - \beta = \left(\frac{1}{3n-1}\right)\beta$.

b. Using the above, $MSE(\hat{\beta}) = MSE(Y_{(1)}) = E(Y_{(1)}^2) - 2\beta E(Y_{(1)}) + \beta^2 = \frac{2}{(3n-1)(3n-2)}\beta^2$.

8.16 It is known that $(n-1)S^2/\sigma^2$ is chi–square with n–1 degrees of freedom.

a. $E(S) = E\left\{\frac{\sigma}{\sqrt{n-1}}\left[\frac{(n-1)S^2}{\sigma^2}\right]^{1/2}\right\} = \frac{\sigma}{\sqrt{n-1}}\int_0^\infty v^{1/2}\frac{1}{\Gamma[(n-1)/2]2^{(n-1)/2}}v^{(n-1)/2}e^{-v/2}dv = \frac{\sigma}{\sqrt{n-1}}\frac{\sqrt{2}\Gamma(n/2)}{\Gamma[(n-1)/2]}$.

b. The estimator $\hat{\sigma} = \frac{\sqrt{n-1}\Gamma[(n-1)/2]}{\sqrt{2}\Gamma(n/2)}S$ is unbiased for σ.

c. Since $E(\overline{Y}) = \mu$, the unbiased estimator of the quantity is $\overline{Y} - z_\alpha\hat{\sigma}$.

8.17 It is given that $\hat{p}_1$ is unbiased, and since $E(Y) = np$, $E(\hat{p}_2) = (np + 1)/(n+2)$.

a. $B(\hat{p}_2) = (np + 1)/(n+2) - p = (1-2p)/(n+2)$.

b. Since $\hat{p}_1$ is unbiased, MSE($\hat{p}_1$) = V($\hat{p}_1$) = $p(1-p)/n$. MSE($\hat{p}_2$) = $V(\hat{p}_2) + B(\hat{p}_2)$ = $\frac{np(1-p)+(1-2p)^2}{(n+2)^2}$.

c. Considering the inequality

$$\frac{np(1-p)+(1-2p)^2}{(n+2)^2} < \frac{p(1-p)}{n},$$

this can be written as

$$(8n+4)p^2 - (8n+4)p + n < 0.$$

Solving for p using the quadratic formula, we have

$$p = \frac{8n+4 \pm \sqrt{(8n+4)^2 - 4(8n+4)n}}{2(8n+4)} = \frac{1}{2} \pm \sqrt{\frac{n+1}{8n+4}}.$$

So, p will be close to .5.

8.18 Using standard techniques from Chapter 6, is can be shown that the density function for $Y_{(1)}$ is given by

$$g_{(1)}(y) = \frac{n}{\theta}\left(1 - \frac{y}{\theta}\right)^{n-1},\ 0 \leq y \leq \theta.$$

So, $E(Y_{(1)}) = \frac{\theta}{n+1}$ and so an unbiased estimator for θ is $(n+1)Y_{(1)}$.

8.19 From the hint, we know that $E(Y_{(1)}) = \beta/n$ so that $\hat{\theta} = nY_{(1)}$ is unbiased for β. Then, $\text{MSE}(\hat{\theta}) = V(\hat{\theta}) + B(\hat{\theta}) = V(nY_{(1)}) = n^2V(Y_{(1)}) = \beta^2$.

8.20 If Y has an exponential distribution with mean θ, then by Ex. 4.11, $E(\sqrt{Y}) = \sqrt{\pi\theta}/2$.

a. Since Y_1 and Y_2 are independent, $E(X) = \pi\theta/4$ so that $(4/\pi)X$ is unbiased for θ.

b. Following part a, it is easily seen that $E(W) = \pi^2\theta^2/16$, so $(4^2/\pi^2)W$ is unbiased for θ^2.

8.21 Using Table 8.1, we can estimate the population mean by $\bar{y} = 11.5$ and use a two–standard–error bound of $2(3.5)/\sqrt{50} = .99$. Thus, we have $11.5 \pm .99$.

8.22 (Similar to Ex. 8.21) The point estimate is $\bar{y} = 7.2\%$ and a bound on the error of estimation is $2(5.6)/\sqrt{200} = .79\%$.

8.23 **a.** The point estimate is $\bar{y} = 11.3$ ppm and an error bound is $2(16.6)/\sqrt{467} = 1.54$ ppm.

b. The point estimate is 46.4 – 45.1 = 1.3 and an error bound is $2\sqrt{\frac{(9.8)^2}{191} + \frac{(10.2)^2}{467}} = 1.7$.

c. The point estimate is .78 – .61 = .17 and an error bound is $2\sqrt{\frac{(.78)(.22)}{467} + \frac{(.61)(.39)}{191}} = .08$.

8.24 Note that by using a two–standard–error bound, $2\sqrt{\frac{(.69)(.31)}{1001}} = .0292 \approx .03$. Constructing this as an interval, this is (.66, .72). We can say that there is little doubt that the true (population) proportion falls in this interval. Note that the value 50% is far from the interval, so it is clear that a majority did feel that the cost of gasoline was a problem.

8.25 We estimate the difference to be 2.4 – 3.1 = –.7 with an error bound of $2\sqrt{\frac{1.44+2.64}{100}} = .404$.

8.26 **a.** The estimate of the true population proportion who think humans should be sent to Mars is .49 with an error bound of $2\sqrt{\frac{.49(.51)}{1093}} = .03$.

b. The standard error is given by $\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$, and this is maximized when $\hat{p} = .5$. So, a conservative error bound that could be used for all sample proportions (with $n = 1093$) is $2\sqrt{\frac{.5(.5)}{1093}} = .0302$ (or 3% as in the above).

8.27 **a.** The estimate of p is the sample proportion: 592/985 = .601, and an error bound is given by $2\sqrt{\frac{.601(.399)}{985}} = .031$.

b. The above can be expressed as the interval (.570, .632). Since this represents a clear majority for the candidate, it appears certain that the republican will be elected. Following Example 8.2, we can be reasonably confident by this statement.

c. The group of "likely voters" is not necessarily the same as "definite voters."

8.28 The point estimate is given by the difference of the sample proportions: .70 – .54 = .16 and an error bound is $2\sqrt{\frac{.7(.3)}{180}+\frac{.54(.46)}{100}}$ = .121.

8.29 **a.** The point estimate is the difference of the sample proportions: .45 – .51 = –.06, and an error bound is $2\sqrt{\frac{.45(.55)}{1001}+\frac{.51(.49)}{1001}}$ = .045.

b. The above can be expressed as the interval (–.06 – .045, –.06 + .045) or (–.105, –.015). Since the value 0 is not contained in the interval, it seems reasonable to claim that fan support for baseball is greater at the end of the season.

8.30 The point estimate is .45 and an error bound is $2\sqrt{\frac{.45(.55)}{1001}}$ = .031. Since 10% is roughly three times the two–standard–error bound, it is not likely (assuming the sample was indeed a randomly selected sample).

8.31 **a.** The point estimate is the difference of the sample proportions: .93 – .96 = –.03, and an error bound is $2\sqrt{\frac{.93(.07)}{200}+\frac{.96(.04)}{450}}$ = .041.

b. The above can be expressed as the interval (–.071, .011). Note that the value zero is contained in the interval, so there is reason to believe that the two pain relievers offer the same relief potential.

8.32 With n = 20, the sample mean amount $\bar{y}$ = 197.1 and the standard deviation s = 90.86.

- The *total accounts receivable* is estimated to be 500($\bar{y}$) = 500(197.1) = 98,550. The standard deviation of this estimate is found by $\sqrt{V(500\bar{Y})} = 500\frac{\sigma}{\sqrt{20}}$. So, this can be estimated by 500(90.86)/$\sqrt{20}$ = 10158.45 and an error bound is given by 2(10158.46) = 20316.9.

- With $\bar{y}$ = 197.1, an error bound is 2(90.86)/$\sqrt{20}$ = 40.63. Expressed as an interval, this is (197.1 – 40.63, 197.1 + 40.63) or (156.47, 237.73). So, it is unlikely that the average amount exceeds $250.

8.33 The point estimate is 6/20 = .3 and an error bound is $2\sqrt{\frac{.3(.7)}{20}}$ = .205. If 80% comply, and 20% fail to comply. This value lies within our error bound of the point estimate, so it is likely.

8.34 An unbiased estimator of λ is $\bar{Y}$, and since $\sqrt{V(\bar{Y})} = \sqrt{\lambda/n}$, an unbiased estimator of the standard error of is $\sqrt{\bar{Y}/n}$.

8.35 Using the result of Ex. 8.34:

a. The point estimate is $\bar{y}$ = 20 and a bound on the error of estimation is $2\sqrt{20/50}$ = 1.265.

b. The point estimate is the difference of the sample mean: 20 – 23 = –3.

8.36 An unbiased estimator of θ is $\overline{Y}$, and since $\sqrt{V(\overline{Y})} = \theta/\sqrt{n}$, an unbiased estimator of the standard error of is $\overline{Y}/\sqrt{n}$.

8.37 Refer to Ex. 8.36: with n = 10, an estimate of θ = $\overline{y}$ = 1020 and an error bound is $2\left(1000/\sqrt{10}\right)$ = 645.1.

8.38 To find an unbiased estimator of $V(Y) = \frac{1}{p^2} - \frac{1}{p}$, note that $E(Y) = \frac{1}{p}$ so Y is an unbiased estimator of $\frac{1}{p}$. Further, $E(Y^2) = V(Y) + [E(Y)]^2 = \frac{2}{p^2} - \frac{1}{p}$ so $E(Y^2 + Y) = \frac{2}{p^2}$. Therefore, an unbiased estimate of $V(Y)$ is $\frac{Y^2+Y}{2} + Y = \frac{Y^2-Y}{2}$.

8.39 Using Table 6 with 4 degrees of freedom, $P\left(.71072 \le 2Y/\beta \le 9.48773\right)$ = .90. So,

$$P\left(\tfrac{2Y}{9.48773} \le \beta \le \tfrac{2Y}{.71072}\right) = .90$$

and $\left(\frac{2Y}{9.48773}, \frac{2Y}{.71072}\right)$ forms a 90% CI for β.

8.40 Use the fact that $Z = \frac{Y-\mu}{\sigma}$ has a standard normal distribution. With σ = 1:

a. The 95% CI is (Y – 1.96, Y + 1.96) since

$$P(-1.96 \le Y - \mu \le 1.96) = P(Y - 1.96 \le \mu \le Y + 1.96) = .95.$$

b. The value Y + 1.645 is the 95% upper limit for μ since

$$P(Y - \mu \le 1.645) = P(\mu \le Y + 1.645) = .95.$$

c. Similarly, Y – 1.645 is the 95% lower limit for μ.

8.41 Using Table 6 with 1 degree of freedom:

a. $.95 = P(.0009821 \le Y^2/\sigma^2 \le 5.02389) = P(Y^2/5.02389 \le \sigma^2 \le Y^2/.0009821)$.

b. $.95 = P(.0039321 \le Y^2/\sigma^2) = P(\sigma^2 \le Y^2/.0039321)$.

c. $.95 = P(Y^2/\sigma^2 \le 3.84146) = P(Y^2/3.84146 \le \sigma^2)$.

8.42 Using the results from Ex. 8.41, the square–roots of the boundaries can be taken to obtain interval estimates σ:

a. $Y/2.24 \le \sigma \le Y/.0313$.

b. $\sigma \le Y/.0627$.

c. $\sigma \ge Y/1.96$.

8.43 **a.** The distribution function for $Y_{(n)}$ is $G_n(y) = \left(\frac{y}{\theta}\right)^n$, $0 \le y \le \theta$, so the distribution function for U is given by

$$F_U(u) = P(U \le u) = P(Y_{(n)} \le \theta u) = G_n(\theta u) = u,\ 0 \le y \le 1.$$

b. (Similar to Example 8.5) We require the value a such that $P\left(\frac{Y_{(n)}}{\theta} \le a\right) = F_U(a) = .95$.
Therefore, $a^n = .95$ so that $a = (.95)^{1/n}$ and the lower confidence bound is $[Y_{(n)}](.95)^{-1/n}$.

8.44 **a.** $F_Y(y) = P(Y \le y) = \int_0^y \frac{2(\theta - t)}{\theta^2} dt = \frac{2y}{\theta} - \frac{y^2}{\theta^2}$, $0 < y < \theta$.

b. The distribution of $U = Y/\theta$ is given by
$F_U(u) = P(U \le u) = P(Y \le \theta u) = F_Y(\theta u) = 2u - u^2 = 2u(1-u)$, $0 < u < 1$. Since this distribution does not depend on θ, $U = Y/\theta$ is a pivotal quantity.

c. Set $P(U \le a) = F_Y(a) = 2a(1 - a) = .9$ so that the quadratic expression is solved at $a = 1 - \sqrt{.10} = .6838$ and then the 90% lower bound for θ is $Y/.6838$.

8.45 Following Ex. 8.44, set $P(U \ge b) = 1 - F_Y(b) = 1 - 2b(1 - b) = .9$, thus $b = 1 - \sqrt{.9} = .05132$ and then the 90% upper bound for θ is $Y/.05132$.

8.46 Let $U = 2Y/\theta$ and let $m_Y(t)$ denote the mgf for the exponential distribution with mean θ. Then:

a. $m_U(t) = E(e^{tU}) = E(e^{t2Y/\theta}) = m_Y(2t/\theta) = (1-2t)^{-1}$. This is the mgf for the chi–square distribution with one degree of freedom. Thus, U has this distribution, and since the distribution does not depend on θ, U is a pivotal quantity.

b. Using Table 6 with 2 degrees of freedom, we have

$$P(.102587 \le 2Y/\theta \le 5.99147) = .90.$$

So, $\left(\frac{2Y}{5.99147}, \frac{2Y}{.102587}\right)$ represents a 90% CI for θ.

c. They are equivalent.

8.47 Note that for all i, the mgf for Y_i is $m_Y(t) = (1 - \theta t)^{-1}$, $t < 1/\theta$.

a. Let $U = 2\sum_{i=1}^n Y_i/\theta$. The mgf for U is

$$m_U(t) = E(e^{tU}) = [m_Y(2t/\theta)]^n = (1-2t)^{-n},\ t < 1/2.$$

This is the mgf for the chi–square distribution with $2n$ degrees of freedom. Thus, U has this distribution, and since the distribution does not depend on θ, U is a pivotal quantity.

b. Similar to part b in Ex. 8.46, let $\chi^2_{.975}, \chi^2_{.025}$ be percentage points from the chi–square distribution with $2n$ degrees of freedom such that

$$P\left(\chi^2_{.975} \le 2\sum_{i=1}^n Y_i/\theta \le \chi^2_{.025}\right) = .95.$$

So, $\left(\frac{2\sum_{i=1}^{n} Y_i}{\chi^2_{.975}}, \frac{2\sum_{i=1}^{n} Y_i}{\chi^2_{.025}}\right)$ represents a 95% CI for θ.

c. The CI is $\left(\frac{2(7)(4.77)}{26.1190}, \frac{2(7)(4.77)}{5.62872}\right)$ or (2.557, 11.864).

8.48 (Similar to Ex. 8.47) Note that for all *i*, the mgf for Y_i is $m_Y(t) = (1-\beta)^{-2}$, $t < 1/\beta$.

a. Let $U = 2\sum_{i=1}^{n} Y_i / \beta$. The mgf for *U* is

$$m_U(t) = E(e^{tU}) = \left[m_Y(2t/\beta)\right]^n = (1-2t)^{-2n},\ t < 1/2.$$

This is the mgf for the chi–square distribution with 4*n* degrees of freedom. Thus, *U* has this distribution, and since the distribution does not depend on θ, *U* is a pivotal quantity.

b. Similar to part b in Ex. 8.46, let $\chi^2_{.975}, \chi^2_{.025}$ be percentage points from the chi–square distribution with 4*n* degrees of freedom such that

$$P\left(\chi^2_{.975} \le 2\sum_{i=1}^{n} Y_i / \beta \le \chi^2_{.025}\right) = .95.$$

So, $\left(\frac{2\sum_{i=1}^{n} Y_i}{\chi^2_{.975}}, \frac{2\sum_{i=1}^{n} Y_i}{\chi^2_{.025}}\right)$ represents a 95% CI for β.

c. The CI is $\left(\frac{2(5)(5.39)}{34.1696}, \frac{2(5)(5.39)}{9.59083}\right)$ or (1.577, 5.620).

8.49 **a.** If α = *m* (a known integer), then $U = 2\sum_{i=1}^{n} Y_i / \beta$ still a pivotal quantity and using a mgf approach it can be shown that *U* has a chi–square distribution with *mn* degrees of freedom. So, the interval is

$$\left(\frac{2\sum_{i=1}^{n} Y_i}{\chi^2_{1-\alpha/2}}, \frac{2\sum_{i=1}^{n} Y_i}{\chi^2_{\alpha/2}}\right),$$

where $\chi^2_{1-\alpha/2}, \chi^2_{\alpha/2}$ are percentage points from the chi–square distribution with *mn* degrees of freedom.

b. The quantity $U = \sum_{i=1}^{n} Y_i / \beta$ is distributed as gamma with shape parameter *cn* and scale parameter 1. Since *c* is known, percentiles from this distribution can be calculated from this gamma distribution (denote these as $\gamma_{1-\alpha/2}, \gamma_{\alpha/2}$) so that similar to part a, the CI is

$$\left(\frac{\sum_{i=1}^{n} Y_i}{\gamma_{1-\alpha/2}}, \frac{\sum_{i=1}^{n} Y_i}{\gamma^2_{\alpha/2}}\right).$$

c. Following the notation in part b above, we generate the percentiles using the Applet:

$$\chi_{.975} = 16.74205,\ \chi_{.025} = 36.54688$$

Thus, the CI is $\left(\frac{10(11.36)}{36.54688}, \frac{10(11.36)}{16.74205}\right)$ or (3.108, 6.785).

8.50 **a.** –.1451
b. .2251
c. Brand *A* has the larger proportion of failures, 22.51% greater than Brand *B*.
d. Brand *B* has the larger proportion of failures, 14.51% greater than Brand *A*.
e. There is no evidence that the brands have different proportions of failures, since we are not confident that the brand difference is strictly positive or negative.

8.51 **a.-f**. Answers vary.

8.52 **a.-c**. Answers vary.
d. The proportion of intervals that capture p should be close to .95 (the confidence level).

8.53 **a.** i. Answers vary. ii. smaller confidence level, larger sample size, smaller value of p.
b. Answers vary.

8.54 **a.** The interval is not calculated because the length is zero (the standard error is zero).
b.-d. Answers vary.
e. The sample size is not large (consider the validity of the normal approximation to the binomial).

8.55 Answers vary, but with this sample size, a normal approximation is appropriate.

8.56 **a.** With $z_{.01} = 2.326$, the 98% CI is $.45 \pm 2.326\sqrt{\frac{.45(.55)}{800}}$ or .45 ± .041.
b. Since the value .50 is not contained in the interval, there is not compelling evidence that a majority of adults feel that movies are getting better.

8.57 With $z_{.005} = 2.576$, the 99% interval is $.51 \pm 2.576\sqrt{\frac{.51(.49)}{1001}}$ or .51 ± .04. We are 99% confident that between 47% and 55% of adults in November, 2003 are baseball fans.

8.58 The parameter of interest is μ = mean number of days required for treatment. The 95% CI is approximately $\bar{y} \pm z_{.025}\left(s/\sqrt{n}\right)$, or $5.4 \pm 1.96(3.1/\sqrt{500})$ or (5.13, 5.67).

8.59 **a.** With $z_{.05} = 1.645$, the 90% interval is $.78 \pm 1.645\sqrt{\frac{.78(.22)}{1030}}$ or .78 ± .021.
b. The lower endpoint of the interval is .78 – .021 = .759, so there is evidence that the true proportion is greater than 75%.

8.60 **a.** With $z_{.005} = 2.576$, the 99% interval is $98.25 \pm 2.576\left(.73/\sqrt{130}\right)$ or 98.25 ± .165.

b. Written as an interval, the above is (98.085, 98.415). So, the "normal" body temperature measurement of 98.6 degrees is not contained in the interval. It is possible that the standard for "normal" is no longer valid.

8.61 With $z_{.025} = 1.96$, the 95% CI is $167.1 - 140.9 \pm 1.96\sqrt{\frac{(24.3)^2 + (17.6)^2}{30}}$ or (15.46, 36.94).

8.62 With $z_{.005} = 2.576$, the approximate 99% CI is $24.8 - 21.3 \pm 2.576\sqrt{\frac{(7.1)^2}{34} + \frac{(8.1)^2}{41}}$ or $(-1.02, 8.02)$. With 99% confidence, the difference in mean molt time for normal males versus those split from their mates is between (–1.02, 8.02).

8.63 **a.** With $z_{.025} = 1.96$, the 95% interval is $.78 \pm 1.96\sqrt{\frac{.78(.22)}{1000}}$ or .78 ± .026 or (.754, .806).

b. The margin of error reported in the article is larger than the 2.6% calculated above. Assuming that a 95% CI was calculated, a value of $p = .5$ gives the margin of error 3.1%.

8.64 **a.** The point estimates are .35 (sample proportion of 18-34 year olds who consider themselves patriotic) and .77 (sample proportion of 60+ year olds who consider themselves patriotic. So, a 98% CI is given by (here, $z_{.01} = 2.326$)

$$.77 - .35 \pm 2.326\sqrt{\frac{(.77)(.23)}{150} + \frac{(.35)(.65)}{340}} \text{ or } .42 \pm .10 \text{ or } (.32, .52).$$

b. Since the value for the difference .6 is outside of the above CI, this is not a likely value.

8.65 **a.** The 98% CI is, with $z_{.01} = 2.326$, is

$$.18 - .12 \pm 2.326\sqrt{\frac{.18(.82) + .12(.88)}{100}} \text{ or } .06 \pm .117 \text{ or } (-.057, .177).$$

b. Since the interval contains both positive and negative values, it is likely that the two assembly lines produce the same proportion of defectives.

8.66 **a.** With $z_{.05} = 1.645$, the 90% CI for the mean posttest score for all BACC students is $18.5 \pm 1.645\left(\frac{8.03}{\sqrt{365}}\right)$ or 18.5 ± .82 or (17.68, 19.32).

b. With $z_{.025} = 1.96$, the 95% CI for the difference in the mean posttest scores for BACC and traditionally taught students is $(18.5 - 16.5) \pm 1.96\sqrt{\frac{(8.03)^2}{365} + \frac{(6.96)^2}{298}}$ or 2.0 ± 1.14.

c. Since 0 is outside of the interval, there is evidence that the mean posttest scores are different.

8.67 **a.** The 95% CI is $7.2 \pm 1.96\sqrt{\frac{8.8}{60}}$ or 7.2 ± .75 or (6.45, 7.95).

b. The 90% CI for the difference in the mean densities is $(7.2 - 4.7) \pm 1.645\sqrt{\frac{8.8}{60} + \frac{4.9}{90}}$ or $2.5 \pm .74$ or (1.76, 3.24).

c. Presumably, the population is ship sightings for all summer and winter months. It is quite possible that the days used in the sample were not randomly selected (the months were chosen in the same year.)

8.68 **a.** Recall that for the multinomial, $V(Y_i) = np_iq_i$ and $\text{Cov}(Y_i,Y_j) = -np_ip_j$ for $i \neq j$. Hence,

$$V(Y_1 - Y_2) = V(Y_1) + V(Y_2) - 2\text{Cov}(Y_1, Y_2) = np_1q_1 + np_2q_2 + 2np_1p_2.$$

b. Since $\hat{p}_1 - \hat{p}_2 = \frac{Y_1 - Y_2}{n}$, using the result in part a we have

$$V(\hat{p}_1 - \hat{p}_2) = \tfrac{1}{n}\left(p_1q_1 + p_2q_2 + 2p_1p_2\right).$$

Thus, an approximate 95% CI is given by

$$\hat{p}_1 - \hat{p}_2 \pm 1.96\sqrt{\tfrac{1}{n}\left(\hat{p}_1\hat{q}_1 + \hat{p}_2\hat{q}_2 + 2\hat{p}_1\hat{p}_2\right)}$$

Using the supplied data, this is

$.06 - .16 \pm 1.96\sqrt{\frac{1}{500}\left(.06(.94) + .16(.84) + 2(.06)(.16)\right)} = -.10 \pm .04$ or $(-.14, -.06)$.

8.69 For the independent counts Y_1, Y_2, Y_3, and Y_4, the sample proportions are $\hat{p}_i = Y_i / n_i$ and $V(\hat{p}_i) = p_iq_i / n_i$ for $i = 1, 2, 3, 4$. The interval of interest can be constructed as

$$(\hat{p}_3 - \hat{p}_1) - (\hat{p}_4 - \hat{p}_2) \pm 1.96\sqrt{V[(\hat{p}_3 - \hat{p}_1) - (\hat{p}_4 - \hat{p}_2)]}\,.$$

By independence, this is

$$(\hat{p}_3 - \hat{p}_1) - (\hat{p}_4 - \hat{p}_2) \pm 1.96\sqrt{\tfrac{1}{n}[\hat{p}_3\hat{q}_3 + \hat{p}_1\hat{q}_1 + \hat{p}_4\hat{q}_4 + \hat{p}_2\hat{q}_2]}\,.$$

Using the sample data, this is

$$(.69 - .65) - (.25 - .43) \pm 1.96\sqrt{\tfrac{1}{500}[.65(.35) + .43(.57) + .69(.31) + .25(.75)}$$

or $.22 \pm .34$ or $(-.12, .56)$

8.70 As with Example 8.9, we must solve the equation $1.96\sqrt{\frac{pq}{n}} = B$ for n.

a. With $p = .9$ and $B = .05$, $n = 139$.
b. If p is unknown, use $p = .5$ so $n = 385$.

8.71 With $B = 2$, $\sigma = 10$, $n = 4\sigma^2/B^2$, so $n = 100$.

8.72 **a.** Since the true proportions are unknown, use .5 for both to compute an error bound (here, we will use a multiple of 1.96 that correlates to a 95% CI):

$$1.96\sqrt{\tfrac{.5(.5)}{1000} + \tfrac{.5(.5)}{1000}} = .044.$$

b. Assuming that the two sample sizes are equal, solve the relation

$$1.645\sqrt{\tfrac{.5(.5)}{n} + \tfrac{.5(.5)}{n}} = .02\,,$$

so $n = 3383$.

8.73 From the previous sample, the proportion of 'tweens who understand and enjoy ads that are silly in nature is .78. Using this as an estimate of p, we estimate the sample size as

$$2.576\sqrt{\tfrac{.78(.22)}{n}} = .02 \text{ or } n = 2847.$$

8.74 With $B = .1$ and $\sigma = .5$, $n = (1.96)^2\sigma^2/B^2$, so $n = 97$. If all of the specimens were selected from a single rainfall, the observations would not be independent.

8.75 Here, $1.645\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = .1$, but $\sigma_1^2 = \sigma_2^2 = .25$, $n_1 = n_2 = n$, so sample $n = 136$ from each location.

8.76 For $n_1 = n_2 = n$ and by using the estimates of population variances given in Ex. 8.61, we can solve $1.645\sqrt{\frac{(24.3)^2+(17.6)^2}{n}} = 5$ so that $n = 98$ adults must be selected from each region.

8.77 Using the estimates $\hat{p}_1 = .7$, $\hat{p}_2 = .54$, the relation is $1.645\sqrt{\frac{.7(.3)+.54(.46)}{n}} = .05$ so $n = 497$.

8.78 Here, we will use the estimates of the true proportions of defectives from Ex. 8.65. So, with a bound $B = (.2)/2 = .1$, the relation is $1.96\sqrt{\frac{.18(.82)+.12(.88)}{n}} = .1$ so $n = 98$.

8.79 **a.** Here, we will use the estimates of the population variances for the two groups of students:

$$2.576\sqrt{\frac{(8.03)^2}{n} + \frac{(6.96)^2}{n}} = .5,$$

so $n = 2998$ students from each group should be sampled.

b. For comparing the mean pretest scores, $s_1 = 5.59$, $s_2 = 5.45$ so $2.576\sqrt{\frac{(5.59)^2}{n} + \frac{(5.45)^2}{n}} = .5$ and thus $n = 1618$ students from each group should be sampled.

c. If it is required that all four sample sizes must be equal, use $n = 2998$ (from part a) to assure an interval width of 1 unit.

8.80 The 95% CI, based on a t–distribution with 21 – 1 = 20 degrees of freedom, is

$$26.6 \pm 2.086\left(7.4/\sqrt{21}\right) = 26.6 \pm 3.37 \text{ or } (23.23, 29.97).$$

8.81 The sample statistics are $\bar{y} = 60.8$, $s = 7.97$. So, the 95% CI is

$$60.8 \pm 2.262\left(7.97/\sqrt{10}\right) = 60.8 \pm 5.70 \text{ or } (55.1, 66.5).$$

8.82 **a.** The 90% CI for the mean verbal SAT score for urban high school seniors is

$$505 \pm 1.729\left(57/\sqrt{20}\right) = 505 \pm 22.04 \text{ or } (482.96, 527.04).$$

b. Since the interval includes the score 508, it is a plausible value for the mean.

c. The 90% CI for the mean math SAT score for urban high school seniors is

$$495 \pm 1.729\left(69/\sqrt{20}\right) = 495 \pm 26.68 \text{ or } (468.32, 521.68).$$

The interval does include the score 520, so the interval supports the stated true mean value.

8.83 **a.** Using the sample–sample CI for $\mu_1 - \mu_2$, using an assumption of normality, we calculate the pooled sample variance

$$s_p^2 = \tfrac{9(3.92)^2 + 9(3.98)^2}{18} = 15.6034$$

Thus, the 95% CI for the difference in mean compartment pressures is

$$14.5–11.1 \pm 2.101\sqrt{15.6034\left(\tfrac{1}{10} + \tfrac{1}{10}\right)} = 3.4 \pm 3.7 \text{ or } (–.3, 7.1).$$

b. Similar to part a, the pooled sample variance for runners and cyclists who exercise at 80% maximal oxygen consumption is given by

$$s_p^2 = \tfrac{9(3.49)^2 + 9(4.95)^2}{18} = 18.3413.$$

The 90% CI for the difference in mean compartment pressures here is

$$12.2–11.5 \pm 1.734\sqrt{18.3413\left(\tfrac{1}{10} + \tfrac{1}{10}\right)} = .7 \pm 3.32 \text{ or } (–2.62, 4.02).$$

c. Since both intervals contain 0, we cannot conclude that the means in either case are different from one another.

8.84 The sample statistics are $\bar{y} = 3.781$, $s = .0327$. So, the 95% CI, with 9 degrees of freedom and $t_{.025} = 2.262$, is

$$3.781 \pm 2.262\left(.0327/\sqrt{10}\right) = 3.781 \pm .129 \text{ or } (3.652, 3.910).$$

8.85 The pooled sample variance is $s_p^2 = \tfrac{15(6)^2 + 19(8)^2}{34} = 51.647$. Then the 95% CI for $\mu_1 - \mu_2$ is

$$11 - 12 \pm 1.96\sqrt{51.647\left(\tfrac{1}{16} + \tfrac{1}{20}\right)} = -1 \pm 4.72 \text{ or } (–5.72, 3.72)$$

(here, we approximate $t_{.025}$ with $z_{.025} = 1.96$).

8.86 **a.** The sample statistics are, with $n = 14$, $\bar{y} = 0.896$, $s = .400$. The 95% CI for μ = mean price of light tuna in water, with 13 degrees of freedom and $t_{.025} = 2.16$ is

$$.896 \pm 2.16\left(.4/\sqrt{14}\right) = .896 \pm .231 \text{ or } (.665, 1.127).$$

b. The sample statistics are, with $n = 11$, $\bar{y} = 1.147$, $s = .679$. The 95% CI for μ = mean price of light tuna in oil, with 10 degrees of freedom and $t_{.025} = 2.228$ is

$$1.147 \pm 2.228\left(.679/\sqrt{11}\right) = 1.147 \pm .456 \text{ or } (.691, 1.603).$$

This CI has a larger width because: s is larger, n is smaller, $t_{\alpha/2}$ is bigger.

8.87 **a.** Following Ex. 8.86, the pooled sample variance is $s_p^2 = \tfrac{13(.4)^2 + 10(.679)^2}{23} = .291$. Then the 90% CI for $\mu_1 - \mu_2$, with 23 degrees of freedom and $t_{.05} = 1.714$ is

$$(.896 - 1.147) \pm 1.714\sqrt{.291\left(\tfrac{1}{14} + \tfrac{1}{11}\right)} = -.251 \pm .373 \text{ or } (-.624, .122).$$

b. Based on the above interval, there is not compelling evidence that the mean prices are different since 0 is contained inside the interval.

8.88 The sample statistics are, with $n = 12$, $\bar{y} = 9$, $s = 6.4$. The 90% CI for μ = mean LC50 for DDT is, with 11 degrees of freedom and $t_{.05} = 1.796$,

$$9 \pm 1.796\left(6.4/\sqrt{12}\right) = 9 \pm 3.32 \text{ or } (5.68, 12.32).$$

8.89 **a.** For the three LC50 measurements of Diazinon, $\bar{y} = 3.57$, $s = 3.67$. The 90% CI for the true mean is (2.62, 9.76).

b. The pooled sample variance is $s_p^2 = \frac{11(6.4)^2 + 2(3.57)^2}{13} = 36.6$. Then the 90% CI for the difference in mean LC50 chemicals, with 15 degrees of freedom and $t_{.025} = 1.771$, is

$$(9 - 3.57) \pm 1.771\sqrt{36.6\left(\tfrac{1}{12} + \tfrac{1}{3}\right)} = 5.43 \pm 6.92 \text{ or } (-1.49, 12.35).$$

We assumed that the sample measurements were independently drawn from normal populations with $\sigma_1 = \sigma_2$.

8.90 **a.** For the 95% CI for the difference in mean verbal scores, the pooled sample variance is $s_p^2 = \frac{14(42)^2 + 14(45)^2}{28} = 1894.5$ and thus

$$446 - 534 \pm 2.048\sqrt{1894.5\left(\tfrac{2}{15}\right)} = -88 \pm 32.55 \text{ or } (-120.55, -55.45).$$

b. For the 95% CI for the difference in mean math scores, the pooled sample variance is $s_p^2 = \frac{14(57)^2 + 14(52)^2}{28} = 2976.5$ and thus

$$548 - 517 \pm 2.048\sqrt{2976.5\left(\tfrac{2}{15}\right)} = 31 \pm 40.80 \text{ or } (-9.80, 71.80).$$

c. At the 95% confidence level, there appears to be a difference in the two mean verbal SAT scores achieved by the two groups. However, a difference is not seen in the math SAT scores.

d. We assumed that the sample measurements were independently drawn from normal populations with $\sigma_1 = \sigma_2$.

8.91 Sample statistics are:

Season	sample mean	sample variance	sample size
spring	15.62	98.06	5
summer	72.28	582.26	4

The pooled sample variance is $s_p^2 = \frac{4(98.06) + 3(582.26)}{7} = 305.57$ and thus the 95% CI is

$$15.62 - 72.28 \pm 2.365\sqrt{305.57\left(\tfrac{1}{5}+\tfrac{1}{4}\right)} = -56.66 \pm 27.73 \text{ or } (-84.39, -28.93).$$

It is assumed that the two random samples were independently drawn from normal populations with equal variances.

8.92 Using the summary statistics, the pooled sample variance is $s_p^2 = \frac{3(.001)+4(.002)}{7} = .0016$ and so the 95% CI is given by

$$.22 - .17 \pm 2.365\sqrt{.0016\left(\tfrac{1}{4}+\tfrac{1}{5}\right)} = .05 \pm .063 \text{ or } (-.013, .113).$$

8.93 **a.** Since the two random samples are assumed to be independent and normally distributed, the quantity $2\overline{X}+\overline{Y}$ is normally distributed with mean $2\mu_1 + \mu_2$ and variance $\left(\frac{4}{n}+\frac{3}{m}\right)\sigma^2$. Thus, is σ^2 is known, then $2\overline{X}+\overline{Y} \pm 1.96\sigma\sqrt{\frac{4}{n}+\frac{3}{m}}$ is a 95% CI for $2\mu_1 + \mu_2$.

b. Recall that $(1/\sigma^2)\sum_{i=1}^{n}(X_i-\overline{X})^2$ has a chi–square distribution with $n - 1$ degrees of freedom. Thus, $[1/(3\sigma^2)]\sum_{i=1}^{m}(Y_i-\overline{Y})^2$ is chi–square with $m - 1$ degrees of freedom and the sum of these is chi–square with $n + m - 2$ degrees of freedom. Then, by using Definition 7.2, the quantity

$$T = \frac{2\overline{X}+\overline{Y}-(2\mu_1+\mu_2)}{\hat{\sigma}\sqrt{\frac{4}{n}+\frac{3}{m}}}, \text{ where}$$

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{n}(X_i-\overline{X})^2 + \frac{1}{3}\sum_{i=1}^{m}(Y_i-\overline{Y})^2}{n+m-2}.$$

Then, the 95% CI is given by $2\overline{X}+\overline{Y} \pm t_{.025}\hat{\sigma}\sqrt{\frac{4}{n}+\frac{3}{m}}$.

8.94 The pivotal quantity is $T = \dfrac{\overline{Y}_1-\overline{Y}_2-(\mu_1-\mu_2)}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$, which has a t–distribution w/ $n_1 + n_2 - 2$ degrees of freedom. By selecting t_α from this distribution, we have that $P(T < t_\alpha) = 1 - \alpha$. Using the same approach to derive the confidence interval, it is found that

$$\overline{Y}_1-\overline{Y}_2 \pm t_\alpha S_p\sqrt{\tfrac{1}{n_1}+\tfrac{1}{n_2}}$$

is a $100(1 - \alpha)$% upper confidence bound for $\mu_1 - \mu_2$.

8.95 From the sample data, $n = 6$ and $s^2 = .503$. Then, $\chi^2_{.95} = 1.145476$ and $\chi^2_{.05} = 11.0705$ with 5 degrees of freedom. The 90% CI for σ^2 is $\left(\frac{5(.503)}{11.0705}, \frac{5(.503)}{1.145476}\right)$ or (.227, 2.196). We are 90% confident that σ^2 lies in this interval.

8.96 From the sample data, $n = 10$ and $s^2 = 63.5$. Then, $\chi^2_{.95} = 3.3251$ and $\chi^2_{.05} = 16.9190$ with 9 degrees of freedom. The 90% CI for σ^2 is $\left(\frac{571.6}{16.9190}, \frac{571.6}{3.3251}\right)$ or (33.79, 171.90).

8.97 **a.** Note that $1-\alpha = P\left(\frac{(n-1)S^2}{\sigma^2} > \chi^2_{1-\alpha}\right) = P\left(\frac{(n-1)S^2}{\chi^2_{1-\alpha}} > \sigma^2\right)$. Then, $\frac{(n-1)S^2}{\chi^2_{1-\alpha}}$ is a 100(1–α)% upper confidence bound for σ^2.

b. Similar to part (a), it can be shown that $\frac{(n-1)S^2}{\chi^2_{\alpha}}$ is a 100(1–α)% lower confidence bound for σ^2.

8.98 The confidence interval for σ^2 is $\left(\frac{(n-1)S^2}{\chi^2_{1-\alpha/2}}, \frac{(n-1)S^2}{\chi^2_{\alpha/2}}\right)$, so since $S^2 > 0$, the confidence interval for σ is simply $\left(\sqrt{\frac{(n-1)S^2}{\chi^2_{1-\alpha/2}}}, \sqrt{\frac{(n-1)S^2}{\chi^2_{\alpha/2}}}\right)$.

8.99 Following Ex. 8.97 and 8.98:

a. 100(1 – α)% upper confidence bound for σ: $\sqrt{\frac{(n-1)S^2}{\chi^2_{1-\alpha}}}$.

b. 100(1 – α)% lower confidence bound for σ: $\sqrt{\frac{(n-1)S^2}{\chi^2_{\alpha}}}$.

8.100 With $n = 20$, the sample variance $s^2 = 34854.4$. From Ex. 8.99, a 99% upper confidence bound for the standard deviation σ is, with $\chi^2_{.99} = 7.6327$,

$$\sqrt{\frac{19(34854.4)}{7.6327}} = 294.55.$$

Since this is an upper bound, it is possible that the true population standard deviation is less than 150 hours.

8.101 With $n = 6$, the sample variance $s^2 = .0286$. Then, $\chi^2_{.95} = 1.145476$ and $\chi^2_{.05} = 11.0705$ with 5 degrees of freedom and the 90% CI for σ^2 is

$$\left(\frac{5(.0286)}{11.0705}, \frac{5(.0286)}{1.145476}\right) = (.013\ .125).$$

8.102 With $n = 5$, the sample variance $s^2 = 144.5$. Then, $\chi^2_{.995} = .20699$ and $\chi^2_{.005} = 14.8602$ with 4 degrees of freedom and the 99% CI for σ^2 is

$$\left(\frac{4(144.5)}{14.8602}, \frac{4(144.5)}{.20699}\right) = (38.90, 2792.41).$$

8.103 With $n = 4$, the sample variance $s^2 = 3.67$. Then, $\chi^2_{.95} = .351846$ and $\chi^2_{.05} = 7.81473$ with 3 degrees of freedom and the 99% CI for σ^2 is

$$\left(\frac{3(3.67)}{7.81473}, \frac{3(3.67)}{.351846}\right) = (1.4, 31.3).$$

An assumption of independent measurements and normality was made. Since the interval implies that the *standard deviation* could be larger than 5 units, it is possible that the instrument could be off by more than two units.

8.104 The only correct interpretation is choice d.

8.105 The difference of the endpoints 7.37 – 5.37 = 2.00 is equal to $2z_{\alpha/2}\sqrt{\frac{\sigma^2}{n}} = 2z_{\alpha/2}\sqrt{\frac{6}{25}}$. Thus, $z_{\alpha/2} \approx 2.04$ so that α/2 = .0207 and the confidence coefficient is 1 – 2(.0207) = .9586.

8.106 **a.** Define: p_1 = proportion of survivors in low water group for male parents
p_2 = proportion of survivors in low nutrient group for male parents

Then, the sample estimates are $\hat{p}_1$ = 522/578 = .903 and $\hat{p}_2$ = 510/568 = .898. The 99% CI for the difference $p_1 - p_2$ is

$$.903 - .898 \pm 2.576\sqrt{\frac{.903(.097)}{578} + \frac{.898(.102)}{568}} = .005 \pm .0456 \text{ or } (-.0406, .0506).$$

b. Define: p_1 = proportion of male survivors in low water group
p_2 = proportion of female survivors in low water group

Then, the sample estimates are $\hat{p}_1$ = 522/578 = .903 and $\hat{p}_2$ = 466/510 = .914. The 99% CI for the difference $p_1 - p_2$ is

$$.903 - .914 \pm 2.576\sqrt{\frac{.903(.097)}{578} + \frac{.914(.086)}{510}} = -.011 \pm .045 \text{ or } (-.056, .034).$$

8.107 With B = .03 and α = .05, we use the sample estimates of the proportions to solve

$$1.96\sqrt{\frac{.903(.097)}{n} + \frac{.898(.102)}{n}} = .03.$$

The solution is n = 764.8, therefore 765 seeds should be used in each environment.

8.108 If it is assumed that p = kill rate = .6, then this can be used in the sample size formula with B = .02 to obtain (since a confidence coefficient was not specified, we are using a multiple of 2 for the error bound)

$$.02 = 2\sqrt{\frac{.6(.4)}{n}}.$$

So, n = 2400.

8.109 **a.** The sample proportion of unemployed workers is 25/400 = .0625, and a two–standard–error bound is given by $2\sqrt{\frac{.0625(.9375)}{400}} = .0242$.

b. Using the same estimate of p, the true proportion of unemployed workers, gives the relation $2\sqrt{\frac{.0625(.9375)}{n}} = .02$. This is solved by n = 585.94, so 586 people should be sampled.

8.110 For an error bound of \$50 and assuming that the population standard deviation σ = 400, the equation to be solved is

$$1.96\frac{400}{\sqrt{n}} = 50.$$

This is solved by n = 245.96, so 246 textile workers should be sampled.

8.111 Assuming that the true proportion p = .5, a confidence coefficient of .95 and desired error of estimation B = .005 gives the relation

$$1.96\sqrt{\tfrac{.5(.5)}{n}} = .005.$$

The solution is n = 38,416.

8.112 The goal is to estimate the difference of

p_1 = proportion of all fraternity men favoring the proposition
p_2 = proportion of all non–fraternity men favoring the proposition

A point estimate of $p_1 - p_2$ is the difference of the sample proportions:

$$300/500 - 64/100 = .6 - .64 = -.04.$$

A two–standard–error bound is

$$2\sqrt{\tfrac{.6(.4)}{500} + \tfrac{.64(.36)}{100}} = .106.$$

8.113 Following Ex. 112, assuming equal sample sizes and population proportions, the equation that must be solved is

$$2\sqrt{\tfrac{.6(.4)}{n} + \tfrac{.6(.4)}{n}} = .05.$$

Here, n = 768.

8.114 The sample statistics are $\bar{y}$ = 795 and s = 8.34 with n = 5. The 90% CI for the mean daily yield is

$$795 \pm 2.132\left(8.34/\sqrt{5}\right) = 795 \pm 7.95 \text{ or } (787.05, 802.85).$$

It was necessary to assume that the process yields follow a normal distribution and that the measurements represent a random sample.

8.115 Following Ex. 8.114 w/ 5 – 1 = 4 degrees of freedom, $\chi^2_{.95}$ = .710721 and $\chi^2_{.05}$ = 9.48773. The 90% CI for σ^2 is (note that $4s^2$ = 278)

$$\left(\tfrac{278}{9.48773}, \tfrac{278}{.710721}\right) \text{ or } (29.30, 391.15).$$

8.116 The 99% CI for μ is given by, with 15 degrees of freedom and $t_{.005}$ = 2.947, is

$$79.47 \pm 2.947\left(25.25/\sqrt{16}\right) = 79.47 \pm 18.60 \text{ or } (60.87, 98.07).$$

We are 99% confident that the true mean long–term word memory score is contained in the interval.

8.117 The 90% CI for the mean annual main stem growth is given by

$$11.3 \pm 1.746\left(3.4/\sqrt{17}\right) = 11.3 \pm 1.44 \text{ or } (9.86, 12.74).$$

8.118 The sample statistics are $\bar{y}$ = 3.68 and s = 1.905 with n = 6. The 90% CI for the mean daily yield is

$$3.68 \pm 2.015\left(1.905/\sqrt{6}\right) = 3.68 \pm 1.57 \text{ or } (2.11, 5.25).$$

8.119 Since both sample sizes are large, we can use the large sample CI for the difference of population means:

$$75 - 72 \pm 1.96\sqrt{\tfrac{10^2}{50} + \tfrac{8^2}{45}} = 3 \pm 3.63 \text{ or } (-.63, 6.63).$$

8.120 Here, we will assume that the two samples of test scores represent random samples from normal distributions with $\sigma_1 = \sigma_2$. The pooled sample variance is $s_p^2 = \frac{10(52)+13(71)}{23} = 62.74$. The 95% CI for $\mu_1 - \mu_2$ is given by

$$64 - 69 \pm 2.069\sqrt{62.74\left(\tfrac{1}{11} + \tfrac{1}{14}\right)} = -5 \pm 6.60 \text{ or } (-11.60, 1.60).$$

8.121 Assume the samples of reaction times represent random sample from normal populations with $\sigma_1 = \sigma_2$. The sample statistics are: $\bar{y}_1 = 1.875$, $s_1^2 = .696$, $\bar{y}_2 = 2.625$, $s_2^2 = .839$. The pooled sample variance is $s_p^2 = \frac{7(.696)+7(.839)}{14} = .7675$ and the 90% CI for $\mu_1 - \mu_2$ is

$$1.875 - 2.625 \pm 1.761\sqrt{.7675\left(\tfrac{2}{8}\right)} = -.75 \pm .77 \text{ or } (-1.52, .02).$$

8.122 A 90% CI for μ = mean time between billing and payment receipt is, with $z_{.05} = 1.645$ (here we can use the large sample interval formula),

$$39.1 \pm 1.645\left(17.3/\sqrt{100}\right) = 39.1 \pm 2.846 \text{ or } (36.25, 41.95).$$

We are 90% confident that the true mean billing time is contained in the interval.

8.123 The sample proportion is 1914/2300 = .832. A 95% CI for p = proportion of all viewers who misunderstand is

$$.832 \pm 1.96\sqrt{\tfrac{.832(.168)}{2300}} = .832 \pm .015 \text{ or } (.817, .847).$$

8.124 The sample proportion is 278/415 = .67. A 95% CI for p = proportion of all corporate executives who consider cash flow the most important measure of a company's financial health is

$$.67 \pm 1.96\sqrt{\tfrac{.67(.33)}{415}} = .67 \pm .045 \text{ or } (.625, .715).$$

8.125 a. From Definition 7.3, the following quantity has an F–distribution with $n_1 - 1$ numerator and $n_2 - 1$ denominator degrees of freedom:

$$F = \frac{\frac{(n_1-1)S_1^2}{\sigma_1^2}\Big/(n_1 - 1)}{\frac{(n_2-1)S_2^2}{\sigma_2^2}\Big/(n_2 - 1)} = \frac{S_1^2}{S_2^2} \times \frac{\sigma_2^2}{\sigma_1^2}.$$

b. By choosing quantiles from the F–distribution with $n_1 - 1$ numerator and $n_2 - 1$ denominator degrees of freedom, we have

$$P(F_{1-\alpha/2} < F < F_{\alpha/2}) = 1 - \alpha.$$

Using the above random variable gives

$$P\left(F_{1-\alpha/2} < \frac{S_1^2}{S_2^2} \times \frac{\sigma_2^2}{\sigma_1^2} < F_{\alpha/2}\right) = P\left(\frac{S_2^2}{S_1^2} F_{1-\alpha/2} < \frac{\sigma_2^2}{\sigma_1^2} < \frac{S_2^2}{S_1^2} F_{\alpha/2}\right) = 1 - \alpha.$$

Thus,

$$\left(\frac{S_2^2}{S_1^2} F_{1-\alpha/2}, \frac{S_2^2}{S_1^2} F_{\alpha/2} \right)$$

is a 100(1 – α)% CI for σ_2^2/σ_1^2.

An alternative expression is given by the following. Let $F_{\nu_2,\alpha}^{\nu_1}$ denote the upper–α critical value from the *F*–distribution with ν_1 numerator and ν_2 denominator degrees of freedom. Because of the relationship (see Ex. 7.29)

$$F_{\nu_2,\alpha}^{\nu_1} = \frac{1}{F_{\nu_1,\alpha}^{\nu_2}},$$

a 100(1 – α)% CI for σ_2^2/σ_1^2 is also given by

$$\left(\frac{1}{F_{\nu_1,\alpha}^{\nu_2}} \frac{S_2^2}{S_1^2}, F_{\nu_2,\alpha}^{\nu_1} \frac{S_2^2}{S_1^2} \right).$$

8.126 Using the CI derived in Ex. 8.126, we have that $F_{9,.025}^{9} = \frac{1}{F_{9,.025}^{9}} = 4.03$. Thus, the CI for the ratio of the true population variances is $\left(\frac{1}{4.03} \cdot \frac{.094}{.273}, \frac{4.03(.094)}{.273} \right)$ = (.085, 1.39).

8.127 It is easy to show (e.g. using the mgf approach) that $\overline{Y}$ has a gamma distribution with shape parameter $100c_0$ and scale parameter (.01)β. In addition the statistic $U = \overline{Y}/\beta$ is a pivotal quantity since the distribution is free of β: the distribution of U is gamma with shape parameter $100c_0$ and scale parameter (.01). Now, $E(U) = c_0$ and $V(U) = (.01)c_0$ and by the Central Limit Theorem,

$$\frac{U - c_0}{.1\sqrt{c_0}} = \frac{\overline{Y}/\beta - c_0}{.1\sqrt{c_0}}$$

has an approximate standard normal distribution. Thus,

$$P\left(-z_{\alpha/2} < \frac{\overline{Y}/\beta - c_0}{.1\sqrt{c_0}} < z_{\alpha/2} \right) \approx 1 - \alpha .$$

Isolating the parameter β in the above inequality yields the desired result.

8.128 **a.** Following the notation of Section 8.8 and the assumptions given in the problem, we know that $\overline{Y}_1 - \overline{Y}_2$ is a normal variable with mean $\mu_1 - \mu_2$ and variance $\frac{\sigma_1^2}{n_1} + \frac{k\sigma_1^2}{n_2}$. Thus, the standardized variable Z^* as defined indeed has a standard normal distribution.

b. The quantities $U_1 = \frac{(n_1 - 1)S_1^2}{\sigma_1^2}$ and $U_2 = \frac{(n_2 - 1)S_2^2}{k\sigma_1^2}$ have independent chi–square distributions with $n_1 - 1$ and $n_2 - 1$ degrees of freedom (respectively). So, $W^* = U_1 + U_2$ has a chi–square distribution with $n_1 + n_2 - 2$ degrees of freedom.

c. By Definition 7.2, the quantity $T^* = \dfrac{Z^*}{\sqrt{W^*/(n_1+n_2-2)}}$ follows a t–distribution with $n_1 + n_2 - 2$ degrees of freedom.

d. A 100(1 – α)% CI for $\mu_1 - \mu_2$ is given by $\overline{Y}_1 - \overline{Y}_2 \pm t_{\alpha/2} S_p^* \sqrt{\frac{1}{n_1} + \frac{k}{n_2}}$, where $t_{\alpha/2}$ is the upper–α/2 critical value from the t–distribution with $n_1 + n_2 - 2$ degrees of freedom and S_p^* is defined in part (c).

e. If $k = 1$, it is equivalent to the result for $\sigma_1 = \sigma_2$.

8.129 Recall that $V(S^2) = \frac{2\sigma^4}{n-1}$.

a. $V(S'^2) = V(\frac{n-1}{n} S^2) = \frac{2(n-1)\sigma^4}{n^2}$.

b. The result follows from $V(S'^2) = V(\frac{n-1}{n} S^2) = \left(\frac{n-1}{n}\right)^2 V(S^2) < V(S^2)$ since $\frac{n-1}{n} < 1$.

8.130 Since S^2 is unbiased,

$\text{MSE}(S^2) = V(S^2) = \frac{2\sigma^4}{n-1}$. Similarly,

$\text{MSE}(S'^2) = V(S'^2) + [B(S'^2)]^2 = \frac{2(n-1)\sigma^4}{n^2} + \left(\frac{n-1}{n}\sigma^2 - \sigma^2\right)^2 = \frac{(2n-1)\sigma^4}{n^2}$.

By considering the ratio of these two MSEs, it can be seen that S'^2 has the smaller MSE and thus possibly a better estimator.

8.131 Define the estimator $\hat{\sigma}^2 = c\sum_{i=1}^{n}(Y_i - \overline{Y})^2$. Therefore, $E(\hat{\sigma}^2) = c(n-1)\sigma^2$ and $V(\hat{\sigma}^2) = 2c^2(n-1)\sigma^4$ so that

$$\text{MSE}(\hat{\sigma}^2) = 2c^2(n-1)\sigma^4 + [c(n-1)\sigma^2 - \sigma^2]^2.$$

Minimizing this quantity with respect to c, we find that the smallest MSE occurs when $c = \frac{1}{n+1}$.

8.132 **a.** The distribution function for $Y_{(n)}$ is given by

$$F_{Y_{(n)}}(y) = P(Y_{(n)} < y) = [F(y)]^n = \left(\frac{y}{\theta}\right)^{cn}, \; 0 \le y \le \theta.$$

b. The distribution of $U = Y_{(n)}/\theta$ is

$$F_U(u) = P(U \le u) = P(Y_{(n)} \le \theta u) = u^{nc}, \; 0 \le u \le 1.$$

Since this distribution is free of θ, $U = Y_{(n)}/\theta$ is a pivotal quantity. Also,

$$P\big(k < Y(n)/\theta \le 1\big) = P\big(k\theta < Y(n) \le \theta\big) = F_{Y_{(n)}}(\theta) - F_{Y_{(n)}}(k\theta) = 1 - k^{cn}.$$

c. i. Using the result from part b with $n = 5$ and $c = 2.4$,

$.95 = 1 - (k)^{12}$ so $k = .779$

ii. Solving the equations .975 = 1 – $(k_1)^{12}$ and .025 = 1 – $(k_2)^{12}$, we obtain k_1 = .73535 and k_2 = .99789. Thus,

$$P(.73535 < Y_{(5)}/\theta < .99789) = P\left(\frac{Y_{(5)}}{.99789} < \theta < \frac{Y_{(5)}}{.73535}\right) = .95.$$

So, $\left(\frac{Y_{(5)}}{.99789}, \frac{Y_{(5)}}{.73535}\right)$ is a 95% CI for θ.

8.133 We know that $E(S_i^2) = \sigma^2$ and $V(S_i^2) = \frac{2\sigma^2}{n_i - 1}$ for i = 1, 2.

a. $E(S_p^2) = \dfrac{(n_1 - 1)E(S_1^2) + (n_2 - 1)E(S_2^2)}{n_1 + n_2 - 2} = \sigma^2$

b. $V(S_p^2) = \dfrac{(n_1 - 1)^2 V(S_1^2) + (n_2 - 1)^2 V(S_2^2)}{(n_1 + n_2 - 2)^2} = \dfrac{2\sigma^4}{n_1 + n_2 - 2}$.

8.134 The width of the small sample CI is $2t_{\alpha/2}\left(\frac{S}{\sqrt{n}}\right)$, and from Ex. 8.16 it was derived that $E(S) = \frac{\sigma}{\sqrt{n-1}}\frac{\sqrt{2}\Gamma(n/2)}{\Gamma[(n-1)/2]}$. Thus,

$$E\left(2t_{\alpha/2}\frac{S}{\sqrt{n}}\right) = 2^{3/2}t_{\alpha/2}\left(\frac{\sigma}{\sqrt{n(n-1)}}\right)\left(\frac{\Gamma(n/2)}{\Gamma[(n-1)/2]}\right).$$

8.135 The midpoint of the CI is given by $M = \frac{1}{2}\left(\frac{(n-1)S^2}{\chi^2_{1-\alpha/2}} + \frac{(n-1)S^2}{\chi^2_{\alpha/2}}\right)$. Therefore, since $E(S^2) = \sigma^2$, we have

$$E(M) = \frac{1}{2}\left(\frac{(n-1)\sigma^2}{\chi^2_{1-\alpha/2}} + \frac{(n-1)\sigma^2}{\chi^2_{\alpha/2}}\right) = \frac{(n-1)\sigma^2}{2}\left(\frac{1}{\chi^2_{1-\alpha/2}} + \frac{1}{\chi^2_{\alpha/2}}\right) \neq \sigma^2.$$

8.136 Consider the quantity $Y_p - \overline{Y}$. Since $Y_1, Y_2, \ldots, Y_n, Y_p$ are independent and identically distributed, we have that

$$E(Y_p - \overline{Y}) = \mu - \mu = 0$$
$$V(Y_p - \overline{Y}) = \sigma^2 + \sigma^2/n = \sigma^2\left(\frac{n+1}{n}\right).$$

Therefore, $Z = \dfrac{Y_p - \overline{Y}}{\sigma\sqrt{\frac{n+1}{n}}}$ has a standard normal distribution. So, by Definition 7.2,

$$\frac{\dfrac{Y_p - \overline{Y}}{\sigma\sqrt{\frac{n+1}{n}}}}{\sqrt{\dfrac{(n-1)S^2}{\sigma^2(n-1)}}} = \frac{Y_p - \overline{Y}}{S\sqrt{\frac{n+1}{n}}}$$

has a t–distribution with n – 1 degrees of freedom. Thus, by using the same techniques as used in Section 8.8, the *prediction interval* is

$$\overline{Y} \pm t_{\alpha/2}S\sqrt{\frac{n+1}{n}},$$

where $t_{\alpha/2}$ is the upper–$\alpha/2$ critical value from the t–distribution with $n - 1$ degrees of freedom.

Chapter 9: Properties of Point Estimators and Methods of Estimation

9.1 Refer to Ex. 8.8 where the variances of the four estimators were calculated. Thus,

$$\text{eff}(\hat{\theta}_1, \hat{\theta}_5) = 1/3 \qquad \text{eff}(\hat{\theta}_2, \hat{\theta}_5) = 2/3 \qquad \text{eff}(\hat{\theta}_3, \hat{\theta}_5) = 3/5.$$

9.2 **a.** The three estimators a unbiased since:

$$E(\hat{\mu}_1) = \tfrac{1}{2}\big(E(Y_1) + E(Y_2)\big) = \tfrac{1}{2}(\mu + \mu) = \mu$$

$$E(\hat{\mu}_2) = \mu/4 + \frac{(n-2)\mu}{2(n-2)} + \mu/4 = \mu$$

$$E(\hat{\mu}_3) = E(\overline{Y}) = \mu.$$

b. The variances of the three estimators are

$$V(\hat{\mu}_1) = \tfrac{1}{4}(\sigma^2 + \sigma^2) = \tfrac{1}{2}\sigma^2$$

$$V(\hat{\mu}_2) = \sigma^2/16 + \frac{(n-2)\sigma^2}{4(n-2)^2} + \sigma^2/16 = \sigma^2/8 + \frac{\sigma^2}{4(n-2)}$$

$$V(\hat{\mu}_3) = \sigma^2/n.$$

Thus, $\text{eff}(\hat{\mu}_3, \hat{\mu}_2) = \dfrac{n^2}{8(n-2)}$, $\text{eff}(\hat{\mu}_3\ \hat{\mu}_1) = n/2$.

9.3 **a.** $E(\hat{\theta}_1) = E(\overline{Y}) - 1/2 = \theta + 1/2 - 1/2 = \theta$. From Section 6.7, we can find the density function of $\hat{\theta}_2 = Y_{(n)}$: $g_n(y) = n(y-\theta)^{n-1}$, $\theta \le y \le \theta + 1$. From this, it is easily shown that $E(\hat{\theta}_2) = E(Y_{(n)}) - n/(n+1) = \theta$.

b. $V(\hat{\theta}_1) = V(\overline{Y}) = \sigma^2/n = 1/(12n)$. With the density in part **a**, $V(\hat{\theta}_2) = V(Y_{(n)}) = \frac{n}{(n+2)(n+1)^2}$. Thus, $\text{eff}(\hat{\theta}_1, \hat{\theta}_2) = \frac{12n^2}{(n+2)(n+1)^2}$.

9.4 See Exercises 8.18 and 6.74. Following those, we have that $V(\hat{\theta}_1) = (n+1)^2 V(Y_{(n)}) = \frac{n}{n+2}\theta^2$. Similarly, $V(\hat{\theta}_2) = \left(\frac{n+1}{n}\right)^2 V(Y_{(n)}) = \frac{1}{n(n+2)}\theta^2$. Thus, the ratio of these variances is as given.

9.5 From Ex. 7.20, we know S^2 is unbiased and $V(S^2) = V(\hat{\sigma}_1^2) = \frac{2\sigma^4}{n-1}$. For $\hat{\sigma}_2^2$, note that $Y_1 - Y_2$ is normal with mean 0 and variance σ^2. So, $\frac{(Y_1 - Y_2)^2}{2\sigma^2}$ is chi–square with one degree of freedom and $E(\hat{\sigma}_2^2) = \sigma^2$, $V(\hat{\sigma}_2^2) = 2\sigma^4$. Thus, we have that $\text{eff}(\hat{\sigma}_1^2, \hat{\sigma}_2^2) = n - 1$.

9.6 Both estimators are unbiased and $V(\hat{\lambda}_1) = \lambda/2$ and $V(\hat{\lambda}_2) = \lambda/n$. The efficiency is $2/n$.

9.7 The estimator $\hat{\theta}_1$ is unbiased so $\text{MSE}(\hat{\theta}_1) = V(\hat{\theta}_1) = \theta^2$. Also, $\hat{\theta}_2 = \overline{Y}$ is unbiased for θ (θ is the mean) and $V(\hat{\theta}_2) = \sigma^2/n = \theta^2/n$. Thus, we have that $\text{eff}(\hat{\theta}_1, \hat{\theta}_2) = 1/n$.

9.8 **a.** It is not difficult to show that $\frac{\partial^2 \ln f(y)}{\partial \mu^2} = -\frac{1}{\sigma^2}$, so $I(\mu) = \sigma^2/n$, Since $V(\overline{Y}) = \sigma^2/n$, $\overline{Y}$ is an efficient estimator of μ.

b. Similarly, $\frac{\partial^2 \ln p(y)}{\partial \lambda^2} = -\frac{y}{\lambda^2}$ and $E(-Y/\lambda^2) = 1/\lambda$. Thus, $I(\lambda) = \lambda/n$. By Ex. 9.6, $\overline{Y}$ is an efficient estimator of λ.

9.9 **a.** $X_6 = 1$.
b.-e. Answers vary.

9.10 **a.-b.** Answers vary.

9.11 **a.-b.** Answers vary.
c. The simulations are different but get close at $n = 50$.

9.12 **a.-b.** Answers vary.

9.13 **a.** Sequences are different but settle down at large *n*.
b. Sequences are different but settle down at large *n*.

9.14 **a.** the mean, 0.
b.-c. the variability of the estimator decreases with *n*.

9.15 Referring to Ex. 9.3, since both estimators are unbiased and the variances go to 0 with as *n* goes to infinity the estimators are consistent.

9.16 From Ex. 9.5, $V(\hat{\sigma}_2^2) = 2\sigma^4$ which is constant for all *n*. Thus, $\hat{\sigma}_2^2$ is not a consistent estimator.

9.17 In Example 9.2, it was shown that both $\overline{X}$ and $\overline{Y}$ are consistent estimators of μ_1 and μ_2, respectively. Using Theorem 9.2, $\overline{X} - \overline{Y}$ is a consistent estimator of $\mu_1 - \mu_2$.

9.18 Note that this estimator is the pooled sample variance estimator S_p^2 with $n_1 = n_2 = n$. In Ex. 8.133 it was shown that S_p^2 is an unbiased estimator. Also, it was shown that the variance of S_p^2 is $\frac{2\sigma^4}{n_1 + n_2 - 2} = \frac{\sigma^4}{n-1}$. Since this quantity goes to 0 with *n*, the estimator is consistent.

9.19 Given $f(y)$, we have that $E(Y) = \frac{\theta}{\theta+1}$ and $V(Y) = \frac{\theta}{(\theta+2)(\theta+1)^2}$ (Y has a beta distribution with parameters $\alpha = \theta$ and $\beta = 1$. Thus, $E(\overline{Y}) = \frac{\theta}{\theta+1}$ and $V(\overline{Y}) = \frac{\theta}{n(\theta+2)(\theta+1)^2}$. Thus, the conditions are satisfied for $\overline{Y}$ to be a consistent estimator.

9.20 Since $E(Y) = np$ and $V(Y) = npq$, we have that $E(Y/n) = p$ and $V(Y/n) = pq/n$. Thus, Y/n is consistent since it is unbiased and its variance goes to 0 with n.

9.21 Note that this is a generalization of Ex. 9.5. The estimator $\hat{\sigma}^2$ can be written as

$$\hat{\sigma}^2 = \frac{1}{k}\left[\frac{(Y_2 - Y_1)^2}{2} + \frac{(Y_4 - Y_3)^2}{2} + \frac{(Y_6 - Y_5)^2}{2} + \ldots + \frac{(Y_n - Y_{n-1})^2}{2}\right].$$

There are k independent terms in the sum, each with mean σ^2 and variance $2\sigma^4$.

a. From the above, $E(\hat{\sigma}^2) = (k\sigma^2)/k = \sigma^2$. So $\hat{\sigma}^2$ is an unbiased estimator.

b. Similarly, $V(\hat{\sigma}^2) = k(2\sigma^4)/k^2 = 2\sigma^4/k$. Since $k = n/2$, $V(\hat{\sigma}^2)$ goes to 0 with n and $\hat{\sigma}^2$ is a consistent estimator.

9.22 Following Ex. 9.21, we have that the estimator $\hat{\lambda}$ can be written as

$$\hat{\lambda} = \frac{1}{k}\left[\frac{(Y_2 - Y_1)^2}{2} + \frac{(Y_4 - Y_3)^2}{2} + \frac{(Y_6 - Y_5)^2}{2} + \ldots + \frac{(Y_n - Y_{n-1})^2}{2}\right].$$

For Y_i, Y_{i-1}, we have that:

$$\frac{E[(Y_i - Y_{i-1})^2]}{2} = \frac{E(Y_i^2) - 2E(Y_i)E(Y_{i-1}) + E(Y_{i-1}^2)}{2} = \frac{(\lambda + \lambda^2) - 2\lambda^2 + (\lambda + \lambda^2)}{2} = \lambda$$

$$\frac{V[(Y_i - Y_{i-1})^2]}{4} < \frac{V(Y_i^2) + V(Y_{i-1}^2)}{4} = \frac{2\lambda + 12\lambda^2 + 8\lambda^3}{4} = \gamma$$, since Y_i and Y_{i-1} are independent and non–negative (the calculation can be performed using the Poisson mgf).

a. From the above, $E(\hat{\lambda}) = (k\lambda)/k = \lambda$. So $\hat{\lambda}$ is an unbiased estimator of λ.

b. Similarly, $V(\hat{\lambda}) < k\gamma/k^2$, where $\gamma < \infty$ is defined above. Since $k = n/2$, $V(\hat{\lambda})$ goes to 0 with n and $\hat{\lambda}$ is a consistent estimator.

9.23 **a.** Note that for $i = 1, 2, \ldots, k$,

$$E(Y_{2i} - Y_{2i-1}) = 0 \qquad V(Y_{2i} - Y_{2i-1}) = 2\sigma^2 = E[(Y_{2i} - Y_{2i-1})^2].$$

Thus, it follows from methods used in Ex. 9.23 that $\hat{\sigma}^2$ is an unbiased estimator.

b. $V(\hat{\sigma}^2) = \frac{1}{4k^2}\sum_{i=1}^{k} V[(Y_{2i} - Y_{2i-1})^2] = \frac{1}{4k}V[(Y_2 - Y_1)^2]$, since the Y's are independent and identically distributed. Now, it is clear that $V[(Y_2 - Y_1)^2] \le E[(Y_2 - Y_1)^4]$, and when this quantity is expanded, only moments of order 4 or less are involved. Since these were assumed to be finite, $E[(Y_2 - Y_1)^4] < \infty$ and so $V(\hat{\sigma}^2) = \frac{1}{4k}V[(Y_2 - Y_1)^2] \to 0$ as $n \to \infty$.

c. This was discussed in part b.

9.24 **a.** From Chapter 6, $\sum_{i=1}^{n} Y_i^2$ is chi–square with with n degrees of freedom.

b. Note that $E(W_n) = 1$ and $V(W_n) = 1/n$. Thus, as $n \to \infty$, $W_n \to E(W_n) = 1$ in probability.

9.25 **a.** Since $E(Y_1) = \mu$, Y_1 is unbiased.

b. $P(|Y_1 - \mu| \le 1) = P(-1 \le Z \le 1) = .6826$.

c. The estimator is not consistent since the probability found in part b does not converge to unity (here, $n = 1$).

9.26 **a.** We have that $P(\theta - \varepsilon \le Y_{(n)} \le \theta + \varepsilon) = F_{(n)}(\theta + \varepsilon) - F_{(n)}(\theta - \varepsilon)$.

- If $\varepsilon > \theta$, $F_{(n)}(\theta + \varepsilon) = 1$ and $F_{(n)}(\theta - \varepsilon) = 0$. Thus, $P(\theta - \varepsilon \le Y_{(n)} \le \theta + \varepsilon) = 1$.
- If $\varepsilon < \theta$, $F_{(n)}(\theta + \varepsilon) = 1$, $F_{(n)}(\theta - \varepsilon) = \left(\frac{\theta-\varepsilon}{\theta}\right)^n$. So, $P(\theta - \varepsilon \le Y_{(n)} \le \theta + \varepsilon) = 1 - \left(\frac{\theta-\varepsilon}{\theta}\right)^n$.

b. The result follows from $\lim_{n\to\infty} P(\theta - \varepsilon \le Y_{(n)} \le \theta + \varepsilon) = \lim_{n\to\infty} \left[1 - \left(\frac{\theta-\varepsilon}{\theta}\right)^n\right] = 1$.

9.27 $P(|Y_{(1)} - \theta| \le \varepsilon) = P(\theta - \varepsilon \le Y_{(1)} \le \theta + \varepsilon) = F_{(1)}(\theta + \varepsilon) - F_{(1)}(\theta - \varepsilon) = 1 - \left(1 - \frac{\theta-\varepsilon}{\theta}\right)^n = \left(\frac{\varepsilon}{\theta}\right)^n$.

But, $\lim_{n\to\infty} \left(\frac{\varepsilon}{\theta}\right)^n = 0$ for $\varepsilon < \theta$. So, $Y_{(1)}$ is not consistent.

9.28 $P(|Y_{(1)} - \beta| \le \varepsilon) = P(\beta - \varepsilon \le Y_{(1)} \le \beta + \varepsilon) = F_{(1)}(\beta + \varepsilon) - F_{(1)}(\beta - \varepsilon) = 1 - \left(\frac{\beta}{\beta+\varepsilon}\right)^{\alpha n}$. Since $\lim_{n\to\infty} \left(\frac{\beta}{\beta+\varepsilon}\right)^{\alpha n} = 0$ for $\varepsilon > 0$, $Y_{(1)}$ is consistent.

9.29 $P(|Y_{(1)} - \theta| \le \varepsilon) = P(\theta - \varepsilon \le Y_{(1)} \le \theta + \varepsilon) = F_{(1)}(\theta + \varepsilon) - F_{(1)}(\theta - \varepsilon) = 1 - \left(\frac{\theta-\varepsilon}{\theta}\right)^{\alpha n}$. Since $\lim_{n\to\infty} \left(\frac{\theta-\varepsilon}{\theta}\right)^{\alpha n} = 0$ for $\varepsilon > 0$, $Y_{(1)}$ is consistent.

9.30 Note that Y is beta with $\mu = 3/4$ and $\sigma^2 = 3/5$. Thus, $E(\overline{Y}) = 3/4$ and $V(\overline{Y}) = 3/(5n)$. Thus, $V(\overline{Y}) \to 0$ and $\overline{Y}$ converges in probability to 3/4.

9.31 Since $\overline{Y}$ is a mean of independent and identically distributed random variables with finite variance, $\overline{Y}$ is consistent and $\overline{Y}$ converges in probability to $E(\overline{Y}) = E(Y) = \alpha\beta$.

9.32 Notice that $E(Y^2) = \int_2^\infty y^2 \frac{2}{y^2} dy = \int_2^\infty 2 dy = \infty$, thus $V(Y) = \infty$ and so the law of large numbers does not apply.

9.33 By the law of large numbers, $\overline{X}$ and $\overline{Y}$ are consistent estimators of λ_1 and λ_2. By Theorem 9.2, $\frac{\overline{X}}{\overline{X}+\overline{Y}}$ converges in probability to $\frac{\lambda_1}{\lambda_1+\lambda_2}$. This implies that observed values of the estimator should be close to the limiting value for large sample sizes, although the variance of this estimator should also be taken into consideration.

9.34 Following Ex. 6.34, Y^2 has an exponential distribution with parameter θ. Thus, $E(Y^2) = \theta$ and $V(Y^2) = \theta^2$. Therefore, $E(W_n) = \theta$ and $V(W_n) = \theta^2/n$. Clearly, W_n is a consistent estimator of θ.

9.35 **a.** $E(\overline{Y}_n) = \frac{1}{n}(\mu + \mu + \cdots + \mu) = \mu$, so $\overline{Y}_n$ is unbiased for μ.

b. $V(\overline{Y}_n) = \frac{1}{n^2}(\sigma_1^2 + \sigma_2^2 + \cdots + \sigma_n^2) = \frac{1}{n^2}\sum_{i=1}^n \sigma_i^2$.

c. In order for $\overline{Y}_n$ to be consistent, it is required that $V(\overline{Y}_n) \to 0$ as $n \to \infty$. Thus, it must be true that all variances must be finite, or simply $\max_i\{\sigma_i^2\} < \infty$.

9.36 Let $X_1, X_2, \ldots, X_n$ be a sequence of Bernoulli trials with success probability p. Thus, it is seen that $Y = \sum_{i=1}^n X_i$. Thus, by the Central Limit Theorem, $U_n = \dfrac{\hat{p}_n - p}{\sqrt{\dfrac{pq}{n}}}$ has a limiting standard normal distribution. By Ex. 9.20, it was shown that $\hat{p}_n$ is consistent for p, so it makes sense that $\hat{q}_n$ is consistent for q, and so by Theorem 9.2 $\hat{p}_n\hat{q}_n$ is consistent for pq. Define $W_n = \sqrt{\dfrac{\hat{p}_n\hat{q}_n}{pq}}$ so that W_n converges in probability to 1. By Theorem 9.3, the quantity $\dfrac{U_n}{W_n} = \dfrac{\hat{p}_n - p}{\sqrt{\dfrac{\hat{p}_n\hat{q}_n}{n}}}$ converges to a standard normal variable.

9.37 The likelihood function is $L(p) = p^{\Sigma x_i}(1-p)^{n-\Sigma x_i}$. By Theorem 9.4, $\sum_{i=1}^n X_i$ is sufficient for p with $g(\Sigma x_i, p) = p^{\Sigma x_i}(1-p)^{n-\Sigma x_i}$ and $h(\boldsymbol{y}) = 1$.

9.38 For this exercise, the likelihood function is given by

$$L = \frac{1}{(2\pi)^{n/2}\sigma^n}\exp\left[-\frac{\sum_{i=1}^n (y_i-\mu)^2}{2\sigma^2}\right] = (2\pi)^{-n/2}\sigma^{-n}\exp\left[\frac{-1}{2\sigma^2}\left(\sum_{i=1}^n y_i^2 - 2\mu n\overline{y} + n\mu^2\right)\right].$$

a. When σ^2 is known, $\overline{Y}$ is sufficient for μ by Theorem 9.4 with

$g(\overline{y},\mu) = \exp\left(\dfrac{2\mu n\overline{y} - n\mu^2}{2\sigma^2}\right)$ and $h(\boldsymbol{y}) = (2\pi)^{-n/2}\sigma^{-n}\exp\left(\dfrac{-1}{2\sigma^2}\sum_{i=1}^n y_i^2\right)$.

b. When μ is known, use Theorem 9.4 with

$g(\sum_{i=1}^n (y_i-\mu)^2, \sigma^2) = (\sigma^2)^{-n/2}\exp\left[-\dfrac{\sum_{i=1}^n (y_i-\mu)^2}{2\sigma^2}\right]$ and $h(\boldsymbol{y}) = (2\pi)^{-n/2}$.

c. When both μ and σ^2 are unknown, the likelihood can be written in terms of the two statistics $U_1 = \sum_{i=1}^n Y_i$ and $U_2 = \sum_{i=1}^n Y_i^2$ with $h(\boldsymbol{y}) = (2\pi)^{-n/2}$. The statistics $\overline{Y}$ and S^2 are also jointly sufficient since they can be written in terms of U_1 and U_2.

9.39 Note that by independence, $U = \sum_{i=1}^{n} Y_i$ has a Poisson distribution with parameter $n\lambda$. Thus, the conditional distribution is expressed as

$$P(Y_1 = y_1, \ldots, Y_n = y_n \mid U = u) = \frac{P(Y_1 = y_1, \ldots, Y_n = y_n)}{P(U = u)} = \frac{\prod_{i=1}^{n} \dfrac{\lambda^{y_i} e^{-\lambda}}{y_i!}}{\dfrac{(n\lambda)^u e^{-n\lambda}}{u!}} = \frac{\dfrac{\lambda^{\Sigma y_i} e^{-n\lambda}}{\Pi y_i!}}{\dfrac{(n\lambda)^u e^{-n\lambda}}{u!}}.$$

We have that $\Sigma y_i = u$, so the above simplifies to

$$P(Y_1 = y_1, \ldots, Y_n = y_n \mid U = u) = \begin{cases} \dfrac{u!}{n^u \Pi y_i!} & \text{if } \Sigma y_i = u \\ 0 & \text{otherwise} \end{cases}.$$

Since the conditional distribution is free of λ, the statistic $U = \sum_{i=1}^{n} Y_i$ is sufficient for λ.

9.40 The likelihood is $L(\theta) = 2^n \theta^{-n} \prod_{i=1}^{n} y_i \exp\left(-\sum_{i=1}^{n} y_i^2 / \theta\right)$. By Theorem 9.4, $U = \sum_{i=1}^{n} Y_i^2$ is sufficient for θ with $g(u, \theta) = \theta^{-n} \exp(-u/\theta)$ and $h(\mathbf{y}) = 2^n \prod_{i=1}^{n} y_i$.

9.41 The likelihood is $L(\alpha) = \alpha^{-n} m^n \left(\prod_{i=1}^{n} y_i\right)^{m-1} \exp\left(-\sum_{i=1}^{n} y_i^m / \alpha\right)$. By Theorem 9.4, $U = \sum_{i=1}^{n} Y_i^m$ is sufficient for α with $g(u, \alpha) = \alpha^{-n} \exp(-u/\alpha)$ and $h(\mathbf{y}) = m^n \left(\prod_{i=1}^{n} y_i\right)^{m-1}$.

9.42 The likelihood function is $L(p) = p^n (1-p)^{\Sigma y_i - n} = p^n (1-p)^{n\bar{y} - n}$. By Theorem 9.4, $\bar{Y}$ is sufficient for p with $g(\bar{y}, p) = p^n (1-p)^{n\bar{y} - n}$ and $h(\mathbf{y}) = 1$.

9.43 With θ known, the likelihood is $L(\alpha) = \alpha^n \theta^{-n\alpha} \left(\prod_{i=1}^{n} y_i\right)^{\alpha-1}$. By Theorem 9.4, $U = \prod_{i=1}^{n} Y_i$ is sufficient for α with $g(u, \alpha) = \alpha^n \theta^{-n\alpha} \left(\prod_{i=1}^{n} y_i\right)^{\alpha-1}$ and $h(\mathbf{y}) = 1$.

9.44 With β known, the likelihood is $L(\alpha) = \alpha^n \beta^{n\alpha} \left(\prod_{i=1}^{n} y_i\right)^{-(\alpha+1)}$. By Theorem 9.4, $U = \prod_{i=1}^{n} Y_i$ is sufficient for α with $g(u, \alpha) = \alpha^n \beta^{n\alpha} (u)^{-(\alpha+1)}$ and $h(\mathbf{y}) = 1$.

9.45 The likelihood function is

$$L(\theta) = \prod_{i=1}^{n} f(y_i \mid \theta) = [a(\theta)]^n \left[\prod_{i=1}^{n} b(y_i)\right] \exp\left[-c(\theta) \sum_{i=1}^{n} d(y_i)\right].$$

Thus, $U = \sum_{i=1}^{n} d(Y_i)$ is sufficient for θ because by Theorem 9.4 $L(\theta)$ can be factored into, where $u = \sum_{i=1}^{n} d(y_i)$, $g(u, \theta) = [a(\theta)]^n \exp[-c(\theta) u]$ and $h(\mathbf{y}) = \prod_{i=1}^{n} b(y_i)$.

9.46 The exponential distribution is in exponential form since $a(\beta) = c(\beta) = 1/\beta$, $b(y) = 1$, and $d(y) = y$. Thus, by Ex. 9.45, $\sum_{i=1}^{n} Y_i$ is sufficient for β, and then so is $\bar{Y}$.

9.47 We can write the density function as $f(y \mid \alpha) = \alpha\theta^{\alpha} \exp[-(\alpha-1)\ln y]$. Thus, the density has exponential form and the sufficient statistic is $\sum_{i=1}^{n} \ln\left(\overline{Y}_i\right)$. Since this is equivalently expressed as $\ln\left(\prod_{i=1}^{n} Y_i\right)$, we have no contradiction with Ex. 9.43.

9.48 We can write the density function as $f(y \mid \alpha) = \alpha\beta^{\alpha} \exp[-(\alpha+1)\ln y]$. Thus, the density has exponential form and the sufficient statistic is $\sum_{i=1}^{n} \ln Y_i$. Since this is equivalently expressed as $\ln \prod_{i=1}^{n} Y_i$, we have no contradiction with Ex. 9.44.

9.49 The density for the uniform distribution on (0, θ) is $f(y \mid \theta) = \frac{1}{\theta}$, $0 \le y \le \theta$. For this problem and several of the following problems, we will use an <u>indicator function</u> to specify the support of y. This is given by, in general, for $a < b$,

$$I_{a,b}(y) = \begin{cases} 1 & \text{if } a \le y \le b \\ 0 & \text{otherwise} \end{cases}.$$

Thus, the previously mentioned uniform distribution can be expressed as

$$f(y \mid \theta) = \frac{1}{\theta} I_{0,\theta}(y).$$

The likelihood function is given by $L(\theta) = \frac{1}{\theta^n} \prod_{i=1}^{n} I_{0,\theta}(y_i) = \frac{1}{\theta^n} I_{0,\theta}(y_{(n)})$, since $\prod_{i=1}^{n} I_{0,\theta}(y_i) = I_{0,\theta}(y_{(n)})$. Therefore, Theorem 9.4 is satisfied with $h(\boldsymbol{y}) = 1$ and $g(y_{(n)}, \theta) = \frac{1}{\theta^n} I_{0,\theta}(y_{(n)})$.

(This problem could also be solved using the conditional distribution definition of sufficiency.)

9.50 As in Ex. 9.49, we will define the uniform distribution on the interval (θ_1, θ_2) as

$$f(y \mid \theta_1, \theta_2) = \frac{1}{(\theta_2 - \theta_1)} I_{\theta_1,\theta_2}(y).$$

The likelihood function, using the same logic as in Ex. 9.49, is

$$L(\theta_1, \theta_2) = \frac{1}{(\theta_2 - \theta_1)^n} \prod_{i=1}^{n} I_{\theta_1,\theta_2}(y_i) = \frac{1}{(\theta_2 - \theta_1)^n} I_{\theta_1,\theta_2}(y_{(1)}) I_{\theta_1,\theta_2}(y_{(n)}).$$

So, Theorem 9.4 is satisfied with $g(y_{(1)}, y_{(n)}, \theta_1, \theta_2) = \frac{1}{(\theta_2 - \theta_1)^n} I_{\theta_1,\theta_2}(y_{(1)}) I_{\theta_1,\theta_2}(y_{(n)})$ and $h(\boldsymbol{y}) = 1$.

9.51 Again, using the indicator notation, the density is

$$f(y \mid \theta) = \exp[-(y - \theta)] I_{a,\infty}(y)$$

(it should be obvious that $y < \infty$ for the indicator function). The likelihood function is

$$L(\theta) = \exp\left(-\sum_{i=1}^{n} y + n\theta\right)\prod_{i=1}^{n} I_{a,\infty}(y_i) = \exp\left(-\sum_{i=1}^{n} y + n\theta\right) I_{a,\infty}(y_{(1)}).$$

Theorem 9.4 is satisfied with $g(y_{(1)},\theta) = \exp(n\theta) I_{a,\infty}(y_{(1)})$ and $h(\boldsymbol{y}) = \exp\left(-\sum_{i=1}^{n} y\right)$.

9.52 Again, using the indicator notation, the density is

$$f(y\,|\,\theta) = \frac{3y^2}{\theta^3} I_{0,\theta}(y).$$

The likelihood function is $L(\theta) = \dfrac{3^n \prod_{i=1}^{n} y_i^2}{\theta^{3n}} \prod_{i=1}^{n} I_{0,\theta}(y_i) = \dfrac{3^n \prod_{i=1}^{n} y_i^2}{\theta^{3n}} I_{0,\theta}(y_{(n)})$. Then, Theorem 9.4 is satisfied with $g(y_{(n)},\theta) = \theta^{-3n} I_{0,\theta}(y_{(n)})$ and $h(\boldsymbol{y}) = 3^n \prod_{i=1}^{n} y_i^2$.

9.53 Again, using the indicator notation, the density is

$$f(y\,|\,\theta) = \frac{2\theta^2}{y^3} I_{\theta,\infty}(y).$$

The likelihood function is $L(\theta) = 2^n \theta^{2n}\left(\prod_{i=1}^{n} y_i^{-3}\right)\prod_{i=1}^{n} I_{\theta,\infty}(y_i) = 2^n \theta^{2n}\left(\prod_{i=1}^{n} y_i^{-3}\right) I_{\theta,\infty}(y_{(1)})$

Theorem 9.4 is satisfied with $g(y_{(1)},\theta) = \theta^{2n} I_{\theta,\infty}(y_{(1)})$ and $h(\boldsymbol{y}) = 2^n\left(\prod_{i=1}^{n} y_i^{-3}\right)$.

9.54 Again, using the indicator notation, the density is

$$f(y\,|\,\alpha,\theta) = \alpha\theta^{-\alpha} y^{\alpha-1} I_{0,\theta}(y).$$

The likelihood function is

$$L(\alpha,\theta) = \alpha^n \theta^{-n\alpha}\left(\prod_{i=1}^{n} y_i\right)^{\alpha-1} \prod_{i=1}^{n} I_{0,\theta}(y_i) = \alpha^n \theta^{-n\alpha}\left(\prod_{i=1}^{n} y_i\right)^{\alpha-1} I_{0,\theta}(y_{(n)}).$$

Theorem 9.4 is satisfied with $g(\prod_{i=1}^{n} y_i, y_{(n)}, \alpha, \theta) = \alpha^n \theta^{-n\alpha}\left(\prod_{i=1}^{n} y_i\right)^{\alpha-1} I_{0,\theta}(y_{(n)})$, $h(\boldsymbol{y}) = 1$ so that $\left(\prod_{i=1}^{n} Y_i, Y_{(n)}\right)$ is jointly sufficient for α and θ.

9.55 Lastly, using the indicator notation, the density is

$$f(y\,|\,\alpha,\beta) = \alpha\beta^{\alpha} y^{-(\alpha+1)} I_{\beta,\infty}(y).$$

The likelihood function is

$$L(\alpha,\beta) = \alpha^n \beta^{n\alpha}\left(\prod_{i=1}^{n} y_i^{-(\alpha+1)}\right)\prod_{i=1}^{n} I_{\beta,\infty}(y_i) = \alpha^n \beta^{n\alpha}\left(\prod_{i=1}^{n} y_i^{-(\alpha+1)}\right) I_{\beta,\infty}(y_{(1)}).$$

Theorem 9.4 is satisfied with $g(\prod_{i=1}^{n} y_i, y_{(1)}, \alpha, \beta) = \alpha^n \beta^{n\alpha}\left(\prod_{i=1}^{n} y_i^{-(\alpha+1)}\right) I_{\beta,\infty}(y_{(1)})$, and $h(\boldsymbol{y}) = 1$ so that $\left(\prod_{i=1}^{n} Y_i, Y_{(1)}\right)$ is jointly sufficient for α and β.

9.56 In Ex. 9.38 (b), it was shown that $\sum_{i=1}^{n} (y_i - \mu)^2$ is sufficient for σ^2. Since the quantity $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} (y_i - \mu)^2$ is unbiased and a function of the sufficient statistic, it is the MVUE of σ^2.

9.57 Note that the estimator can be written as

$$\hat{\sigma}^2 = \frac{S_X^2 + S_Y^2}{2},$$

where $S_X^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X})^2$, $S_Y^2 = \frac{1}{n-1}\sum_{i=1}^n (Y_i - \overline{Y})^2$. Since both of these estimators are the MVUE (see Example 9.8) for σ^2 and E($\hat{\sigma}^2$) = σ^2, $\hat{\sigma}^2$ is the MVUE for σ^2.

9.58 From Ex. 9.34 and 9.40, $\sum_{i=1}^n Y_i^2$ is sufficient for θ and $E(Y^2) = \theta$. Thus, the MVUE is $\hat{\theta} = \frac{1}{n}\sum_{i=1}^n Y_i^2$.

9.59 Note that $E(C) = E(3Y^2) = 3E(Y^2) = 3[V(Y) + (E(Y))^2] = 3(\lambda + \lambda^2)$. Now, from Ex. 9.39, it was determined that $\sum_{i=1}^n Y_i$ is sufficient for λ, so if an estimator can be found that is unbiased for $3(\lambda + \lambda^2)$ and a function of the sufficient statistic, it is the MVUE. Note that $\sum_{i=1}^n Y_i$ is Poisson with parameter $n\lambda$, so

$$E(\overline{Y}^2) = V(\overline{Y}) + [E(\overline{Y})]^2 = \tfrac{\lambda}{n} + \lambda^2, \text{ and}$$
$$E(\overline{Y}/n) = \lambda/n.$$

Thus $\lambda^2 = E(\overline{Y}^2) - E(\overline{Y}/n)$ so that the MVUE for $3(\lambda + \lambda^2)$ is

$$3\left[\overline{Y}^2 - \overline{Y}/n + \overline{Y}\right] = 3\left[\overline{Y}^2 + \overline{Y}\left(1 - \tfrac{1}{n}\right)\right].$$

9.60 **a.** The density can be expressed as $f(y\,|\,\theta) = \theta\exp[(\theta - 1)\ln y]$. Thus, the density has exponential form and $-\sum_{i=1}^n \ln y_i$ is sufficient for θ.

b. Let $W = -\ln Y$. The distribution function for W is

$$F_W(w) = P(W \le w) = P(-\ln Y \le w) = 1 - P(Y \le e^{-w}) = 1 - \int_0^{e^{-w}} \theta y^{\theta-1}dy = 1 - e^{-\theta w},\ w > 0.$$

This is the exponential distribution function with mean 1/θ.

c. For the transformation $U = 2\theta W$, the distribution function for U is

$$F_U(u) = P(U \le u) = P(2\theta W \le u) = P(W \le \tfrac{u}{2\theta}) = F_W(\tfrac{u}{2\theta}) = 1 - e^{-u/2},\ u > 0.$$

Note that this is the exponential distribution with mean 2, but this is equivalent to the chi–square distribution with 2 degrees of freedom. Therefore, by property of independent chi–square variables, $2\theta\sum_{i=1}^n W_i$ is chi–square with $2n$ degrees of freedom.

d. From Ex. 4.112, the expression for the expected value of the reciprocal of a chi–square variable is given. Thus, it follows that $E\left[\left(2\theta\sum_{i=1}^n W_i\right)^{-1}\right] = \dfrac{1}{2n-2} = \dfrac{1}{2(n-1)}$.

e. From part d, $\dfrac{n-1}{\sum_{i=1}^n W_i} = \dfrac{n-1}{-\sum_{i=1}^n \ln Y_i}$ is unbiased and thus the MVUE for θ.

9.61 It has been shown that $Y_{(n)}$ is sufficient for θ and $E(Y_{(n)}) = \left(\frac{n}{n+1}\right)\theta$. Thus, $\left(\frac{n+1}{n}\right)Y_{(n)}$ is the MVUE for θ.

9.62 Calculate $E(Y_{(1)}) = \int_{\theta}^{\infty} nye^{-n(y-\theta)}dy = \int_{0}^{\infty} n(u+\theta)e^{-nu}du = \theta + \frac{1}{n}$. Thus, $Y_{(1)} - \frac{1}{n}$ is the MVUE for θ.

9.63 **a.** The distribution function for Y is $F(y) = y^3/\theta^3$, $0 \le y \le \theta$. So, the density function for $Y_{(n)}$ is $f_{(n)}(y) = n[F(y)]^{n-1}f(y) = 3ny^{3n-1}/\theta^{3n}$, $0 \le y \le \theta$.

b. From part **a**, it can be shown that $E(Y_{(n)}) = \frac{3n}{3n+1}\theta$. Since $Y_{(n)}$ is sufficient for θ, $\frac{3n+1}{3n}Y_{(n)}$ is the MVUE for θ.

9.64 **a.** From Ex. 9.38, $\overline{Y}$ is sufficient for μ. Also, since σ = 1, $\overline{Y}$ has a normal distribution with mean μ and variance $1/n$. Thus, $E(\overline{Y}^2) = V(\overline{Y}) + [E(\overline{Y})]^2 = 1/n + \mu^2$. Therefore, the MVUE for μ^2 is $\overline{Y}^2 - 1/n$.

b. $V(\overline{Y}^2 - 1/n) = V(\overline{Y}^2) = E(\overline{Y}^4) - [E(\overline{Y}^2)]^2 = E(\overline{Y}^4) - [1/n + \mu^2]^2$. It can be shown that $E(\overline{Y}^4) = \frac{3}{n^2} + \frac{6\mu^2}{n} + \mu^4$ (the mgf for $\overline{Y}$ can be used) so that

$$V(\overline{Y}^2 - 1/n) = \frac{3}{n^2} + \frac{6\mu^2}{n} + \mu^4 -- [1/n + \mu^2]^2 = (2 + 4n\mu^2)/n^2.$$

9.65 **a.** $E(T) = P(T=1) = P(Y_1 = 1, Y_2 = 0) = P(Y_1 = 1)P(Y_2 = 0) = p(1-p)$.

b. $P(T = 1 \mid W = w) = \dfrac{P(Y_1 = 1, Y_2 = 0, W = w)}{P(W = w)} = \dfrac{P(Y_1 = 1, Y_2 = 0, \sum_{i=3}^{n} Y_i = w-1)}{P(W = w)}$

$$= \frac{P(Y_1 = 1)P(Y_2 = 0)P(\sum_{i=3}^{n} Y_i = w-1)}{P(W = w)} = \frac{p(1-p)\binom{n-2}{w-1}p^{w-1}(1-p)^{n-(w-1)}}{\binom{n}{w}p^w(1-p)^{n-w}}$$

$$= \frac{w(n-w)}{n(n-1)}.$$

c. $E(T \mid W) = P(T = 1 \mid W) = \dfrac{W}{n}\left(\dfrac{n-W}{n-1}\right) = \left(\dfrac{n}{n-1}\right)\dfrac{W}{n}\left(1 - \dfrac{W}{n}\right)$. Since T is unbiased by part (a) above and W is sufficient for p and so also for $p(1-p)$, $n\overline{Y}(1-\overline{Y})/(n-1)$ is the MVUE for $p(1-p)$.

9.66 **a.** **i.** The ratio of the likelihoods is given by

$$\frac{L(\mathbf{x} \mid p)}{L(\mathbf{y} \mid p)} = \frac{p^{\Sigma x_i}(1-p)^{n-\Sigma x_i}}{p^{\Sigma y_i}(1-p)^{n-\Sigma y_i}} = \frac{p^{\Sigma x_i}(1-p)^{-\Sigma x_i}}{p^{\Sigma y_i}(1-p)^{-\Sigma y_i}} = \left(\frac{p}{1-p}\right)^{\Sigma x_i - \Sigma y_i}$$

ii. If $\Sigma x_i = \Sigma y_i$, the ratio is 1 and free of p. Otherwise, it will not be free of p.

iii. From the above, it must be that $g(Y_1,\ldots,Y_n) = \sum_{i=1}^{n} Y_i$ is the minimal sufficient statistic for p. This is the same as in Example 9.6.

b. **i.** The ratio of the likelihoods is given by

$$\frac{L(\boldsymbol{x}\mid\theta)}{L(\boldsymbol{y}\mid\theta)} = \frac{2^n(\prod_{i=1}^{n} x_i)\theta^{-n}\exp(-\sum_{i=1}^{n} x_i^2/\theta)}{2^n(\prod_{i=1}^{n} y_i)\theta^{-n}\exp(-\sum_{i=1}^{n} y_i^2/\theta)} = \frac{\prod_{i=1}^{n} x_i}{\prod_{i=1}^{n} y_i}\exp\left[-\frac{1}{\theta}\left(\sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} y_i^2\right)\right]$$

ii. The above likelihood ratio will only be free of θ if $\sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} y_i^2$, so that $\sum_{i=1}^{n} Y_i^2$ is a minimal sufficient statistic for θ.

9.67 The likelihood is given by

$$L(\boldsymbol{y}\mid\mu,\sigma^2) = \frac{1}{(2\pi)^{n/2}\sigma^n}\exp\left[-\frac{\sum_{i=1}^{n}(y_i-\mu)^2}{2\sigma^2}\right].$$

The ratio of the likelihoods is

$$\frac{L(\boldsymbol{x}\mid\mu,\sigma^2)}{L(\boldsymbol{y}\mid\mu,\sigma^2)} = \exp\left\{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{n}(x_i-\mu)^2 - \sum_{i=1}^{n}(y_i-\mu)^2\right]\right\} =$$

$$\exp\left\{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} y_i^2 - 2\mu\left(\sum_{i=1}^{n} x_i - \sum_{i=1}^{n} y_i\right)\right]\right\}.$$

This ratio is free of (μ, σ^2) only if both $\sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} y_i^2$ and $\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i$, so $\sum_{i=1}^{n} Y_i$ and $\sum_{i=1}^{n} Y_i^2$ form jointly minimal sufficient statistics for μ and σ^2.

9.68 For unbiased estimators $g_1(U)$ and $g_2(U)$, whose values only depend on the data through the sufficient statistic U, we have that $E[g_1(U) - g_2(U)] = 0$. Since the density for U is complete, $g_1(U) - g_2(U) \equiv 0$ by definition so that $g_1(U) = g_2(U)$. Therefore, there is only one unbiased estimator for θ based on U, and it must also be the MVUE.

9.69 It is easy to show that $\mu = \frac{\theta+1}{\theta+2}$ so that $\theta = \frac{2\mu-1}{1-\mu}$. Thus, the MOM estimator is $\hat{\theta} = \frac{2\overline{Y}-1}{1-\overline{Y}}$. Since $\overline{Y}$ is a consistent estimator of μ, by the Law of Large Numbers $\hat{\theta}$ converges in probability to θ. However, this estimator is not a function of the sufficient statistic so it can't be the MVUE.

9.70 Since μ = λ, the MOM estimator of λ is $\hat{\lambda} = m_1' = \overline{Y}$.

9.71 Since $E(Y) = \mu_1' = 0$ and $E(Y^2) = \mu_2' = V(Y) = \sigma^2$, we have that $\hat{\sigma}^2 = m_2' = \frac{1}{n}\sum_{i=1}^{n} Y_i^2$.

9.72 Here, we have that $\mu_1' = \mu$ and $\mu_2' = \sigma^2 + \mu^2$. Thus, $\hat{\mu} = m_1' = \overline{Y}$ and $\hat{\sigma}^2 = m_2' - \overline{Y}^2 = \frac{1}{n}\sum_{i=1}^{n} Y_i^2 - \overline{Y}^2 = \frac{1}{n}\sum_{i=1}^{2}(Y_i - \overline{Y})^2$.

9.73 Note that our sole observation Y is hypergeometric such that $E(Y) = n\theta/N$. Thus, the MOM estimator of θ is $\hat{\theta} = NY/n$.

9.74 **a.** First, calculate $\mu_1' = E(Y) = \int_0^\theta 2y(\theta - y)/\theta^2 dy = \theta/3$. Thus, the MOM estimator of θ is $\hat{\theta} = 3\overline{Y}$.

b. The likelihood is $L(\theta) = 2^n \theta^{-2n} \prod_{i=1}^{n}(\theta - y_i)$. Clearly, the likelihood can't be factored into a function that only depends on $\overline{Y}$, so the MOM is not a sufficient statistic for θ.

9.75 The density given is a beta density with $\alpha = \beta = \theta$. Thus, $\mu_1' = E(Y) = .5$. Since this doesn't depend on θ, we turn to $\mu_2' = E(Y^2) = \frac{\theta+1}{2(2\theta+1)}$ (see Ex. 4.200). Hence, with $m_2' = \frac{1}{n}\sum_{i=1}^{n} Y_i^2$, the MOM estimator of θ is $\hat{\theta} = \frac{1-2m_2'}{4m_2'-1}$.

9.76 Note that $\mu_1' = E(Y) = 1/p$. Thus, the MOM estimator of p is $\hat{p} = 1/Y$.

9.77 Here, $\mu_1' = E(Y) = \frac{3}{2}\theta$. So, the MOM estimator of θ is $\hat{\theta} = \frac{2}{3}\overline{Y}$.

9.78 For Y following the given power family distribution,

$$E(Y) = \int_0^3 \alpha y^\alpha 3^{-\alpha} dy = \alpha 3^{-\alpha} \left.\frac{y^{\alpha+1}}{\alpha+1}\right|_0^3 = \frac{3\alpha}{\alpha+1}.$$

Thus, the MOM estimator of θ is $\hat{\theta} = \frac{\overline{Y}}{3-\overline{Y}}$.

9.79 For Y following the given Pareto distribution,

$$E(Y) = \int_\beta^\infty \alpha\beta^\alpha y^{-\alpha} dy = \alpha\beta^\alpha \left.\frac{y^{-\alpha+1}}{-\alpha+1}\right|_\beta^\infty = \alpha\beta/(\alpha - 1).$$

The mean is not defined if $\alpha < 1$. Thus, a generalized MOM estimator for α cannot be expressed.

9.80 **a.** The MLE is easily found to be $\hat{\lambda} = \overline{Y}$.
b. $E(\hat{\lambda}) = \lambda$, $V(\hat{\lambda}) = \lambda/n$.
c. Since $\hat{\lambda}$ is unbiased and has a variance that goes to 0 with increasing n, it is consistent.
d. By the invariance property, the MLE for $P(Y = 0)$ is $\exp(-\hat{\lambda})$.

9.81 The MLE is $\hat{\theta} = \overline{Y}$. By the invariance property of MLEs, the MLE of θ^2 is $\overline{Y}^2$.

9.82 The likelihood function is $L(\theta) = \theta^{-n} r^n \left(\prod_{i=1}^n y_i\right)^{r-1} \exp\left(-\sum_{i=1}^n y_i^r / \theta\right)$.

a. By Theorem 9.4, a sufficient statistic for θ is $\sum_{i=1}^n Y_i^r$.

b. The log–likelihood is

$$\ln L(\theta) = -n\ln\theta + n\ln r + (r-1)\ln\left(\prod_{i=1}^n y_i\right) - \sum_{i=1}^n y_i^r / \theta .$$

By taking a derivative w.r.t. θ and equating to 0, we find $\hat{\theta} = \frac{1}{n}\sum_{i=1}^n Y_i^r$.

c. Note that $\hat{\theta}$ is a function of the sufficient statistic. Since it is easily shown that $E(Y^r) = \theta$, $\hat{\theta}$ is then unbiased and the MVUE for θ.

9.83 **a.** The likelihood function is $L(\theta) = (2\theta+1)^{-n}$. Let $\gamma = \gamma(\theta) = 2\theta + 1$. Then, the likelihood can be expressed as $L(\gamma) = \gamma^{-n}$. The likelihood is maximized for small values of γ. The smallest value that can safely maximize the likelihood (see Example 9.16) without violating the support is $\hat{\gamma} = Y_{(n)}$. Thus, by the invariance property of MLEs,

$$\hat{\theta} = \tfrac{1}{2}\left(Y_{(n)} - 1\right).$$

b. Since $V(Y) = \frac{(2\theta+1)^2}{12}$. By the invariance principle, the MLE is $\left(Y_{(n)}\right)^2/12$.

9.84 This exercise is a special case of Ex. 9.85, so we will refer to those results.

a. The MLE is $\hat{\theta} = \overline{Y}/2$, so the maximum likelihood estimate is $\overline{y}/2 = 63$.

b. $E(\hat{\theta}) = \theta$, $V(\hat{\theta}) = V(\overline{Y}/2) = \theta^2/6$.

c. The bound on the error of estimation is $2\sqrt{V(\hat{\theta})} = 2\sqrt{(130)^2/6} = 106.14$.

d. Note that $V(Y) = 2\theta^2 = 2(130)^2$. Thus, the MLE for $V(Y) = 2(\hat{\theta})^2$.

9.85 **a.** For $\alpha > 0$ known the likelihood function is

$$L(\theta) = \frac{1}{[\Gamma(\alpha)]^n \theta^{n\alpha}} \left(\prod_{i=1}^n y_i\right)^{\alpha-1} \exp\left(-\sum_{i=1}^n y_i / \theta\right).$$

The log–likelihood is then

$$\ln L(\theta) = -n\ln[\Gamma(\alpha)] - n\alpha\ln\theta + (\alpha-1)\ln\left(\prod_{i=1}^n y_i\right) - \sum_{i=1}^n y_i/\theta$$

so that

$$\tfrac{d}{d\theta}\ln L(\theta) = -n\alpha/\theta + \sum_{i=1}^n y_i/\theta^2 .$$

Equating this to 0 and solving for θ, we find the MLE of θ to be

$$\hat{\theta} = \tfrac{1}{n\alpha}\sum_{i=1}^n Y_i = \tfrac{1}{\alpha}\overline{Y} .$$

b. Since $E(Y) = \alpha\theta$ and $V(Y) = \alpha\theta^2$, $E(\hat{\theta}) = \theta$, $V(\hat{\theta}) = \theta^2/(n\alpha)$.

c. Since $\overline{Y}$ is a consistent estimator of $\mu = \alpha\theta$, it is clear that $\hat{\theta}$ must be consistent for θ.

d. From the likelihood function, it is seen from Theorem 9.4 that $U = \sum_{i=1}^{n} Y_i$ is a sufficient statistic for θ. Since the gamma distribution is in the exponential family of distributions, U is also the minimal sufficient statistic.

e. Note that U has a gamma distribution with shape parameter $n\alpha$ and scale parameter θ. The distribution of $2U/\theta$ is chi–square with $2n\alpha$ degrees of freedom. With $n = 5$, $\alpha = 2$, $2U/\theta$ is chi–square with 20 degrees of freedom. So, with $\chi^2_{.95} = 10.8508$, $\chi^2_{.05} = 31.4104$, a 90% CI for θ is $\left(\frac{2\sum_{i=1}^{n} Y_i}{31.4104}, \frac{2\sum_{i=1}^{n} Y_i}{10.8508} \right)$.

9.86 First, similar to Example 9.15, the MLEs of μ_1 and μ_2 are $\hat{\mu}_1 = \overline{X}$ and $\hat{\mu}_2 = \overline{Y}$. To estimate σ^2, the likelihood is

$$L(\sigma^2) = \frac{1}{(2\pi)^{(m+n)/2}\sigma^{m+n}} \exp\left\{ -\tfrac{1}{2}\left[\sum_{i=1}^{m}\left(\frac{x_i - \mu_1}{\sigma}\right)^2 - \sum_{i=1}^{n}\left(\frac{y_i - \mu_2}{\sigma}\right)^2 \right] \right\}.$$

The log–likelihood is

$$\ln L(\sigma^2) = K - (m+n)\ln\sigma - \tfrac{1}{2\sigma^2}\left[\sum_{i=1}^{m}(x_i - \mu_1)^2 - \sum_{i=1}^{n}(y_i - \mu_2)^2 \right]$$

By differentiating and setting this quantity equal to 0, we obtain

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{m}(x_i - \mu_1)^2 - \sum_{i=1}^{n}(y_i - \mu_2)^2}{m+n}.$$

As in Example 9.15, the MLEs of μ_1 and μ_2 can be used in the above to arrive at the MLE for σ^2:

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{m}(X_i - \overline{X})^2 - \sum_{i=1}^{n}(Y_i - \overline{Y})^2}{m+n}.$$

9.87 Let Y_1 = # of candidates favoring candidate A, Y_2 = # of candidate favoring candidate B, and Y_3 = # of candidates favoring candidate C. Then, (Y_1, Y_2, Y_3) is trinomial with parameters (p_1, p_2, p_3) and sample size n. Thus, the likelihood $L(p_1, p_2)$ is simply the probability mass function for the trinomial (recall that $(p_3 = 1 - p_1 - p_2)$:

$$L(p_1, p_2) = \tfrac{n!}{n_1!n_2!n_3!} p_1^{y_1} p_2^{y_2} (1 - p_1 - p_2)^{y_3}$$

This can easily be jointly maximized with respect to p_1 and p_2 to obtain the MLEs $\hat{p}_1 = Y_1/n$, $\hat{p}_2 = Y_2/n$, and so $\hat{p}_3 = Y_3/n$.

For the given data, we have $\hat{p}_1 = .30$, $\hat{p}_2 = .38$, and $\hat{p}_3 = .32$. Thus, the point estimate of $p_1 - p_2$ is $.30 - .38 = -.08$. From Theorem 5.13, we have that $V(Y_i) = np_iq_i$ and $\text{Cov}(Y_i, Y_j) = -np_ip_j$. A two–standard–deviation error bound can be found by

$$2\sqrt{V(\hat{p}_1 - \hat{p}_2)} = 2\sqrt{V(\hat{p}_1) + V(\hat{p}_2) - 2Cov(\hat{p}_1, \hat{p}_2)} = 2\sqrt{p_1q_1/n + p_2q_2/n + 2p_1p_2/n}.$$

This can be estimated by using the MLEs found above. By plugging in the estimates, error bound of .1641 is obtained.

9.88 The likelihood function is $L(\theta) = (\theta+1)^n \left(\prod_{i=1}^n y_i\right)^\theta$. The MLE is $\hat{\theta} = -n / \sum_{i=1}^n \ln Y_i$. This is a different estimator that the MOM estimator from Ex. 9.69, however note that the MLE is a function of the sufficient statistic.

9.89 Note that the likelihood is simply the mass function for *Y*: $L(p) = \binom{2}{y} p^y (1-p)^{2-y}$. By the ML criteria, we choose the value of *p* that maximizes the likelihood. If $Y = 0$, $L(p)$ is maximized at $p = .25$. If $Y = 2$, $L(p)$ is maximized at $p = .75$. But, if $Y = 1$, $L(p)$ has the same value at both $p = .25$ and $p = .75$; that is, $L(.25) = L(.75)$ for $y = 1$. Thus, for this instance the MLE is not unique.

9.90 Under the hypothesis that $p_W = p_M = p$, then Y = # of people in the sample who favor the issue is binomial with success probability *p* and $n = 200$. Thus, by Example 9.14, the MLE for *p* is $\hat{p} = Y/n$ and the sample estimate is 55/200.

9.91 Refer to Ex. 9.83 and Example 9.16. Let $\gamma = 2\theta$. Then, the MLE for γ is $\hat{\gamma} = Y_{(n)}$ and by the invariance principle the MLE for θ is $\hat{\theta} = Y_{(n)}/2$.

9.92 **a.** Following the hint, the MLE of θ is $\hat{\theta} = Y_{(n)}$.

b. From Ex. 9.63, $f_{(n)}(y) = 3ny^{3n-1}/\theta^{3n}$, $0 \le y \le \theta$. The distribution of $T = Y_{(n)}/\theta$ is

$$f_T(t) = 3nt^{3n-1},\ 0 \le t \le 1.$$

Since this distribution doesn't depend on θ, *T* is a pivotal quantity.

c. (Similar to Ex. 8.132) Constants *a* and *b* can be found to satisfy $P(a < T < b) = 1 - \alpha$ such that $P(T < a) = P(T > b) = \alpha/2$. Using the density function from part b, these are given by $a = (\alpha/2)^{1/(3n)}$ and $b = (1-\alpha/2)^{1/(3n)}$. So, we have

$$1 - \alpha = P(a < Y_{(n)}/\theta < b) = P\left(Y_{(n)}/b < \theta < Y_{(n)}/a\right).$$

Thus, $\left(\dfrac{Y_{(n)}}{(1-\alpha/2)^{1/(3n)}}, \dfrac{Y_{(n)}}{(\alpha/2)^{1/(3n)}}\right)$ is a $(1-\alpha)100\%$ CI for θ.

9.93 **a.** Following the hint, the MLE for θ is $\hat{\theta} = Y_{(1)}$.

b. Since $F(y \mid \theta) = 1 - 2\theta^2 y^{-2}$, the density function for $Y_{(1)}$ is easily found to be

$$g_{(1)}(y) = 2n\theta^{2n} y^{-(2n+1)},\ y > \theta.$$

If we consider the distribution of $T = \theta/Y_{(1)}$, the density function of *T* can be found to be

$$f_T(t) = 2nt^{2n-1},\ 0 < t < 1.$$

c. (Similar to Ex. 9.92) Constants *a* and *b* can be found to satisfy $P(a < T < b) = 1 - \alpha$ such that $P(T < a) = P(T > b) = \alpha/2$. Using the density function from part b, these are given by $a = (\alpha/2)^{1/(2n)}$ and $b = (1-\alpha/2)^{1/(2n)}$. So, we have

$$1-\alpha = P(a < \theta/Y_{(1)} < b) = P\left(aY_{(1)} < \theta < bY_{(1)}\right).$$

Thus, $\left[(\alpha/2)^{1/(2n)}Y_{(1)}, (1-\alpha/2)^{1/(2n)}Y_{(1)}\right]$ is a (1 – α)100% CI for θ.

9.94 Let β = t(θ) so that θ = $t^{-1}(\beta)$. If the likelihood is maximized at $\hat{\theta}$, then $L(\hat{\theta}) \geq L(\theta)$ for all θ. Define $\hat{\beta} = t(\hat{\theta})$ and denote the likelihood as a function of β as $L_1(\beta) = L(t^{-1}(\beta))$. Then, for any β,

$$L_1(\beta) = L(t^{-1}(\beta)) = L(\theta) \leq L(\hat{\theta}) = L(t^{-1}(\hat{\beta})) = L_1(\hat{\beta}).$$

So, the MLE of β is $\hat{\beta}$ and so the MLE of t(θ) is $t(\hat{\theta})$.

9.95 The quantity to be estimated is $R = p/(1-p)$. Since $\hat{p} = Y/n$ is the MLE of p, by the invariance principle the MLE for R is $\hat{R} = \hat{p}/(1-\hat{p})$.

9.96 From Ex. 9.15, the MLE for σ^2 was found to be $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \overline{Y})^2$. By the invariance property, the MLE for σ is $\hat{\sigma} = \sqrt{\hat{\sigma}^2} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - \overline{Y})^2}$.

9.97 **a.** Since $\mu_1' = 1/p$, the MOM estimator for p is $\hat{p} = 1/m_1' = 1/\overline{Y}$.

b. The likelihood function is $L(p) = p^n(1-p)^{\Sigma y_i - n}$ and the log–likelihood is

$$\ln L(p) = n\ln p + \left(\sum_{i=1}^{n} y_i - n\right)\ln(1-p).$$

Differentiating, we have

$$\tfrac{d}{dp}\ln L(p) = \tfrac{n}{p} - \tfrac{1}{1-p}\left(\sum_{i=1}^{n} y_i - n\right).$$

Equating this to 0 and solving for p, we obtain the MLE $\hat{p} = 1/\overline{Y}$, which is the same as the MOM estimator found in part a.

9.98 Since $\ln p(y \mid p) = \ln p + (y-1)\ln(1-p)$,

$$\tfrac{d}{dp}\ln p(y \mid p) = 1/p - (y-1)/(1-p)$$

$$\tfrac{d^2}{dp^2}\ln p(y \mid p) = -1/p^2 - (y-1)/(1-p)^2.$$

Then,

$$-E\left[\tfrac{d^2}{dp^2}\ln p(Y \mid p)\right] = -E\left[-1/p^2 - (Y-1)/(1-p)^2\right] = \frac{1}{p^2(1-p)}.$$

Therefore, the approximate (limiting) variance of the MLE (as given in Ex. 9.97) is given by

$$V(\hat{p}) \approx \frac{p^2(1-p)}{n}.$$

9.99 From Ex. 9.18, the MLE for $t(p) = p$ is $\hat{p} = Y/n$ and with $-E\left[\frac{d^2}{dp^2}\ln p(Y \mid p)\right] = \frac{1}{p(1-p)}$, a 100(1 – α)% CI for p is $\hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$. This is the same CI for p derived in Section 8.6.

9.100 In Ex. 9.81, it was shown that $\overline{Y}^2$ is the MLE of $t(\theta) = \theta^2$. It is easily found that for the exponential distribution with mean θ,

$$-E\left[\frac{d^2}{dp^2}\ln f(Y \mid \theta)\right] = \frac{1}{\theta^2}.$$

Thus, since $\frac{d}{d\theta}t(\theta) = 2\theta$, we have an approximate (large sample) 100(1 – α)% CI for θ as

$$\overline{Y}^2 \pm z_{\alpha/2}\sqrt{\left.\left(\frac{(2\theta)^2}{n\frac{1}{\theta^2}}\right)\right|_{\theta=\hat{\theta}}} = \overline{Y}^2 \pm z_{\alpha/2}\left(\frac{2\overline{Y}^2}{\sqrt{n}}\right).$$

9.101 From Ex. 9.80, the MLE for $t(\lambda) = \exp(-\lambda)$ is $t(\hat{\lambda}) = \exp(-\hat{\lambda}) = \exp(-\overline{Y})$. It is easily found that for the Poisson distribution with mean λ,

$$-E\left[\frac{d^2}{dp^2}\ln p(Y \mid \lambda)\right] = \frac{1}{\lambda}.$$

Thus, since $\frac{d}{d\lambda}t(\lambda) = -\exp(-\lambda)$, we have an approximate 100(1 – α)% CI for λ as

$$\exp(-\overline{Y}) \pm z_{\alpha/2}\sqrt{\left.\frac{\exp(-2\lambda)}{n\frac{1}{\lambda}}\right|_{\lambda=\overline{Y}}} = \exp(-\overline{Y}) \pm z_{\alpha/2}\sqrt{\frac{\overline{Y}\exp(-2\overline{Y})}{n}}.$$

9.102 With $n = 30$ and $\bar{y} = 4.4$, the maximum likelihood estimate of p is 1/(4.4) = .2273 and an approximate 90% CI for p is

$$\hat{p} \pm z_{.025}\sqrt{\frac{\hat{p}^2(1-\hat{p})}{n}} = .2273 \pm 1.96\sqrt{\frac{(.2273)^2(.7727)}{30}} = .2273 \pm .0715 \text{ or } (.1558, .2988).$$

9.103 The Rayleigh distribution is a special case of the (Weibull) distribution from Ex. 9.82. Also see Example 9.7

a. From Ex. 9.82 with $r = 2$, $\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} Y_i^2$.

b. It is easily found that for the Rayleigh distribution with parameter θ,

$$\frac{d^2}{dp^2}\ln f(Y \mid \theta) = \frac{1}{\theta^2} - \frac{2Y^2}{\theta^3}.$$

Since $E(Y^2) = \theta$, $-E\left[\frac{d^2}{dp^2}\ln f(Y \mid \theta)\right] = \frac{1}{\theta^2}$ and so $V(\hat{\theta}) \approx \theta^2/n$.

9.104 **a.** MOM: $\mu_1' = E(Y) = \theta + 1$, so $\hat{\theta}_1 = m_1' - 1 = \overline{Y} - 1$.

b. MLE: $\hat{\theta}_2 = Y_{(1)}$, the first order statistic.

c. The estimator $\hat{\theta}_1$ is unbiased since $E(\hat{\theta}_1) = E(\overline{Y}) - 1 = \theta + 1 - 1 = \theta$. The distribution of $Y_{(1)}$ is $g_{(1)}(y) = ne^{-n(y-\theta)}$, $y > \theta$. So, $E(Y_{(1)}) = E(\hat{\theta}_2) = \frac{1}{n} + \theta$. Thus, $\hat{\theta}_2$ is not unbiased but $\hat{\theta}_2^* = Y_{(1)} - \frac{1}{n}$ is unbiased for θ.

The efficiency of $\hat{\theta}_1 = \overline{Y} - 1$ relative to $\hat{\theta}_2^* = Y_{(1)} - \frac{1}{n}$ is given by

$$\text{eff}(\hat{\theta}_1, \hat{\theta}_2^*) = \frac{V(\hat{\theta}_2^*)}{V(\hat{\theta}_1)} = \frac{V(Y_{(1)} - \frac{1}{n})}{V(\overline{Y} - 1)} = \frac{V(Y_{(1)})}{V(\overline{Y})} = \frac{\frac{1}{n^2}}{\frac{1}{n}} = \frac{1}{n}.$$

9.105 From Ex. 9.38, we must solve

$$\frac{d^2 \ln L}{d\sigma^2} = \frac{-n}{2\sigma^2} + \frac{\Sigma(y_i - \mu)^2}{2\sigma^4} = 0, \quad \text{so} \quad \hat{\sigma}^2 = \frac{\Sigma(y_i - \mu)^2}{n}.$$

9.106 Following the method used in Ex. 9.65, construct the random variable T such that

$T = 1$ if $Y_1 = 0$ and $T = 0$ otherwise

Then, $E(T) = P(T = 1) = P(Y_1 = 0) = \exp(-\lambda)$. So, T is unbiased for $\exp(-\lambda)$. Now, we know that $W = \sum_{i=1}^n Y_i$ is sufficient for λ, and so it is also sufficient for $\exp(-\lambda)$. Recalling that W has a Poisson distribution with mean $n\lambda$,

$$E(T \mid W = w) = P(T = 1 \mid W = w) = P(Y_1 = 0 \mid W = w) = \frac{P(Y_1 = 0, W = w)}{P(W = w)}$$

$$= \frac{P(Y_1 = 0) P\left(\sum_{i=2}^n Y_i = w\right)}{P(W = w)} = \frac{e^{-\lambda}\left(e^{-(n-1)\lambda} \frac{[(n-1)\lambda]^w}{w!}\right)}{e^{-n\lambda} \frac{(n\lambda)^w}{w!}} = \left(1 - \tfrac{1}{n}\right)^w.$$

Thus, the MVUE is $\left(1 - \frac{1}{n}\right)^{\Sigma Y_i}$. Note that in the above we used the result that $\sum_{i=2}^n Y_i$ is Poisson with mean $(n-1)\lambda$.

9.107 The MLE of θ is $\hat{\theta} = \overline{Y}$. By the invariance principle for MLEs, the MLE of $\overline{F}(t)$ is $\hat{\overline{F}}(t) = \exp(-t/\overline{Y})$.

9.108 **a.** $E(V) = P(Y_1 > t) = 1 - F(t) = \exp(-t/\theta)$. Thus, V is unbiased for $\exp(-t/\theta)$.

b. Recall that U has a gamma distribution with shape parameter n and scale parameter θ. Also, $U - Y_1 = \sum_{i=2}^n Y_i$ is gamma with shape parameter $n - 1$ and scale parameter θ, and since Y_1 and $U - Y_1$ are independent,

$$f(y_1, u - y_1) = \left(\tfrac{1}{\theta} e^{-y_1/\theta}\right) \tfrac{1}{\Gamma(n-1)\theta^{n-1}} (u - y_1)^{n-2} e^{-(u-y_1)/\theta}, \ 0 \le y_1 \le u < \infty.$$

Next, apply the transformation $z = u - y_1$ such that $u = z + y_1$ to get the joint distribution

$$f(y_1, u) = \tfrac{1}{\Gamma(n-1)\theta^n} (u - y_1)^{n-2} e^{-u/\theta}, \ 0 \le y_1 \le u < \infty.$$

Now, we have

$$f(y_1 \mid u) = \frac{f(y_1, u)}{f(u)} = \left(\frac{n-1}{u^{n-1}}\right)(u - y_1)^{n-2}, \ 0 \le y_1 \le u < \infty.$$

c. $E(V|U) = P(Y_1 > t|U = u) = \int_t^u \left(\frac{n-1}{u^{n-1}}\right)(u - y_1)^{n-2} dy_1 = \int_t^u \left(\frac{n-1}{u}\right)\left(1 - \frac{y_1}{u}\right)^{n-2} dy_1$

$$= -\left(1 - \frac{y_1}{u}\right)^{n-1}\Bigg|_t^u = \left(1 - \frac{t}{u}\right)^{n-1}.$$

So, the MVUE is $\left(1 - \frac{t}{U}\right)^{n-1}$.

9.109 Let $Y_1, Y_2, \ldots, Y_n$ represent the (independent) values drawn on each of the n draws. Then, the probability mass function for each Y_i is

$$P(Y_i = k) = \tfrac{1}{N},\ k = 1, 2, \ldots, N.$$

a. Since $\mu_1' = E(Y) = \sum_{k=1}^N kP(Y = k) = \sum_{k=1}^N k\frac{1}{N} = \frac{N(N+1)}{2N} = \frac{N+1}{2}$, the MOM estimator of N is $\frac{\hat{N}_1+1}{2} = \overline{Y}$ or $\hat{N}_1 = 2\overline{Y} - 1$.

b. First, $E(\hat{N}_1) = 2E(\overline{Y}) - 1 = 2\left(\frac{N+1}{2}\right) - 1 = N$, so $\hat{N}_1$ is unbiased. Now, since $E(Y^2) = \sum_{k=1}^N k^2 \frac{1}{N} = \frac{N(N+1)(2N+1)}{6N} = \frac{(N+1)(2N+1)}{6}$, we have that $V(Y) = \frac{(N+1)(N-1)}{12}$. Thus, $V(\hat{N}_1) = 4V(\overline{Y}) = 4\left(\frac{(N+1)(N-1)}{12n}\right) = \frac{N^2-1}{3n}$.

9.110 **a.** Following Ex. 9.109, the likelihood is

$$L(N) = \tfrac{1}{N^n}\prod_{i=1}^n I(y_i \in \{1, 2, \ldots, N\}) = \tfrac{1}{N^n} I(y_{(n)} \le N).$$

In order to maximize L, N should be chosen as small as possible subject to the constraint that $y_{(n)} \le N$. Thus $\hat{N}_2 = Y_{(n)}$.

b. Since $P(\hat{N}_2 \le k) = P(Y_{(n)} \le k) = P(Y_1 \le k)\cdots P(Y_n \le k) = \left(\frac{k}{N}\right)^n$, so $P(\hat{N}_2 \le k-1) = \left(\frac{k-1}{N}\right)^n$ and $P(\hat{N}_2 = k) = \left(\frac{k}{N}\right)^n - \left(\frac{k-1}{N}\right)^n = N^{-n}[k^n - (k-1)^n]$. So,

$$E(\hat{N}_2) = N^{-n}\sum_{k=1}^N k[k^n - (k-1)^n] = N^{-n}\sum_{k=1}^N [k^{n+1} - (k-1)^{n+1} - (k-1)^n]$$
$$= N^{-n}\left[N^{n+1} - \sum_{k=1}^N (k-1)^n\right].$$

Consider $\sum_{k=1}^N (k-1)^n = 0^n + 1^n + 2^n + \ldots + (N-1)^n$. For large N, this is approximately the area beneath the curve $f(x) = x^n$ from $x = 0$ to $x = N$, or $\sum_{k=1}^N (k-1)^n \approx \int_0^N x^n dx = \frac{N^{n+1}}{n+1}$.

Thus, $E(\hat{N}_2) \approx N^{-n}[N^{n+1} - \frac{N^{n+1}}{n+1}] = \frac{n}{n+1}N$ and $\hat{N}_3 = \frac{n+1}{n}\hat{N}_2 = \frac{n+1}{n}Y_{(n)}$ is approximately unbiased for N.

c. $V(\hat{N}_2)$ is given, so $V(\hat{N}_3) = \left(\frac{n+1}{n}\right)^2 V(\hat{N}_2) = \frac{N^2}{n(n+2)}$.

d. Note that, for $n > 1$,

$$\frac{V(\hat{N}_1)}{V(\hat{N}_3)} = \frac{n(n+2)}{3n}\frac{(N^2-1)}{N^2} = \frac{n+2}{3}\left(1-\frac{1}{N^2}\right) > 1,$$

since for large N, $\left(1-\frac{1}{N^2}\right) \approx 1$

9.111 The (approximately) unbiased estimate of N is $\hat{N}_3 = \frac{n+1}{n}Y_{(n)} = \frac{6}{5}(210) = 252$ and an approximate error bound is given by

$$2\sqrt{V(\hat{N}_3)} \approx 2\sqrt{\frac{N^2}{n(n+2)}} \approx 2\sqrt{\frac{(252)^2}{5(7)}} = 85.192.$$

9.112 **a.** (Refer to Section 9.3.) By the Central Limit Theorem, $\frac{\overline{Y}-\lambda}{\sqrt{\lambda/n}}$ converges to a standard normal variable. Also, $\overline{Y}/\lambda$ converges in probability to 1 by the Law of Large Numbers, as does $\sqrt{\overline{Y}/\lambda}$. So, the quantity

$$W_n = \frac{\dfrac{\overline{Y}-\lambda}{\sqrt{\lambda/n}}}{\sqrt{\overline{Y}/\lambda}} = \frac{\overline{Y}-\lambda}{\sqrt{\overline{Y}/n}}$$

converges to a standard normal distribution.

b. By part **a**, an approximate $(1-\alpha)100\%$ CI for λ is $\overline{Y} \pm z_{\alpha/2}\sqrt{\overline{Y}/n}$.

Chapter 10: Hypothesis Testing

10.1 See Definition 10.1.

10.2 Note that Y is binomial with parameters $n = 20$ and p.

a. If the experimenter concludes that less than 80% of insomniacs respond to the drug when actually the drug induces sleep in 80% of insomniacs, a type I error has occurred.

b. $\alpha = P(\text{reject } H_0 \mid H_0 \text{ true}) = P(Y \le 12 \mid p = .8) = .032$ (using Appendix III).

c. If the experimenter does not reject the hypothesis that 80% of insomniacs respond to the drug when actually the drug induces sleep in fewer than 80% of insomniacs, a type II error has occurred.

d. $\beta(.6) = P(\text{fail to reject } H_0 \mid H_a \text{ true}) = P(Y > 12 \mid p = .6) = 1 - P(Y \le 12 \mid p = .6) = .416$.

e. $\beta(.4) = P(\text{fail to reject } H_0 \mid H_a \text{ true}) = P(Y > 12 \mid p = .4) = .021$.

10.3 **a.** Using the Binomial Table, $P(Y \le 11 \mid p = .8) = .011$, so $c = 11$.

b. $\beta(.6) = P(\text{fail to reject } H_0 \mid H_a \text{ true}) = P(Y > 11 \mid p = .6) = 1 - P(Y \le 11 \mid p = .6) = .596$.

c. $\beta(.4) = P(\text{fail to reject } H_0 \mid H_a \text{ true}) = P(Y > 11 \mid p = .4) = .057$.

10.4 The parameter p = proportion of ledger sheets with errors.

a. If it is concluded that the proportion of ledger sheets with errors is larger than .05, when actually the proportion is equal to .05, a type I error occurred.

b. By the proposed scheme, H_0 will be rejected under the following scenarios (let E = error, N = no error):

Sheet 1	Sheet 2	Sheet 3
N	N	.
N	E	N
E	N	N
E	E	N

With $p = .05$, $\alpha = P(NN) + P(NEN) + P(ENN) + P(EEN) = (.95)^2 + 2(.05)(.95)^2 + (.05)^2(.95) = .995125$.

c. If it is concluded that $p = .05$, but in fact $p > .05$, a type II error occurred.

d. $\beta(p_a) = P(\text{fail to reject } H_0 \mid H_a \text{ true}) = \text{P}(EEE, NEE, \text{ or } ENE \mid p_a) = 2p_a^2(1 - p_a) + p_a^3$.

10.5 Under H_0, Y_1 and Y_2 are uniform on the interval (0, 1). From Example 6.3, the distribution of $U = Y_1 + Y_2$ is

$$g(u) = \begin{cases} u & 0 \le u \le 1 \\ 2 - u & 1 < u \le 2 \end{cases}$$

Test 1: $P(Y_1 > .95) = .05 = \alpha$.

Test 2: $\alpha = .05 = P(U > c) = \int_c^2 (2 - u)du = 2 = 2c + .5c^2$. Solving the quadratic gives the plausible solution of $c = 1.684$.

10.6 The test statistic Y is binomial with $n = 36$.

a. $\alpha = P(\text{reject } H_0 \mid H_0 \text{ true}) = P(|Y - 18| \geq 4 \mid p = .5) = P(Y \leq 14) + P(Y \geq 22) = .243$.

b. $\beta = P(\text{fail to reject } H_0 \mid H_a \text{ true}) = P(|Y - 18| \leq 3 \mid p = .7) = P(15 \leq Y \leq 21 \mid p = .7) = .09155$.

10.7 **a.** False, H_0 is not a statement involving a random quantity.
b. False, for the same reason as part a.
c. True.
d. True.
e. False, this is given by α.
f. **i.** True.
ii. True.
iii. False, β and α behave inversely to each other.

10.8 Let Y_1 and Y_2 have binomial distributions with parameters $n = 15$ and p.

a. $\alpha = P(\text{reject } H_0 \text{ in stage 1} \mid H_0 \text{ true}) + \text{P}(\text{reject } H_0 \text{ in stage 2} \mid H_0 \text{ true})$

$$= P(Y_1 \geq 4) + P(Y_1 + Y_2 \geq 6,\, Y_1 \leq 3) = P(Y_1 \geq 4) + \sum_{i=0}^{3} P(Y_1 + Y_2 \geq 6,\, Y_1 \leq i)$$

$$= P(Y_1 \geq 4) + \sum_{i=0}^{3} P(Y_2 \geq 6 - i) P(Y_1 \leq i) = .0989 \text{ (calculated with } p = .10).$$

Using R, this is found by:

```
> 1 - pbinom(3,15,.1)+sum((1-pbinom(5-0:3,15,.1))*dbinom(0:3,15,.1))
[1] 0.0988643
```

b. Similar to part a with $p = .3$: $\alpha = .9321$.

c. $\beta = P(\text{fail to reject } H_0 \mid p = .3)$

$$= \sum_{i=0}^{3} P(Y_1 = i,\, Y_1 + Y_2 \leq 5) = \sum_{i=0}^{3} P(Y_2 = 5 - i) P(\,Y_1 = i) = .0679.$$

10.9 **a.** The simulation is performed with a known $p = .5$, so rejecting H_0 is a type I error.
b.-e. Answers vary.
f. This is because of part a.
g.-h. Answers vary.

10.10 **a.** An error is the rejection of H_0 (type I).
b. Here, the error is failing to reject H_0 (type II).
c. H_0 is rejected more frequently the further the true value of p is from .5.
d. Similar to part **c**.

10.11 **a.** The error is failing to reject H_0 (type II).
b.-d. Answers vary.

10.12 Since β and α behave inversely to each other, the simulated value for β should be smaller for $\alpha = .10$ than for $\alpha = .05$.

10.13 The simulated values of β and α should be closer to the nominal levels specified in the simulation.

10.14 **a.** The smallest value for the test statistic is –.75. Therefore, since the RR is {z < –.84}, the null hypothesis will never be rejected. The value of n is far too small for this large–sample test.
b. Answers vary.
c. H_0 is rejected when $\hat{p} = 0.00$. $P(Y = 0 \mid p = .1) = .349 > .20$.
d. Answers vary, but n should be large enough.

10.15 **a.** Answers vary.
b. Answers vary.

10.16 **a.** Incorrect decision (type I error).
b. Answers vary.
c. The simulated rejection (error) rate is .000, not close to $\alpha = .05$.

10.17 **a.** H_0: $\mu_1 = \mu_2$, H_a: $\mu_1 > \mu_2$.
b. Reject if $Z > 2.326$, where Z is given in Example 10.7 ($D_0 = 0$).
c. $z = .075$.
d. Fail to reject H_0 – not enough evidence to conclude the mean distance for breaststroke is larger than individual medley.
e. The sample variances used in the test statistic were too large to be able to detect a difference.

10.18 H_0: $\mu = 13.20$, H_a: $\mu < 13.20$. Using the large sample test for a mean, $z = –2.53$, and with $\alpha = .01$, $–z_{.01} = –2.326$. So, H_0 is rejected: there is evidence that the company is paying substandard wages.

10.19 H_0: $\mu = 130$, H_a: $\mu < 130$. Using the large sample test for a mean, $z = \frac{128.6-130}{2.1/\sqrt{40}} = -4.22$ and with $–z_{.05} = –1.645$, H_0 is rejected: there is evidence that the mean output voltage is less than 130.

10.20 H_0: $\mu \geq 64$, H_a: $\mu < 64$. Using the large sample test for a mean, $z = –1.77$, and w/ $\alpha = .01$, $–z_{.01} = –2.326$. So, H_0 is not rejected: there is not enough evidence to conclude the manufacturer's claim is false.

10.21 Using the large–sample test for two means, we obtain $z = 3.65$. With $\alpha = .01$, the test rejects if $|z| > 2.576$. So, we can reject the hypothesis that the soils have equal mean shear strengths.

10.22 **a.** The mean pretest scores should probably be equal, so letting μ_1 and μ_2 denote the mean pretest scores for the two groups, H_0: $\mu_1 = \mu_2$, H_a: $\mu_1 \neq \mu_2$.
b. This is a two–tailed alternative: reject if $|z| > z_{\alpha/2}$.
c. With $\alpha = .01$, $z_{.005} = 2.576$. The computed test statistic is $z = 1.675$, so we fail to reject H_0: we cannot conclude the there is a difference in the pretest mean scores.

10.23 **a.-b.** Let μ_1 and μ_2 denote the mean distances. Since there is no prior knowledge, we will perform the test H_0: $\mu_1 - \mu_2 = 0$ vs. H_a: $\mu_1 - \mu_2 \neq 0$, which is a two–tailed test.

c. The computed test statistic is $z = -.954$, which does not lead to a rejection with $\alpha = .10$: there is not enough evidence to conclude the mean distances are different.

10.24 Let p = proportion of overweight children and adolescents. Then, H_0: $p = .15$, H_a: $p < .15$ and the computed large sample test statistic for a proportion is $z = -.56$. This does not lead to a rejection at the $\alpha = .05$ level.

10.25 Let p = proportion of adults who always vote in presidential elections. Then, H_0: $p = .67$, H_a: $p \neq .67$ and the large sample test statistic for a proportion is $|z| = 1.105$. With $z_{.005} = 2.576$, the null hypothesis cannot be rejected: there is not enough evidence to conclude the reported percentage is false.

10.26 Let p = proportion of Americans with brown eyes. Then, H_0: $p = .45$, H_a: $p \neq .45$ and the large sample test statistic for a proportion is $z = -.90$. We fail to reject H_0.

10.27 Define: p_1 = proportion of English–fluent Riverside students
p_2 = proportion of English–fluent Palm Springs students.

To test H_0: $p_1 - p_2 = 0$, versus H_a: $p_1 - p_2 \neq 0$, we can use the large–sample test statistic

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}}.$$

However, this depends on the (unknown) values p_1 and p_2. Under H_0, $p_1 = p_2 = p$ (i.e. they are samples from the same binomial distribution), so we can "pool" the samples to estimate p:

$$\hat{p}_p = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2} = \frac{Y_1 + Y_2}{n_1 + n_2}.$$

So, the test statistic becomes

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}_p \hat{q}_p\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}.$$

Here, the value of the test statistic is $z = -.1202$, so a significant difference cannot be supported.

10.28 **a.** (Similar to 10.27) Using the large–sample test derived in Ex. 10.27, the computed test statistic is $z = -2.254$. Using a two–sided alternative, $z_{.025} = 1.96$ and since $|z| > 1.96$, we can conclude there is a significant difference between the proportions.

b. Advertisers should consider targeting females.

10.29 Note that color A is preferred over B and C if it has the highest probability of being purchased. Thus, let p = probability customer selects color A. To determine if A is preferred, consider the test H_0: $p = 1/3$, H_a: $p > 1/3$. With $\hat{p} = 400/1000 = .4$, the test statistic is $z = 4.472$. This rejects H_0 with $\alpha = .01$, so we can safely conclude that color A is preferred (note that it was assumed that "the first 1000 washers sold" is a random sample).

10.30 Let $\hat{p}$ = sample percentage preferring the product. With $\alpha = .05$, we reject H_0 if

$$\frac{\hat{p}-.2}{\sqrt{.2(.8)/100}} < -1.645\,.$$

Solving for $\hat{p}$, the solution is $\hat{p} < .1342$.

10.31 The assumptions are: (1) a random sample (2) a (limiting) normal distribution for the pivotal quantity (3) known population variance (or sample estimate can be used for large n).

10.32 Let p = proportion of U.S. adults who feel the environment quality is fair or poor. To test H_0: $p = .50$ vs. H_a: $p > 50$, we have that $\hat{p} = .54$ so the large–sample test statistic is $z = 2.605$ and with $z_{.05} = 1.645$, we reject H_0 and conclude that there is sufficient evidence to conclude that a majority of the nation's adults think the quality of the environment is fair or poor.

10.33 (Similar to Ex. 10.27) Define:

p_1 = proportion of Republicans strongly in favor of the death penalty
p_2 = proportion of Democrats strongly in favor of the death penalty

To test H_0: $p_1 - p_2 = 0$ vs. H_a: $p_1 - p_2 > 0$, we can use the large–sample test derived in Ex. 10.27 with $\hat{p}_1 = .23$, $\hat{p}_2 = .17$, and $\hat{p}_p = .20$. Thus, $z = 1.50$ and for $z_{.05} = 1.645$, we fail to reject H_0: there is not enough evidence to support the researcher's belief.

10.34 Let μ = mean length of stay in hospitals. Then, for H_0: $\mu = 5$, H_a: $\mu > 5$, the large sample test statistic is $z = 2.89$. With $\alpha = .05$, $z_{.05} = 1.645$ so we can reject H_0 and support the agency's hypothesis.

10.35 (Similar to Ex. 10.27) Define:

p_1 = proportion of currently working homeless men
p_2 = proportion of currently working domiciled men

The hypotheses of interest are H_0: $p_1 - p_2 = 0$, H_a: $p_1 - p_2 < 0$, and we can use the large–sample test derived in Ex. 10.27 with $\hat{p}_1 = .30$, $\hat{p}_2 = .38$, and $\hat{p}_p = .355$. Thus, $z = -1.48$ and for $-z_{.01} = -2.326$, we fail to reject H_0: there is not enough evidence to support the claim that the proportion of working homeless men is less than the proportion of working domiciled men.

10.36 (similar to Ex. 10.27) Define:

p_1 = proportion favoring complete protection
p_2 = proportion desiring destruction of nuisance alligators
Using the large–sample test for H_0: $p_1 - p_2 = 0$ versus H_a: $p_1 - p_2 \neq 0$, $z = -4.88$. This value leads to a rejections at the $\alpha = .01$ level so we conclude that there is a difference.

10.37 With H_0: $\mu = 130$, this is rejected if $z = \frac{\bar{y}-130}{\sigma/\sqrt{n}} < -1.645$, or if $\bar{y} < 130 - \frac{1.645\sigma}{\sqrt{n}} = 129.45$. If $\mu = 128$, then $\beta = P(\bar{Y} > 129.45 \mid \mu = 128) = P(Z > \frac{129.45-128}{2.1/\sqrt{40}}) = P(Z > 4.37) = .0000317$.

10.38 With H_0: $\mu \geq 64$, this is rejected if $z = \frac{\bar{y}-64}{\sigma/\sqrt{n}} < -2.326$, or if $\bar{y} < 64 - \frac{2.326\sigma}{\sqrt{n}} = 61.36$. If μ = 60, then $\beta = P(\bar{Y} > 61.36 \mid \mu = 60) = P(Z > \frac{61.36-60}{8/\sqrt{50}}) = P(Z > 1.2) = .1151$.

10.39 In Ex. 10.30, we found the rejection region to be: $\{\hat{p} < .1342\}$. For p = .15, the type II error rate is $\beta = P(\hat{p} > .1342 \mid p = .15) = P\left(Z > \frac{.1342-.15}{\sqrt{.15(.85)/100}}\right) = P(Z > -.4424) = .6700$.

10.40 Refer to Ex. 10.33. The null and alternative tests were H_0: $p_1 - p_2 = 0$ vs. H_a: $p_1 - p_2 > 0$. We must find a common sample size n such that $\alpha = P(\text{reject } H_0 \mid H_0 \text{ true}) = .05$ and $\beta = P(\text{fail to reject } H_0 \mid H_a \text{ true}) \leq .20$. For $\alpha = .05$, we use the test statistic

$$Z = \frac{\hat{p}_1 - \hat{p}_2 - 0}{\sqrt{\frac{p_1q_1}{n} + \frac{p_2q_2}{n}}}$$ such that we reject H_0 if $Z \geq z_{.05} = 1.645$. In other words,

$$\text{Reject } H_0 \text{ if: } \hat{p}_1 - \hat{p}_2 \geq 1.645\sqrt{\frac{p_1q_1}{n} + \frac{p_2q_2}{n}}.$$

For β, we fix it at the largest acceptable value so $P(\hat{p}_1 - \hat{p}_2 \leq c \mid p_1 - p_2 = .1) = .20$ for some c, or simply

$$\text{Fail to reject } H_0 \text{ if: } \frac{\hat{p}_1 - \hat{p}_2 - .1}{\sqrt{\frac{p_1q_1}{n} + \frac{p_2q_2}{n}}} = -.84, \text{ where } -.84 = z_{.20}.$$

Let $\hat{p}_1 - \hat{p}_2 = 1.645\sqrt{\frac{p_1q_1}{n} + \frac{p_2q_2}{n}}$ and substitute this in the above statement to obtain

$$-.84 = \frac{1.645\sqrt{\frac{p_1q_1}{n} + \frac{p_2q_2}{n}} - .1}{\sqrt{\frac{p_1q_1}{n} + \frac{p_2q_2}{n}}} = 1.645 - \frac{.1}{\sqrt{\frac{p_1q_1}{n} + \frac{p_2q_2}{n}}}, \text{ or simply } 2.485 = \frac{.1}{\sqrt{\frac{p_1q_1}{n} + \frac{p_2q_2}{n}}}.$$

Using the hint, we set $p_1 = p_2 = .5$ as a "worse case scenario" and find that

$$2.485 = \frac{.1}{\sqrt{.5(.5)\left[\frac{1}{n} + \frac{1}{n}\right]}}.$$

The solution is $n = 308.76$, so the common sample size for the researcher's test should be $n = 309$.

10.41 Refer to Ex. 10.34. The rejection region, written in terms of $\bar{y}$, is

$$\left\{\frac{\bar{y}-5}{3.1/\sqrt{500}} > 1.645\right\} \Leftrightarrow \{\bar{y} > 5.228\}.$$

Then, $\beta = P(\bar{y} \leq 5.228 \mid \mu = 5.5) = P\left(Z \leq \frac{5.228-5.5}{3.1/\sqrt{500}}\right) = P(Z \leq 1.96) = .025$.

10.42 Using the sample size formula given in this section, we have

$$n = \frac{(z_\alpha + z_\beta)^2 \sigma^2}{(\mu_a - \mu_0)^2} = 607.37,$$

so a sample size of 608 will provide the desired levels.

10.43 Let μ_1 and μ_2 denote the mean dexterity scores for those students who did and did not (respectively) participate in sports.

a. For H_0: $\mu_1 - \mu_2 = 0$ vs. H_a: $\mu_1 - \mu_2 > 0$ with $\alpha = .05$, the rejection region is $\{z > 1.645\}$ and the computed test statistic is

$$z = \frac{32.19 - 31.68}{\sqrt{\frac{(4.34)^2}{37} + \frac{(4.56)^2}{37}}} = .49.$$

Thus H_0 is not rejected: there is insufficient evidence to indicate the mean dexterity score for students participating in sports is larger.

b. The rejection region, written in terms of the sample means, is

$$\overline{Y}_1 - \overline{Y}_2 > 1.645\sqrt{\frac{(4.34)^2}{37} + \frac{(4.56)^2}{37}} = 1.702.$$

Then, $\beta = P(\overline{Y}_1 - \overline{Y}_2 \le 1.702 \mid \mu_1 - \mu_2 = 3) = P\left(Z \le \frac{1.703 - 3}{\hat{\sigma}_{\overline{Y}_1 - \overline{Y}_2}}\right) = P(Z < -1.25) = .1056.$

10.44 We require $\alpha = P(\overline{Y}_1 - \overline{Y}_2 > c \mid \mu_1 - \mu_2 = 0) = P\left(Z > \frac{c - 0}{\sqrt{(\sigma_1^2 + \sigma_2^2)/n}}\right)$, so that $z_\alpha = \frac{c\sqrt{n}}{\sqrt{\sigma_1^2 + \sigma_2^2}}$. Also,

$\beta = P(\overline{Y}_1 - \overline{Y}_2 \le c \mid \mu_1 - \mu_2 = 3) = P\left(Z \le \frac{(c-3)\sqrt{n}}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right)$, so that $-z_\beta = \frac{(c-3)\sqrt{n}}{\sqrt{\sigma_1^2 + \sigma_2^2}}$. By eliminating c

in these two expressions, we have $z_\alpha \sqrt{\frac{\sigma_1^2 + \sigma_2^2}{n}} = 3 - z_\beta \sqrt{\frac{\sigma_1^2 + \sigma_2^2}{n}}$. Solving for n, we have

$$n = \frac{2(1.645)^2[(4.34)^2 + (4.56)^2]}{3^2} = 47.66.$$

A sample size of 48 will provide the required levels of α and β.

10.45 The 99% CI is $1.65 - 1.43 \pm 2.576\sqrt{\frac{(.26)^2}{30} + \frac{(.22)^2}{35}} = .22 \pm .155$ or (.065, .375). Since the interval does not contain 0, the null hypothesis should be rejected (same conclusion as Ex. 10.21).

10.46 The rejection region is $\frac{\hat{\theta} - \theta_0}{\hat{\sigma}_\theta} > z_\alpha$, which is equivalent to $\theta_0 < \hat{\theta} - z_\alpha \hat{\sigma}_{\hat{\theta}}$. The left–hand side is the 100(1 – α)% lower confidence bound for θ.

10.47 (Refer to Ex. 10.32) The 95% lower confidence bound is $.54 - 1.645\sqrt{\frac{.54(.46)}{1060}} = .5148$. Since the value $p = .50$ is less than this lower bound, it does not represent a plausible value for p. This is equivalent to stating that the hypothesis H_0: $p = .50$ should be rejected.

10.48 (Similar to Ex. 10.46) The rejection region is $\frac{\hat{\theta}-\theta_0}{\hat{\sigma}_\theta} < -z_\alpha$, which is equivalent to $\theta_0 > \hat{\theta} + z_\alpha \hat{\sigma}_{\hat{\theta}}$. The left–hand side is the 100(1 – α)% upper confidence bound for θ.

10.49 (Refer to Ex. 10.19) The upper bound is $128.6 + 1.645\left(\frac{2.1}{\sqrt{40}}\right)$ = 129.146. Since this bound is less than the hypothesized value of 130, H_0 should be rejected as in Ex. 10.19.

10.50 Let μ = mean occupancy rate. To test H_0: μ ≥ .6, H_a: μ < .6, the computed test statistic is $z = \frac{.58-.6}{.11/\sqrt{120}} = -1.99$. The p–value is given by $P(Z < -1.99) = .0233$. Since this is less than the significance level of .10, H_0 is rejected.

10.51 To test H_0: $\mu_1 - \mu_2 = 0$ vs. H_a: $\mu_1 - \mu_2 \neq 0$, where μ_1, μ_2 represent the two mean reading test scores for the two methods, the computed test statistic is

$$z = \frac{74-71}{\sqrt{\frac{9^2}{50}+\frac{10^2}{50}}} = 1.58.$$

The p–value is given by $P(|Z| > 1.58) = 2P(Z > 1.58) = .1142$, and since this is larger than α = .05, we fail to reject H_0.

10.52 The null and alternative hypotheses are H_0: $p_1 - p_2 = 0$ vs. H_a: $p_1 - p_2 > 0$, where p_1 and p_2 correspond to normal cell rates for cells treated with .6 and .7 (respectively) concentrations of actinomycin D.

a. Using the sample proportions .786 and .329, the test statistic is (refer to Ex. 10.27)
$z = \frac{.786-.329}{\sqrt{(.557)(.443)\frac{2}{70}}}$ = 5.443. The p–value is $P(Z > 5.443) \approx 0$.

b. Since the p–value is less than .05, we can reject H_0 and conclude that the normal cell rate is lower for cells exposed to the higher actinomycin D concentration.

10.53 **a.** The hypothesis of interest is H_0: $\mu_1 = 3.8$, H_a: $\mu_1 < 3.8$, where μ_1 represents the mean drop in FVC for men on the physical fitness program. With $z = -.996$, we have p–value $= P(Z < -1) = .1587$.

b. With α = .05, H_0 cannot be rejected.

c. Similarly, we have H_0: $\mu_2 = 3.1$, H_a: $\mu_2 < 3.1$. The computed test statistic is $z = -1.826$ so that the p–value is $P(Z < -1.83) = .0336$.

d. Since α = .05 is greater than the p–value, we can reject the null hypothesis and conclude that the mean drop in FVC for women is less than 3.1.

10.54 **a.** The hypotheses are H_0: $p = .85$, H_a: $p > .85$, where p = proportion of right–handed executives of large corporations. The computed test statistic is $z = 5.34$, and with α = .01, $z_{.01} = 2.326$. So, we reject H_0 and conclude that the proportion of right–handed executives at large corporations is greater than 85%

b. Since p–value = $P(Z > 5.34) < .000001$, we can safely reject H_0 for any significance level of .000001 or more. This represents strong evidence against H_0.

10.55 To test H_0: $p = .05$, H_a: $p < .05$, with $\hat{p} = 45/1124 = .040$, the computed test statistic is $z = -1.538$. Thus, p–value = $P(Z < -1.538) = .0616$ and we fail to reject H_0 with $\alpha = .01$. There is not enough evidence to conclude that the proportion of bad checks has decreased from 5%.

10.56 To test H_0: $\mu_1 - \mu_2 = 0$ vs. H_a: $\mu_1 - \mu_2 > 0$, where μ_1, μ_2 represent the two mean recovery times for treatments {no supplement} and {500 mg Vitamin C}, respectively. The computed test statistic is $z = \frac{6.9-5.8}{\sqrt{[(2.9)^2+(1.2)^2]/35}} = 2.074$. Thus, p–value = $P(Z > 2.074) = .0192$ and so the company can reject the null hypothesis at the .05 significance level conclude the Vitamin C reduces the mean recovery times.

10.57 Let p = proportion who renew. Then, the hypotheses are H_0: $p = .60$, H_a: $p \neq .60$. The sample proportion is $\hat{p} = 108/200 = .54$, and so the computed test statistic is $z = -1.732$. The p–value is given by $2P(Z < -1.732) = .0836$.

10.58 The null and alternative hypotheses are H_0: $p_1 - p_2 = 0$ vs. H_a: $p_1 - p_2 > 0$, where p_1 and p_2 correspond to, respectively, the proportions associated with groups A and B. Using the test statistic from Ex. 10.27, its computed value is $z = \frac{.74-.46}{\sqrt{.6(.4)\frac{2}{50}}} = 2.858$. Thus, p–value = $P(Z > 2.858) = .0021$. With $\alpha = .05$, we reject H0 and conclude that a greater fraction feel that a female model used in an ad increases the perceived cost of the automobile.

10.59 **a.-d.** Answers vary.

10.60 **a.-d.** Answers vary.

10.61 If the sample size is small, the test is only appropriate if the random sample was selected from a normal population. Furthermore, if the population is not normal and σ is unknown, the estimate s should only be used when the sample size is large.

10.62 For the test statistic to follow a t–distribution, the random sample should be drawn from a normal population. However, the test does work satisfactorily for similar populations that possess mound–shaped distributions.

10.63 The sample statistics are $\bar{y} = 795$, $s = 8.337$.

a. The hypotheses to be tested are H_0: $\mu = 800$, H_a: $\mu < 800$, and the computed test statistic is $t = \frac{795-800}{8.337/\sqrt{5}} = -1.341$. With 5 – 1 = 4 degrees of freedom, $-t_{.05} = -2.132$ so we fail to reject H_0 and conclude that there is not enough evidence to conclude that the process has a lower mean yield.

b. From Table 5, we find that p–value > .10 since $-t_{.10} = -1.533$.

c. Using the Applet, p–value = .1255.

d. The conclusion is the same.

10.64 The hypotheses to be tested are H_0: μ = 7, H_a: μ ≠ 7, where μ = mean beverage volume.

a. The computed test statistic is $t = \frac{7.1-7}{.12/\sqrt{10}}$ = 2.64 and with 10 –1 = 9 degrees of freedom, we find that $t_{.025}$ = 2.262. So the null hypothesis could be rejected if α = .05 (recall that this is a two–tailed test).

b. Using the Applet, $2P(T > 2.64) = 2(.01346) = .02692$.

c. Reject H_0.

10.65 The sample statistics are $\bar{y}$ = 39.556, s = 7.138.

a. To test H_0: μ = 45, H_a: μ < 45, where μ = mean cost, the computed test statistic is t = –3.24. With 18 – 1 = 17 degrees of freedom, we find that $-t_{.005} = -2.898$, so the p–value must be less than .005.

b. Using the Applet, $P(T < -3.24) = .00241$.

c. Since $t_{.025}$ = 2.110, the 95% CI is $39.556 \pm 2.11\left(\frac{7.138}{\sqrt{18}}\right)$ or (36.006, 43.106).

10.66 The sample statistics are $\bar{y}$ = 89.855, s = 14.904.

a. To test H_0: μ = 100, H_a: μ < 100, where μ = mean DL reading for current smokers, the computed test statistic is t = –3.05. With 20 – 1 = 19 degrees of freedom, we find that $-t_{.01} = -2.539$, so we reject H_0 and conclude that the mean DL reading is less than 100.

b. Using Appendix 5, $-t_{.005} = -2.861$, so p–value < .005.

c. Using the Applet, $P(T < -3.05) = .00329$.

10.67 Let μ = mean calorie content. Then, we require H_0: μ = 280, H_a: μ > 280.

a. The computed test statistic is $t = \frac{358-280}{54/\sqrt{10}}$ = 4.568. With 10 – 1 = 9 degrees of freedom, $t_{.01}$ = 2.821 so H_0 can be rejected: it is apparent that the mean calorie content is greater than advertised.

b. The 99% lower confidence bound is $358 - 2.821\frac{54}{\sqrt{10}}$ = 309.83 cal.

c. Since the value 280 is below the lower confidence bound, it is unlikely that μ = 280 (same conclusion).

10.68 The random samples are drawn independently from two normal populations with common variance.

10.69 The hypotheses are H_0: $\mu_1 - \mu_2 = 0$ vs. H_a: $\mu_1 - \mu_2 \neq 0$.

a. The computed test statistic is, where $s_p^2 = \frac{10(52)+13(71)}{23} = 62.74$, is given by

$$t = \frac{64-69}{\sqrt{62.74\left(\frac{1}{11}+\frac{1}{14}\right)}} = -1.57.$$

i. With 11 + 14 – 2 = 23 degrees of freedom, $-t_{.10} = -1.319$ and $-t_{.05} = -1.714$. Thus, since we have a two–sided alternative, .10 < p–value < .20.

ii. Using the Applet, $2P(T < -1.57) = 2(.06504) = .13008$.

b. We assumed that the two samples were selected independently from normal populations with common variance.

c. Fail to reject H_0.

10.70 **a.** The hypotheses are H_0: $\mu_1 - \mu_2 = 0$ vs. H_a: $\mu_1 - \mu_2 > 0$. The computed test statistic is t = 2.97 (here, $s_p^2 = .0001444$). With 21 degrees of freedom, $t_{.05} = 1.721$ so we reject H_0.

b. For this problem, the hypotheses are H_0: $\mu_1 - \mu_2 = .01$ vs. H_a: $\mu_1 - \mu_2 > .01$. Then, $t = \frac{(.041-.026)-.01}{\sqrt{s_p^2\left(\frac{1}{9}+\frac{1}{12}\right)}} = .989$ and p–value > .10. Using the Applet, $P(T > .989) = .16696$.

10.71 **a.** The summary statistics are: $\bar{y}_1 = 97.856$, $s_1^2 = .3403$, $\bar{y}_2 = 98.489$, $s_2^2 = .3011$. To test: H_0: $\mu_1 - \mu_2 = 0$ vs. H_a: $\mu_1 - \mu_2 \neq 0$, t = -2.3724 with 16 degrees of freedom. We have that $-t_{.01} = -2.583$, $-t_{.025} = -2.12$, so $.02 < p$–value $< .05$.

b. Using the Applet, $2P(T < -2.3724) = 2(.01527) = .03054$.

R output:

```
> t.test(temp~sex,var.equal=T)

        Two Sample t-test

data:  temp by sex
t = -2.3724, df = 16, p-value = 0.03055
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -1.19925448 -0.06741219
sample estimates:
mean in group 1 mean in group 2
       97.85556        98.48889
```

10.72 To test: H_0: $\mu_1 - \mu_2 = 0$ vs. H_a: $\mu_1 - \mu_2 \neq 0$, $t = 1.655$ with 38 degrees of freedom. Since we have that $\alpha = .05$, $t_{.025} \approx z_{.025} = 1.96$ so fail to reject H_0 and p–value $= 2P(T > 1.655) = 2(.05308) = .10616$.

10.73 **a.** To test: H_0: $\mu_1 - \mu_2 = 0$ vs. H_a: $\mu_1 - \mu_2 \neq 0$, $t = 1.92$ with 18 degrees of freedom. Since we have that $\alpha = .05$, $t_{.025} = 2.101$ so fail to reject H_0 and p–value $= 2P(T > 1.92) = 2(.03542) = .07084$.

b. To test: H_0: $\mu_1 - \mu_2 = 0$ vs. H_a: $\mu_1 - \mu_2 \neq 0$, $t = .365$ with 18 degrees of freedom. Since we have that $\alpha = .05$, $t_{.025} = 2.101$ so fail to reject H_0 and p–value $= 2P(T > .365) = 2(.35968) = .71936$.

10.74 The hypotheses are H_0: $\mu = 6$ vs. H_a: $\mu < 6$ and the computed test statistic is $t = 1.62$ with 11 degrees of freedom (note that here $\bar{y} = 9$, so H_0 could never be rejected). With α = .05, the critical value is $-t_{.05} = -1.796$ so fail to reject H_0.

10.75 Define μ = mean trap weight. The sample statistics are $\bar{y}$ = 28.935, s = 9.507. To test H_0: μ = 30.31 vs. H_a: μ < 30.31, t = –.647 with 19 degrees of freedom. With α = .05, the critical value is $-t_{.05} = -1.729$ so fail to reject H_0: we cannot conclude that the mean trap weight has decreased. R output:

```
> t.test(lobster,mu=30.31, alt="less")

        One Sample t-test

data:  lobster
t = -0.6468, df = 19, p-value = 0.2628
alternative hypothesis: true mean is less than 30.31
95 percent confidence interval:
     -Inf 32.61098
```

10.76 **a.** To test H_0: $\mu_1 - \mu_2 = 0$ vs. H_a: $\mu_1 - \mu_2 > 0$, where μ_1, μ_2 represent mean plaque measurements for the control and antiplaque groups, respectively.

b. The pooled sample variance is $s_p^2 = \frac{6(.32)^2 + 6(.32)^2}{12} = .1024$ and the computed test statistic is $t = \frac{1.26 - .78}{\sqrt{.1024\left(\frac{2}{7}\right)}} = 2.806$ with 12 degrees of freedom. Since α = .05, $t_{.05} = 1.782$ and H_0 is rejected: there is evidence that the antiplaque rinse reduces the mean plaque measurement.

c. With $t_{.01} = 2.681$ and $t_{.005} = 3.005$, .005 < p–value < .01 (exact: .00793).

10.77 **a.** To test: H_0: $\mu_1 - \mu_2 = 0$ vs. H_a: $\mu_1 - \mu_2 \neq 0$, where μ_1, μ_2 are the mean verbal SAT scores for students intending to major in engineering and language (respectively), the pooled sample variance is $s_p^2 = \frac{14(42)^2 + 14(45)^2}{28} = 1894.5$ and the computed test statistic is $t = \frac{446 - 534}{\sqrt{1894.5\left(\frac{2}{15}\right)}} = -5.54$ with 28 degrees of freedom. Since $-t_{.005} = -2.763$, we can reject H_0 and p–value < .01 (exact: 6.35375e-06).

b. Yes, the CI approach agrees.

c. To test: H_0: $\mu_1 - \mu_2 = 0$ vs. H_a: $\mu_1 - \mu_2 \neq 0$, where μ_1, μ_2 are the mean math SAT scores for students intending to major in engineering and language (respectively), the pooled sample variance is $s_p^2 = \frac{14(57)^2 + 14(52)^2}{28} = 2976.5$ and the computed test statistic is $t = \frac{548 - 517}{\sqrt{2976.5\left(\frac{2}{15}\right)}} = 1.56$ with 28 degrees of freedom. From Table 5, .10 < p–value < .20 (exact: 0.1299926).

d. Yes, the CI approach agrees.

10.78 **a.** We can find $P(Y > 1000) = P(Z > \frac{1000-800}{40}) = P(Z > 5) \approx 0$, so it is very unlikely that the force is greater than 1000 lbs.

b. Since $n = 40$, the large–sample test for a mean can be used: H_0: $\mu = 800$ vs. H_a: $\mu > 800$ and the test statistic is $z = \frac{825-800}{\sqrt{2350/40}} = 3.262$. With p–value $= P(Z > 3.262) < .00135$, we reject H_0.

c. Note that if $\sigma = 40$, $\sigma^2 = 1600$. To test: H_0: $\sigma^2 = 1600$ vs. H_a: $\sigma^2 > 1600$. The test statistic is $\chi^2 = \frac{39(2350)}{1600} = 57.281$. With $40 - 1 = 39$ degrees of freedom (approximated with 40 degrees of freedom in Table 6), $\chi^2_{.05} = 55.7585$. So, we can reject H_0 and conclude there is sufficient evidence that σ exceeds 40.

10.79 **a.** The hypotheses are: H_0: $\sigma^2 = .01$ vs. H_a: $\sigma^2 > .01$. The test statistic is $\chi^2 = \frac{7(.018)}{.01} = 12.6$ with 7 degrees of freedom. With $\alpha = .05$, $\chi^2_{.05} = 14.07$ so we fail to reject H_0. We must assume the random sample of carton weights were drawn from a normal population.

b. i. Using Table 6, $.05 < p$–value $< .10$.
ii. Using the Applet, $P(\chi^2 > 12.6) = .08248$.

10.80 The two random samples must be independently drawn from normal populations.

10.81 For this exercise, refer to Ex. 8.125.

a. The rejection region is $\left\{S_1^2/S_2^2 > F^{\nu_1}_{\nu_2,\alpha/2}\right\} \cup \left\{S_1^2/S_2^2 < \left(F^{\nu_2}_{\nu_1,\alpha/2}\right)^{-1}\right\}$. If the reciprocal is taken in the second inequality, we have $S_2^2/S_1^2 > F^{\nu_1}_{\nu_2,\alpha/2}$.

b. $P\left(S_L^2/S_S^2 > F^{\nu_L}_{\nu_S,\alpha/2}\right) = P\left(S_1^2/S_2^2 > F^{\nu_1}_{\nu_2,\alpha/2}\right) + P\left(S_2^2/S_1^2 > F^{\nu_2}_{\nu_1,\alpha/2}\right) = \alpha$, by part a.

10.82 **a.** Let σ_1^2, σ_2^2 denote the variances for compartment pressure for resting runners and cyclists, respectively. To test H_0: $\sigma_1^2 = \sigma_2^2$ vs. H_a: $\sigma_1^2 \neq \sigma_2^2$, the computed test statistic is $F = (3.98)^2/(3.92)^2 = 1.03$. With $\alpha = .05$, $F^9_{9,.025} = 4.03$ and we fail to reject H_0.

b. i. From Table 7, p–value $> .1$.
ii. Using the Applet, $2P(F > 1.03) = 2(.4828) = .9656$.

c. Let σ_1^2, σ_2^2 denote the population variances for compartment pressure for 80% maximal O_2 consumption for runners and cyclists, respectively. To test H_0: $\sigma_1^2 = \sigma_2^2$ vs. H_a: $\sigma_1^2 \neq \sigma_2^2$, the computed test statistic is $F = (16.9)^2/(4.67)^2 = 13.096$ and we reject H_0: the is sufficient evidence to claim a difference in variability.

d. i. From Table 7, p–value $< .005$.
ii. Using the Applet, $2P(F > 13.096) = 2(.00036) = .00072$.

10.83 **a.** The manager of the dairy is concerned with determining if there is a *difference* in the two variances, so a two–sided alternative should be used.

b. The salesman for company *A* would prefer $H_a: \sigma_1^2 < \sigma_2^2$, since if this hypothesis is accepted, the manager would choose company *A*'s machine (since it has a smaller variance).

c. For similar logic used in part b, the salesman for company B would prefer $H_a: \sigma_1^2 > \sigma_2^2$.

10.84 Let σ_1^2, σ_2^2 denote the variances for measurements corresponding to 95% ethanol and 20% bleach, respectively. The desired hypothesis test is $H_0: \sigma_1^2 = \sigma_2^2$ vs. $H_a: \sigma_1^2 \neq \sigma_2^2$ and the computed test statistic is $F = \left(2.78095/.17143\right) = 16.222$.

a. i. With 14 numerator and 14 denominator degrees of freedom, we can approximate the critical value in Table 7 by $F_{14,.005}^{15} = 4.25$, so *p*–value < .01 (two–tailed test).

ii. Using the Applet, $2P(F > 16.222) \approx 0$.

b. We would reject H_0 and conclude the variances are different.

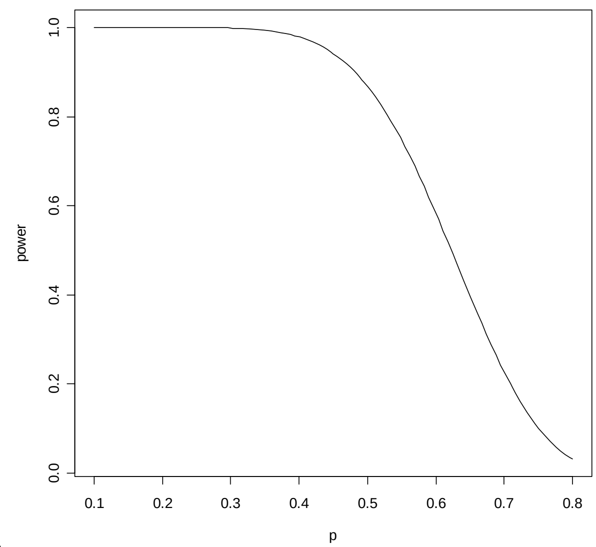
10.85 Since $(.7)^2 = .49$, the hypotheses are: H_0: $\sigma^2 = .49$ vs. H_a: $\sigma^2 > .49$. The sample variance $s^2 = 3.667$ so the computed test statistic is $\chi^2 = \frac{3(3.667)}{.49} = 22.45$ with 3 degrees of freedom. Since $\chi_{.05}^2 = 12.831$, *p*–value < .005 (exact: .00010).

10.86 The hypotheses are: H_0: $\sigma^2 = 100$ vs. H_a: $\sigma^2 > 100$. The computed test statistic is $\chi^2 = \frac{19(144)}{100} = 27.36$. With $\alpha = .01$, $\chi_{.01}^2 = 36.1908$ so we fail to reject H_0: there is not enough evidence to conclude the variability for the new test is higher than the standard.

10.87 Refer to Ex. 10.87. Here, the test statistic is $(.017)^2/(.006)^2 = 8.03$ and the critical value is $F_{12,.05}^{9} = 2.80$. Thus, we can support the claim that the variance in measurements of DDT levels for juveniles is greater than it is for nestlings.

10.88 Refer to Ex. 10.2. Table 1 in Appendix III is used to find the binomial probabilities.

a. power(.4) = $P(Y \leq 12 \mid p = .4) = .979$.
b. power(.5) = $P(Y \leq 12 \mid p = .5) = .86$
c. power(.6) = $P(Y \leq 12 \mid p = .6) = .584$.
d. power(.7) = $P(Y \leq 12 \mid p = .7) = .228$



e. The power function is above.

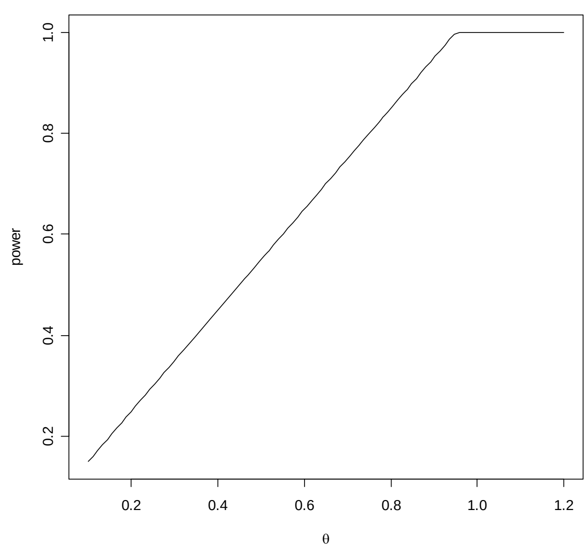
10.89 Refer to Ex. 10.5: $Y_1 \sim \text{Unif}(\theta, \theta + 1)$.

a. $\theta = .1$, so $Y_1 \sim \text{Unif}(.1, 1.1)$ and power$(.1) = P(Y_1 > .95) = \int_{.95}^{1.1} dy = .15$

b. $\theta = .4$: power$(.4) = P(Y > .95) = .45$

c. $\theta = .7$: power$(.7) = P(Y > .95) = .75$

d. $\theta = 1$: power$(1) = P(Y > .95) = 1$



e. The power function is above.

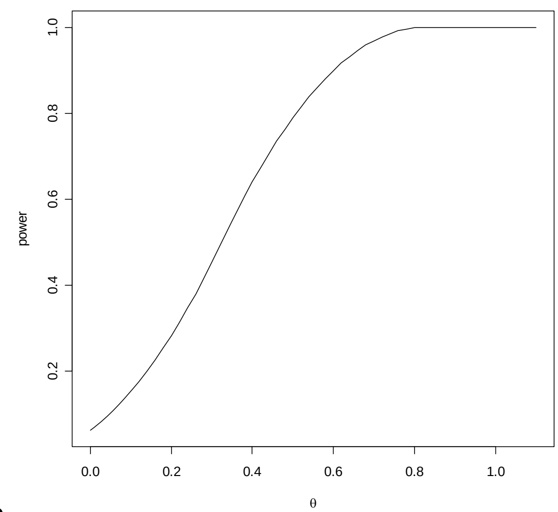
10.90 Following Ex. 10.5, the distribution function for Test 2, where $U = Y_1 + Y_2$, is

$$F_U(u) = \begin{cases} 0 & u < 0 \\ .5u^2 & 0 \le u \le 1 \\ 2u - .5u^2 - 1 & 1 < u \le 2 \\ 1 & u > 2 \end{cases}.$$

The test rejects when $U > 1.684$. The power function is given by:

$$\text{power}(\theta) = P_\theta(Y_1 + Y_2 > 1.684) = P(Y_1 + Y_2 - 2\theta > 1.684 - 2\theta)$$
$$= P(U > 1.684 - 2\theta) = 1 - F_U(1.684 - 2\theta).$$

a. power(.1) = 1 – F_U(1.483) = .133 power(.4) = 1 – F_U(.884) = .609
power(.7) = 1 – F_U(.284) = .960 power(1) = 1 – F_U(–.316) = 1.



b. The power function is above.

c. Test 2 is a more powerful test.

10.91 Refer to Example 10.23 in the text. The hypotheses are H_0: μ = 7 vs. H_a: μ > 7.

a. The uniformly most powerful test is identically the Z–test from Section 10.3. The rejection region is: reject if $Z = \frac{\bar{Y}-7}{\sqrt{5/20}} > z_{.05} = 1.645$, or equivalently, reject if

$\bar{Y} > 1.645\sqrt{.25} + 7 = 7.82$.

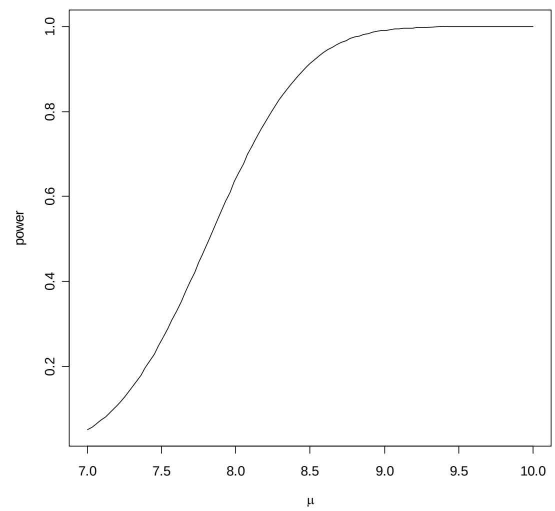
b. The power function is: power(μ) = $P(\bar{Y} > 7.82 \mid \mu) = P\left(Z > \frac{7.82-\mu}{\sqrt{5/20}}\right)$. Thus:

power(7.5) = $P(\bar{Y} > 7.82 \mid 7.5)$ = $P(Z > .64)$ = .2611.

power(8.0) = $P(\bar{Y} > 7.82 \mid 8.0)$ = $P(Z > –.36)$ = .6406.

power(8.5) = $P(\bar{Y} > 7.82 \mid 8.5)$ = $P(Z > –1.36)$ = .9131

power(9.0) = $P(\bar{Y} > 7.82 \mid 9.0)$ = $P(Z > –2.36)$ = .9909.



c. The power function is above.

10.92 Following Ex. 10.91, we require power(8) = $P(\overline{Y} > 7.82 \mid 8) = P\left(Z > \frac{7.82-8}{\sqrt{5/n}}\right)$ = .80. Thus, $\frac{7.82-8}{\sqrt{5/n}} = z_{.80} = -.84$. The solution is n = 108.89, or 109 observations must be taken.

10.93 Using the sample size formula from the end of Section 10.4, we have $n = \frac{(1.96+1.96)^2(25)}{(10-5)^2}$ = 15.3664, so 16 observations should be taken.

10.94 The most powerful test for H_0: $\sigma^2 = \sigma_0^2$ vs. H_a: $\sigma^2 = \sigma_1^2$, $\sigma_1^2 > \sigma_0^2$, is based on the likelihood ratio:

$$\frac{L(\sigma_0^2)}{L(\sigma_1^2)} = \left(\frac{\sigma_1}{\sigma_0}\right)^n \exp\left[-\left(\frac{\sigma_1^2 - \sigma_0^2}{2\sigma_0^2\sigma_1^2}\sum_{i=1}^n (y_i - \mu)^2\right)\right] < k.$$

This simplifies to

$$T = \sum_{i=1}^n (y_i - \mu)^2 > \left[n\ln\left(\frac{\sigma_1}{\sigma_0}\right) - \ln k\right]\frac{2\sigma_0^2\sigma_1^2}{\sigma_1^2 - \sigma_0^2} = c,$$

which is to say we should reject if the statistic T is large. To find a rejection region of size α, note that

$\frac{T}{\sigma_0^2} = \frac{\sum_{i=1}^n (Y_i - \mu)^2}{\sigma_0^2}$ has a chi–square distribution with n degrees of freedom. Thus, the most powerful test is equivalent to the chi–square test, and this test is UMP since the RR is the same for any $\sigma_1^2 > \sigma_0^2$.

10.95 **a.** To test H_0: $\theta = \theta_0$ vs. H_a: $\theta = \theta_a$, $\theta_0 < \theta_a$, the best test is

$$\frac{L(\theta_0)}{L(\theta_a)} = \left(\frac{\theta_a}{\theta_0}\right)^{12} \exp\left[-\left(\frac{1}{\theta_0} - \frac{1}{\theta_a}\right)\sum_{i=1}^4 y_i\right] < k.$$

This simplifies to

$$T = \sum_{i=1}^{4} y_i > \ln k\left(\frac{\theta_0}{\theta_a}\right)^{12}\left[\frac{1}{\theta_0} - \frac{1}{\theta_a}\right]^{-1} = c\,,$$

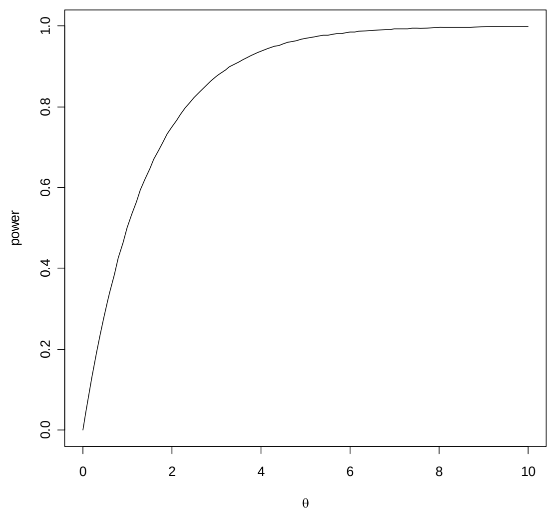
so H_0 should be rejected if T is large. Under H_0, Y has a gamma distribution with a shape parameter of 3 and scale parameter θ_0. Likewise, T is gamma with shape parameter of 12 and scale parameter θ_0, and $2T/\theta_0$ is chi–square with 24 degrees of freedom. The critical region can be written as

$$\frac{2T}{\theta_0} = \frac{2\sum_{i=1}^{4} Y_i}{\theta_0} > \frac{2c}{\theta_0} = c_1\,,$$

where c_1 will be chosen (from the chi–square distribution) so that the test is of size α.

b. Since the critical region doesn't depend on any specific $\theta_a < \theta_0$, the test is UMP.

10.96 **a.** The power function is given by power(θ) = $\int_{.5}^{1} \theta y^{\theta-1} dy = 1 - .5^{\theta}$. The power function is graphed below.



b. To test H_0: θ = 1 vs. H_a: $\theta = \theta_a$, $1 < \theta_a$, the likelihood ratio is

$$\frac{L(1)}{L(\theta_a)} = \frac{1}{\theta_a y^{\theta_a - 1}} < k.$$

This simplifies to

$$y > \left(\frac{1}{\theta_a k}\right)^{\frac{1}{\theta_a - 1}} = c\,,$$

where c is chosen so that the test is of size α. This is given by

$$P(Y \ge c \mid \theta = 1) = \int_c^1 dy = 1 - c = \alpha\,,$$

so that $c = 1 - \alpha$. Since the RR does not depend on a specific $\theta_a > 1$, it is UMP.

10.97 Note that (N_1, N_2, N_3) is trinomial (multinomial with $k = 3$) with cell probabilities as given in the table.

a. The likelihood function is simply the probability mass function for the trinomial:

$$L(\theta) = \binom{n}{n_1\ n_2\ n_3}\theta^{2n_1}\left[2\theta(1-\theta)\right]^{n_2}(1-\theta)^{2n_3}, \ 0 < \theta < 1,\ n = n_1 + n_2 + n_3.$$

b. Using part a, the best test for testing H_0: $\theta = \theta_0$ vs. H_a: $\theta = \theta_a$, $\theta_0 < \theta_a$, is

$$\frac{L(\theta_0)}{L(\theta_a)} = \left(\frac{\theta_0}{\theta_a}\right)^{2n_1+n_2}\left(\frac{1-\theta_0}{1-\theta_a}\right)^{n_2+2n_3} < k.$$

Since we have that $n_2 + 2n_3 = 2n - (2n_1 + n_2)$, the RR can be specified for certain values of $S = 2N_1 + N_2$. Specifically, the log–likelihood ratio is

$$s\ln\left(\frac{\theta_0}{\theta_a}\right) + (2n-s)\ln\left(\frac{1-\theta_0}{1-\theta_a}\right) < \ln k,$$

or equivalently

$$s > \left[\ln k - 2n\ln\left(\frac{1-\theta_0}{1-\theta_a}\right)\right] \times \left[\ln\left(\frac{\theta_0(1-\theta_a)}{\theta_a(1-\theta_0)}\right)\right]^{-1} = c.$$

So, the rejection region is given by $\{S = 2N_1 + N_2 > c\}$.

c. To find a size α rejection region, the distribution of (N_1, N_2, N_3) is specified and with $S = 2N_1 + N_2$, a null distribution for S can be found and a critical value specified such that $P(S \ge c \mid \theta_0) = \alpha$.

d. Since the RR doesn't depend on a specific $\theta_a > \theta_0$, it is a UMP test.

10.98 The density function that for the Weibull with shape parameter m and scale parameter θ.

a. The best test for testing H_0: $\theta = \theta_0$ vs. H_a: $\theta = \theta_a$, where $\theta_0 < \theta_a$, is

$$\frac{L(\theta_0)}{L(\theta_a)} = \left(\frac{\theta_a}{\theta_0}\right)^n \exp\left[-\left(\frac{1}{\theta_0} - \frac{1}{\theta_a}\right)\sum_{i=1}^n y_i^m\right] < k,$$

This simplifies to

$$\sum_{i=1}^n y_i^m > -\left[\ln k + n\ln\left(\frac{\theta_0}{\theta_a}\right)\right] \times \left[\frac{1}{\theta_0} - \frac{1}{\theta_a}\right]^{-1} = c.$$

So, the RR has the form $\left\{T = \sum_{i=1}^m Y_i^m > c\right\}$, where c is chosen so the RR is of size α. To do so, note that the distribution of Y^m is exponential so that under H_0,

$$\frac{2T}{\theta_0} = \frac{2\sum_{i=1}^n Y_i^m}{\theta_0} > \frac{2c}{\theta_0}$$

is chi–square with $2n$ degrees of freedom. So, the critical value can be selected from the chi–square distribution and this does not depend on the specific $\theta_a > \theta_0$, so the test is UMP.

b. When H_0 is true, $T/50$ is chi–square with $2n$ degrees of freedom. Thus, $\chi^2_{.05}$ can be selected from this distribution so that the RR is $\{T/50 > \chi^2_{.05}\}$ and the test is of size $\alpha = .05$. If H_a is true, $T/200$ is chi–square with $2n$ degrees of freedom. Thus, we require

$$\beta = P(T/50 \le \chi^2_{.05} \mid \theta = 400) = P(T/200 \le \tfrac{1}{4}\chi^2_{.05} \mid \theta = 400) = P(\chi^2 \le \tfrac{1}{4}\chi^2_{.05}) = .05.$$

Thus, we have that $\frac{1}{4}\chi^2_{.05} = \chi^2_{.95}$. From Table 6 in Appendix III, it is found that the degrees of freedom necessary for this equality is $12 = 2n$, so $n = 6$.

10.99 **a.** The best test is

$$\frac{L(\lambda_0)}{L(\lambda_a)} = \left(\frac{\lambda_0}{\lambda_a}\right)^T \exp[n(\lambda_a - \lambda_0)] < k,$$

where $T = \sum_{i=1}^n Y_i$. This simplifies to

$$T > \frac{\ln k - n(\lambda_a - \lambda_0)}{\ln(\lambda_0/\lambda_a)} = c,$$

and c is chosen so that the test is of size α.

b. Since under H_0 $T = \sum_{i=1}^n Y_i$ is Poisson with mean $n\lambda$, c can be selected such that

$$P(T > c \mid \lambda = \lambda_0) = \alpha.$$

c. Since this critical value does not depend on the specific $\lambda_a > \lambda_0$, so the test is UMP.

d. It is easily seen that the UMP test is: reject if $T < k'$.

10.100 Since **X** and **Y** are independent, the likelihood function is the product of all marginal mass function. The best test is given by

$$\frac{L_0}{L_1} = \frac{2^{\Sigma x_i + \Sigma y_i}\exp(-2m - 2n)}{\left(\frac{1}{2}\right)^{\Sigma x_i} 3^{\Sigma y_i}\exp(-m/2 - 3n)} = 4^{\Sigma x_i}\left(\tfrac{2}{3}\right)^{\Sigma y_i}\exp(-3m/2 + n) < k.$$

This simplifies to

$$(\ln 4)\sum_{i=1}^m x_i + \ln(2/3)\sum_{i=1}^n y_i < k',$$

and k' is chosen so that the test is of size α.

10.101 **a.** To test H_0: $\theta = \theta_0$ vs. H_a: $\theta = \theta_a$, where $\theta_a < \theta_0$, the best test is

$$\frac{L(\theta_0)}{L(\theta_a)} = \left(\frac{\theta_a}{\theta_0}\right)^n \exp\left[-\left(\frac{1}{\theta_0} - \frac{1}{\theta_a}\right)\sum_{i=1}^n y_i\right] < k.$$

Equivalently, this is

$$\sum_{i=1}^{n} y_i < \left[n \ln\left(\frac{\theta_0}{\theta_a}\right) + \ln k \right] \times \left[\frac{1}{\theta_a} - \frac{1}{\theta_0} \right]^{-1} = c \,,$$

and c is chosen so that the test is of size α (the chi–square distribution can be used – see Ex. 10.95).

b. Since the RR does not depend on a specific value of $\theta_a < \theta_0$, it is a UMP test.

10.102 **a.** The likelihood function is the product of the mass functions:

$$L(p) = p^{\Sigma y_i}(1-p)^{n-\Sigma y_i}\,.$$

i. It follows that the likelihood ratio is

$$\frac{L(p_0)}{L(p_a)} = \frac{p_0^{\Sigma y_i}(1-p_0)^{n-\Sigma y_i}}{p_a^{\Sigma y_i}(1-p_a)^{n-\Sigma y_i}} = \left(\frac{p_0(1-p_a)}{p_a(1-p_0)}\right)^{\Sigma y_i}\left(\frac{1-p_0}{1-p_a}\right)^{n}.$$

ii. Simplifying the above, the test rejects when

$$\sum_{i=1}^{n} y_i \ln\left(\frac{p_0(1-p_a)}{p_a(1-p_0)}\right) + n\ln\left(\frac{1-p_0}{1-p_a}\right) < \ln k.$$

Equivalently, this is

$$\sum_{i=1}^{n} y_i > \left[\ln k - n\ln\left(\frac{1-p_0}{1-p_a}\right)\right] \times \left[\ln\left(\frac{p_0(1-p_a)}{p_a(1-p_0)}\right)\right]^{-1} = c.$$

iii. The rejection region is of the form $\{\sum_{i=1}^{n} y_i > c\}$.

b. For a size α test, the critical value c is such that $P(\sum_{i=1}^{n} Y_i > c \mid p_0) = \alpha$. Under H_0, $\sum_{i=1}^{n} Y_i$ is binomial with parameters n and p_0.

c. Since the critical value can be specified without regard to a specific value of p_a, this is the UMP test.

10.103 Refer to Section 6.7 and 9.7 for this problem.

a. The likelihood function is $L(\theta) = \theta^{-n} I_{0,\theta}(y_{(n)})$. To test H_0: $\theta = \theta_0$ vs. H_a: $\theta = \theta_a$, where $\theta_a < \theta_0$, the best test is

$$\frac{L(\theta_0)}{L(\theta_a)} = \left(\frac{\theta_a}{\theta_0}\right)^n \frac{I_{0,\theta_0}(y_{(n)})}{I_{0,\theta_a}(y_{(n)})} < k.$$

So, the test only depends on the value of the largest order statistic $Y_{(n)}$, and the test rejects whenever $Y_{(n)}$ is small. The density function for $Y_{(n)}$ is $g_n(y) = ny^{n-1}\theta^{-n}$, for $0 \le y \le \theta$. For a size α test, select c such that

$$\alpha = P(Y_{(n)} < c \mid \theta = \theta_0) = \int_0^c ny^{n-1}\theta_0^{-n}dy = \frac{c^n}{\theta_0^n},$$

so $c = \theta_0\alpha^{1/n}$. So, the RR is $\{Y_{(n)} < \theta_0\alpha^{1/n}\}$.

b. Since the RR does not depend on the specific value of $\theta_a < \theta_0$, it is UMP.

10.104 Refer to Ex. 10.103.

a. As in Ex. 10.103, the test can be based on $Y_{(n)}$. In the case, the rejection region is of the form $\{Y_{(n)} > c\}$. For a size α test select c such that

$$\alpha = P(Y_{(n)} > c \mid \theta = \theta_0) = \int_c^{\theta_0} ny^{n-1}\theta_0^{-n}dy = 1 - \frac{c^n}{\theta_0^n},$$

so $c = \theta_0(1 - \alpha)^{1/n}$.

b. As in Ex. 10.103, the test is UMP.

c. It is not unique. Another interval for the RR can be selected so that it is of size α and the power is the same as in part a and independent of the interval. Example: choose the rejection region $C = (a,b) \cup (\theta_0, \infty)$, where $(a,b) \subset (0, \theta_0)$. Then,

$$\alpha = P(a < Y_{(n)} < b \mid \theta_0) = \frac{b^n - a^n}{\theta_0^n},$$

The power of this test is given by

$$P(a < Y_{(n)} < b \mid \theta_a) + P(Y_{(n)} > \theta_0 \mid \theta_a) = \frac{b^n - a^n}{\theta_a^n} + \frac{\theta_a^n - \theta_0^n}{\theta_a^n} = (\alpha - 1)\frac{\theta_0^n}{\theta_a^n} + 1,$$

which is independent of the interval (a, b) and has the same power as in part a.

10.105 The hypotheses are H_0: $\sigma^2 = \sigma_0^2$ vs. H_a: $\sigma^2 > \sigma_0^2$. The null hypothesis specifies $\Omega_0 = \{\sigma^2 : \sigma^2 = \sigma_0^2\}$, so in this restricted space the MLEs are $\hat{\mu} = \bar{y}$, σ_0^2. For the unrestricted space Ω, the MLEs are $\hat{\mu} = \bar{y}$, while

$$\hat{\sigma}^2 = \max\left[\sigma_0^2, \frac{1}{n}\sum_{i=1}^n (y_i - \bar{y})^2\right].$$

The likelihood ratio statistic is

$$\lambda = \frac{L(\hat{\Omega}_0)}{L(\hat{\Omega})} = \left(\frac{\hat{\sigma}^2}{\sigma_0^2}\right)^{n/2} \exp\left[-\frac{\sum_{i=1}^n (y_i - \bar{y})^2}{2\sigma_0^2} + \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{2\hat{\sigma}^2}\right].$$

If $\hat{\sigma}^2 = \sigma_0^2$, $\lambda = 1$. If $\hat{\sigma}^2 > \sigma_0^2$,

$$\lambda = \frac{L(\hat{\Omega}_0)}{L(\hat{\Omega})} = \left(\frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n\sigma_0^2}\right)^{n/2} \exp\left[-\frac{\sum_{i=1}^n (y_i - \bar{y})^2}{2\sigma_0^2} + \frac{n}{2}\right],$$

and H_0 is rejected when $\lambda \le k$. This test is a function of the chi–square test statistic $\chi^2 = (n-1)S^2/\sigma_0^2$ and since the function is monotonically decreasing function of χ^2, the test $\lambda \le k$ is equivalent to $\chi^2 \ge c$, where c is chosen so that the test is of size α.

10.106 The hypothesis of interest is H_0: $p_1 = p_2 = p_3 = p_4 = p$. The likelihood function is

$$L(\boldsymbol{p}) = \prod_{i=1}^{4} \binom{200}{y_i} p_i^{y_i} (1-p_i)^{200-y_i}.$$

Under H_0, it is easy to verify that the MLE of p is $\hat{p} = \sum_{i=1}^{4} y_i / 800$. For the unrestricted space, $\hat{p}_i = y_i / 200$ for i = 1, 2, 3, 4. Then, the likelihood ratio statistic is

$$\lambda = \frac{\left(\frac{\Sigma y_i}{800}\right)^{\Sigma y_i} \left(1 - \frac{\Sigma y_i}{800}\right)^{800-\Sigma y_i}}{\prod_{i=1}^{4} \left(\frac{y_i}{200}\right)^{y_i} \left(1 - \frac{y_i}{200}\right)^{200-y_i}}.$$

Since the sample sizes are large, Theorem 10.2 can be applied so that $-2\ln\lambda$ is approximately distributed as chi–square with 3 degrees of freedom and we reject H_0 if $-2\ln\lambda > \chi^2_{.05} = 7.81$. For the data in this exercise, y_1 = 76, y_2 = 53, y_3 = 59, and y_4 = 48. Thus, $-2\ln\lambda$ = 10.54 and we reject H_0: the fraction of voters favoring candidate A is not the sample in all four wards.

10.107 Let $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_m$ denote the two samples. Under H_0, the quantity

$$V = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2 + \sum_{i=1}^{n}(Y_i - \overline{Y})^2}{\sigma_0^2} = \frac{(n-1)S_1^2 + (m-1)S_2^2}{\sigma_0^2}$$

has a chi–square distribution with $n + m - 2$ degrees of freedom. If H_a is true, then both S_1^2 and S_2^2 will tend to be larger than σ_0^2. Under H_0, the maximized likelihood is

$$L(\hat{\Omega}_0) = \frac{1}{(2\pi)^{n/2}} \frac{1}{\sigma_0^n} \exp\left(-\tfrac{1}{2}V\right).$$

In the unrestricted space, the likelihood is either maximized at σ_0 or σ_a. For the former, the likelihood ratio will be equal to 1. But, for $k < 1$, $\frac{L(\hat{\Omega}_0)}{L(\hat{\Omega})} < k$ only if $\hat{\sigma} = \sigma_a$. In this case,

$$\lambda = \frac{L(\hat{\Omega}_0)}{L(\hat{\Omega})} = \left(\frac{\sigma_a}{\sigma_0}\right)^n \exp\left[-\tfrac{1}{2}V + \tfrac{1}{2}V\left(\tfrac{\sigma_0^2}{\sigma_a^2}\right)\right] = \left(\frac{\sigma_a}{\sigma_0}\right)^n \exp\left[-\tfrac{1}{2}V\left(1 - \tfrac{\sigma_0^2}{\sigma_a^2}\right)\right],$$

which is a decreasing function of V. Thus, we reject H_0 if V is too large, and the rejection region is $\{V > \chi^2_\alpha\}$.

10.108 The likelihood is the product of all $n = n_1 + n_2 + n_3$ normal densities:

$$L(\Theta) = \tfrac{1}{(2\pi)^n} \tfrac{1}{\sigma_1^{n_1}\sigma_2^{n_2}\sigma_3^{n_3}} \exp\left\{-\tfrac{1}{2}\sum_{i=1}^{n_1}\left(\tfrac{x_i-\mu_1}{\sigma_1}\right)^2 - \tfrac{1}{2}\sum_{i=1}^{n_2}\left(\tfrac{y_i-\mu_2}{\sigma_2}\right)^2 - \tfrac{1}{2}\sum_{i=1}^{n_3}\left(\tfrac{w_i-\mu_3}{\sigma_3}\right)^2\right\}$$

a. Under H_a (unrestricted), the MLEs for the parameters are:

$\hat{\mu}_1 = \overline{X}$, $\hat{\mu}_2 = \overline{Y}$, $\hat{\mu}_3 = \overline{W}$, $\hat{\sigma}_1^2 = \frac{1}{n_1}\sum_{i=1}^{n_1}(X_i - \overline{X})^2$, $\hat{\sigma}_2^2$, $\hat{\sigma}_3^2$ defined similarly.

Under H_0, $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma^2$ and the MLEs are

$$\hat{\mu}_1 = \overline{X},\ \hat{\mu}_2 = \overline{Y},\ \hat{\mu}_3 = \overline{W},\ \hat{\sigma}^2 = \frac{n_1\hat{\sigma}_1^2 + n_2\hat{\sigma}_2^2 + n_3\hat{\sigma}_3^2}{n}.$$

By defining the LRT, it is found to be equal to

$$\lambda = \frac{\left(\hat{\sigma}_1^2\right)^{n_1/2}\left(\hat{\sigma}_2^2\right)^{n_2/2}\left(\hat{\sigma}_3^2\right)^{n_3/2}}{\left(\hat{\sigma}^2\right)^{n/2}}.$$

b. For large values of n_1, n_2, and n_3, the quantity $-2\ln\lambda$ is approximately chi–square with 3–1=2 degrees of freedom. So, the rejection region is: $-2\ln\lambda > \chi^2_{.05} = 5.99$.

10.109 The likelihood function is $L(\Theta) = \frac{1}{\theta_1^m\theta_2^n}\exp\left[-\left(\sum_{i=1}^m x_i/\theta_1 + \sum_{i=1}^n y_i/\theta_2\right)\right]$.

a. Under H_a (unrestricted), the MLEs for the parameters are:

$$\hat{\theta}_1 = \overline{X},\ \hat{\theta}_2 = \overline{Y}.$$

Under H_0, $\theta_1 = \theta_2 = \theta$ and the MLE is

$$\hat{\theta} = (m\overline{X} + n\overline{Y})/(m+n).$$

By defining the LRT, it is found to be equal to

$$\lambda = \frac{\overline{X}^m\overline{Y}^n}{\left(\frac{m\overline{X}+n\overline{Y}}{m+n}\right)^{m+n}}$$

b. Since $2\sum_{i=1}^m X_i/\theta_1$ is chi–square with $2m$ degrees of freedom and $2\sum_{i=1}^n Y_i/\theta_2$ is chi–square with $2n$ degrees of freedom, the distribution of the quantity under H_0

$$F = \frac{\dfrac{\left(2\sum_{i=1}^m X_i/\theta\right)}{2m}}{\dfrac{\left(2\sum_{i=1}^n Y_i/\theta\right)}{2n}} = \frac{\overline{X}}{\overline{Y}}$$

has an F–distribution with $2m$ numerator and $2n$ denominator degrees of freedom. This test can be seen to be equivalent to the LRT in part a by writing

$$\lambda = \frac{\overline{X}^m\overline{Y}^n}{\left(\frac{m\overline{X}+n\overline{Y}}{m+n}\right)^{m+n}} = \left[\frac{m\overline{X}+n\overline{Y}}{\overline{X}(m+n)}\right]^{-m}\left[\frac{m\overline{X}+n\overline{Y}}{\overline{Y}(m+n)}\right]^{-n} = \left[\frac{m}{m+n} + \frac{n}{F(m+n)}\right]^{-m}\left[\frac{m}{m+n}F + \frac{n}{m+n}\right]^{-n}.$$

So, λ is small if F is too large or too small. Thus, the rejection region is equivalent to $F > c_1$ and $F < c_2$, where c_1 and c_2 are chosen so that the test is of size α.

10.110 This is easily proven by using Theorem 9.4: write the likelihood function as a function of the sufficient statistic, so therefore the LRT must also only be a function of the sufficient statistic.

10.111 **a.** Under H_0, the likelihood is maximized at θ_0. Under the alternative (unrestricted) hypothesis, the likelihood is maximized at either θ_0 or θ_a. Thus, $L(\hat{\Omega}_0) = L(\theta_0)$ and $L(\hat{\Omega}) = \max\{L(\theta_0), L(\theta_a)\}$. Thus,

$$\lambda = \frac{L(\hat{\Omega}_0)}{L(\hat{\Omega})} = \frac{L(\theta_0)}{\max\{L(\theta_0), L(\theta_a)\}} = \frac{1}{\max\{1,\ L(\theta_a)/L(\theta_0)\}}.$$

b. Since $\dfrac{1}{\max\{1,\ L(\theta_a)/L(\theta_0)\}} = \min\{1,\ L(\theta_0)/L(\theta_a)\}$, we have $\lambda < k < 1$ if and only if

$$L(\theta_0)/L(\theta_a) < k.$$

c. The results are consistent with the Neyman–Pearson lemma.

10.112 Denote the samples as $X_1, \ldots, X_{n_1}$, and $Y_1, \ldots, Y_{n_2}$, where $n = n_1 + n_2$.

Under H_a (unrestricted), the MLEs for the parameters are:

$$\hat{\mu}_1 = \overline{X},\ \hat{\mu}_2 = \overline{Y},\ \hat{\sigma}^2 = \tfrac{1}{n}\left(\textstyle\sum_{i=1}^{n_1}(X_i - \overline{X})^2 + \sum_{i=1}^{n_2}(Y_i - \overline{Y})^2\right).$$

Under H_0, $\mu_1 = \mu_2 = \mu$ and the MLEs are

$$\hat{\mu} = \tfrac{n_1\overline{X} + n_2\overline{Y}}{n},\ \hat{\sigma}_0^2 = \tfrac{1}{n}\left(\textstyle\sum_{i=1}^{n_1}(X_i - \hat{\mu})^2 + \sum_{i=1}^{n_2}(Y_i - \hat{\mu})^2\right).$$

By defining the LRT, it is found to be equal to

$$\lambda = \left(\frac{\hat{\sigma}^2}{\hat{\sigma}_0^2}\right)^{n/2} \le k, \text{ or equivalently reject if } \left(\frac{\hat{\sigma}_0^2}{\hat{\sigma}^2}\right) \ge k'.$$

Now, write

$$\textstyle\sum_{i=1}^{n_1}(X_i - \hat{\mu})^2 = \sum_{i=1}^{n_1}(X_i - \overline{X} + \overline{X} - \hat{\mu})^2 = \sum_{i=1}^{n_1}(X_i - \overline{X})^2 + n_1(\overline{X} - \hat{\mu})^2,$$

$$\textstyle\sum_{i=1}^{n_2}(Y_i - \hat{\mu})^2 = \sum_{i=1}^{n_2}(Y_i - \overline{Y} + \overline{Y} - \hat{\mu})^2 = \sum_{i=1}^{n_2}(Y_i - \overline{Y})^2 + n_2(\overline{Y} - \hat{\mu})^2,$$

and since $\hat{\mu} = \frac{n_1}{n}\overline{X} + \frac{n_2}{n}\overline{Y}$, and alternative expression for $\hat{\sigma}_0^2$ is

$$\textstyle\sum_{i=1}^{n_1}(X_i - \overline{X})^2 + \sum_{i=1}^{n_2}(Y_i - \overline{Y})^2 + \frac{n_1 n_2}{n}(\overline{X} - \overline{Y})^2.$$

Thus, the LRT rejects for large values of

$$1 + \tfrac{n_1 n_2}{n}\left(\frac{(\overline{X} - \overline{Y})^2}{\sum_{i=1}^{n_1}(X_i - \overline{X})^2 + \sum_{i=1}^{n_2}(Y_i - \overline{Y})^2}\right).$$

Now, we are only concerned with $\mu_1 > \mu_2$ in H_a, so we could only reject if $\overline{X} - \overline{Y} > 0$.

Thus, the test is equivalent to rejecting if $\dfrac{\overline{X} - \overline{Y}}{\sqrt{\sum_{i=1}^{n_1}(X_i - \overline{X})^2 + \sum_{i=1}^{n_2}(Y_i - \overline{Y})^2}}$ is large.

This is equivalent to the two–sample t test statistic (σ^2 unknown) except for the constants that do not depend on the data.

10.113 Following Ex. 10.112, the LRT rejects for large values of

$$1 + \tfrac{n_1 n_2}{n}\left(\frac{(\overline{X} - \overline{Y})^2}{\sum_{i=1}^{n_1}(X_i - \overline{X})^2 + \sum_{i=1}^{n_2}(Y_i - \overline{Y})^2}\right).$$

Equivalently, the test rejects for large values of

$$\frac{\left|\overline{X} - \overline{Y}\right|}{\sqrt{\sum_{i=1}^{n_1}(X_i - \overline{X})^2 + \sum_{i=1}^{n_2}(Y_i - \overline{Y})^2}}.$$

This is equivalent to the two–sample t test statistic (σ^2 unknown) except for the constants that do not depend on the data.

10.114 Using the sample notation $Y_{11}, \ldots, Y_{1n_1}, Y_{21}, \ldots, Y_{2n_2}, Y_{31}, \ldots, Y_{3n_3}$, with $n = n_1 + n_2 + n_3$, we have that under H_a (unrestricted hypothesis), the MLEs for the parameters are:

$$\hat{\mu}_1 = \overline{Y}_1,\ \hat{\mu}_2 = \overline{Y}_2,\ \hat{\mu}_3 = \overline{Y}_3,\ \hat{\sigma}^2 = \tfrac{1}{n}\left(\sum_{i=1}^{3}\sum_{j=1}^{n_i}(Y_{ij} - \overline{Y}_i)^2\right).$$

Under H_0, $\mu_1 = \mu_2 = \mu_3 = \mu$ so the MLEs are

$$\hat{\mu} = \tfrac{1}{n}\sum_{i=1}^{3}\sum_{j=1}^{n_i} Y_{ij} = \tfrac{n_1\overline{Y}_1 + n_2\overline{Y}_2 + n_3\overline{Y}_3}{n},\quad \hat{\sigma}_0^2 = \tfrac{1}{n}\left(\sum_{i=1}^{3}\sum_{j=1}^{n_i}(Y_{ij} - \hat{\mu})^2\right).$$

Similar to Ex. 10.112, ny defining the LRT, it is found to be equal to

$$\lambda = \left(\frac{\hat{\sigma}^2}{\hat{\sigma}_0^2}\right)^{n/2} \le k\text{, or equivalently reject if } \left(\frac{\hat{\sigma}_0^2}{\hat{\sigma}^2}\right) \ge k'.$$

In order to show that this test is equivalent to and exact F test, we refer to results and notation given in Section 13.3 of the text. In particular,

$$n\hat{\sigma}^2 = \text{SSE}$$
$$n\hat{\sigma}_0^2 = \text{TSS} = \text{SST} + \text{SSE}$$

Then, we have that the LRT rejects when

$$\frac{\text{TSS}}{\text{SSE}} = \frac{\text{SSE}+\text{SST}}{\text{SSE}} = 1 + \frac{\text{SST}}{\text{SSE}} = 1 + \frac{\text{MST}}{\text{MSE}}\tfrac{2}{n-3} + 1 + F\tfrac{2}{n-3} \ge k',$$

where the statistic $F = \dfrac{\text{MST}}{\text{MSE}} = \dfrac{\text{SST/2}}{\text{SSE/}(n\text{-}3)}$ has an F–distribution with 2 numerator and n–3 denominator degrees of freedom under H_0. The LRT rejects when the statistic F is large and so the tests are equivalent,

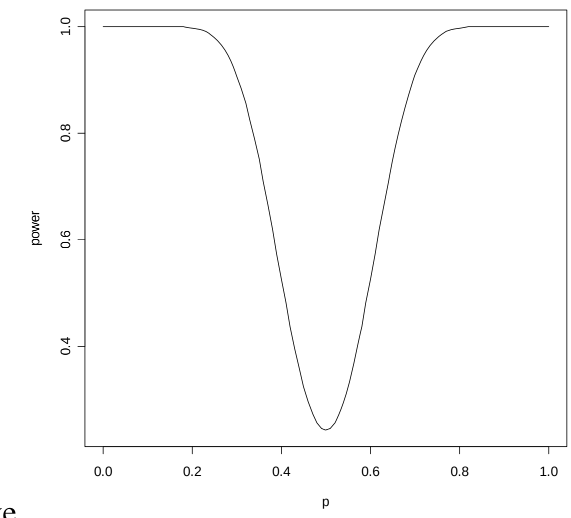
10.115 **a.** True
b. False: H_0 is not a statement regarding a random quantity.
c. False: "large" is a relative quantity
d. True
e. False: power is computed for specific values in H_a
f. False: it must be true that p–value $\le \alpha$
g. False: the UMP test has the highest power against all other α–level tests.
h. False: it always holds that $\lambda \le 1$.
i. True.

10.116 From Ex. 10.6, we have that

$$\text{power}(p) = 1 - \beta(p) = 1 - P(|Y - 18| \le 3 \mid p) = 1 - P(15 \le Y \le 21 \mid p).$$

Thus,

power(.2) = .9975 power(.3) = .9084 power(.4) = .5266
power(.5) = .2430
power(.6) = .9975 power(.7) = .9084 power(.8) = .5266



A graph of the power function is above.

10.117 **a.** The hypotheses are H_0: $\mu_1 - \mu_2 = 0$ vs. H_a: $\mu_1 - \mu_2 \neq 0$, where μ_1 = mean nitrogen density for chemical compounds and μ_2 = mean nitrogen density for air. Then, $s_p^2 = \frac{9(.00131)^2 + 8(.000574)^2}{17} = .000001064$ and $|t| = 22.17$ with 17 degrees of freedom. The p–value is far less than 2(.005) = .01 so H_0 should be rejected.

b. The 95% CI for $\mu_1 - \mu_2$ is (–.01151, –.00951).
c. Since the CI do not contain 0, there is evidence that the mean densities are different.
d. The two approaches agree.

10.118 The hypotheses are H_0: $\mu_1 - \mu_2 = 0$ vs. H_a: $\mu_1 - \mu_2 < 0$, where μ_1 = mean alcohol blood level for sea level and μ_2 = mean alcohol blood level for 12,000 feet. The sample statistics are $\bar{y}_1 = .10$, $s_1 = .0219$, $\bar{y}_2 = .1383$, $s_2 = .0232$. The computed value of the test statistic is $t = -2.945$ and with 10 degrees of freedom, $-t_{.10} = -1.383$ so H_0 should be rejected.

10.119 **a.** The hypotheses are H_0: $p = .20$, H_a: $p > .20$.
b. Let Y = # who prefer brand A. The significance level is
$\alpha = P(Y \geq 92 \mid p = .20) = P(Y > 91.5 \mid p = .20) \approx P(Z > \frac{91.5-80}{8}) = P(Z > 1.44) = .0749$.

10.120 Let μ = mean daily chemical production.
a. H_0: $\mu = 1100$, H_a: $\mu < 1100$.
b. With .05 significance level, we can reject H_0 if $Z < -1.645$.
c. For this large sample test, $Z = -1.90$ and we reject H_0: there is evidence that suggests there has been a drop in mean daily production.

10.121 The hypotheses are H_0: $\mu_1 - \mu_2 = 0$ vs. H_a: $\mu_1 - \mu_2 \neq 0$, where μ_1, μ_2 are the mean breaking distances. For this large–sample test, the computed test statistic is

$|z| = \frac{|118-109|}{\sqrt{\frac{102}{64}+\frac{87}{64}}} = 5.24$. Since p–value ≈ $2P(Z > 5.24)$ is approximately 0, we can reject the null hypothesis: the mean braking distances are different.

10.122 **a.** To test $H_0: \sigma_1^2 = \sigma_2^2$ vs. $H_a: \sigma_1^2 > \sigma_2^2$, where σ_1^2, σ_2^2 represent the population variances for the two lines, the test statistic is $F = (92{,}000)/(37{,}000) = 2.486$ with 49 numerator and 49 denominator degrees of freedom. So, with $F_{.05} = 1.607$ we can reject the null hypothesis.

b. p–value = $P(F > 2.486) = .0009$
Using R:

```
> 1-pf(2.486,49,49)
[1] 0.0009072082
```

10.123 **a.** Our test is $H_0: \sigma_1^2 = \sigma_2^2$ vs. $H_a: \sigma_1^2 \neq \sigma_2^2$, where σ_1^2, σ_2^2 represent the population variances for the two suppliers. The computed test statistic is $F = (.273)/(.094) = 2.904$ with 9 numerator and 9 denominator degrees of freedom. With $\alpha = .05$, $F_{.05} = 3.18$ so H_0 is not rejected: we cannot conclude that the variances are different.

b. The 90% CI is given by $\left(\frac{9(.094)}{16.919}, \frac{9(.094)}{3.32511}\right) = (.050, .254)$. We are 90% confident that the true variance for Supplier B is between .050 and .254.

10.124 The hypotheses are H_0: $\mu_1 - \mu_2 = 0$ vs. H_a: $\mu_1 - \mu_2 \neq 0$, where μ_1, μ_2 are the mean strengths for the two materials. Then, $s_p^2 = .0033$ and $t = \frac{1.237-.978}{\sqrt{.0033\left(\frac{2}{9}\right)}} = 9.568$ with 17 degrees of freedom. With $\alpha = .10$, the critical value is $t_{.05} = 1.746$ and so H_0 is rejected.

10.125 **a.** The hypotheses are H_0: $\mu_A - \mu_B = 0$ vs. H_a: $\mu_A - \mu_B \neq 0$, where μ_A, μ_B are the mean efficiencies for the two types of heaters. The two sample means are 73.125, 77.667, and $s_p^2 = 10.017$. The computed test statistic is $\frac{73.125-77.667}{\sqrt{10.017\left(\frac{1}{8}+\frac{1}{6}\right)}} = -2.657$ with 12 degrees of freedom. Since p–value = $2P(T > 2.657)$, we obtain $.02 < p\text{–value} < .05$ from Table 5 in Appendix III.

b. The 90% CI for $\mu_A - \mu_B$ is

$$73.125 - 77.667 \pm 1.782\sqrt{10.017\left(\tfrac{1}{8}+\tfrac{1}{6}\right)} = -4.542 \pm 3.046 \text{ or } (-7.588, -1.496).$$

Thus, we are 90% confident that the difference in mean efficiencies is between –7.588 and –1.496.

10.126 **a.** $SE(\hat{\theta}) = \sqrt{V(\hat{\theta})} = \sqrt{a_1^2 V(\overline{X}) + a_2^2 V(\overline{Y}) + a_3^2 V(\overline{W})} = \sigma\sqrt{\frac{a_1^2}{n_1} + \frac{a_2^2}{n_1} + \frac{a_3^2}{n_3}}$.

b. Since $\hat{\theta}$ is a linear combination of normal random variables, $\hat{\theta}$ is normally distributed with mean θ and standard deviation given in part **a**.

c. The quantity $(n_1 + n_2 + n_3)S_p^2 / \sigma^2$ is chi–square with $n_1+n_2+n_3 - 3$ degrees of freedom and by Definition 7.2, T has a t–distribution with $n_1+n_2+n_3 - 3$ degrees of freedom.

d. A 100(1 – α)% CI for θ is $\hat{\theta} \pm t_{\alpha/2} s_p \sqrt{\frac{a_1^2}{n_1} + \frac{a_2^2}{n_1} + \frac{a_3^2}{n_3}}$, where $t_{\alpha/2}$ is the upper–α/2 critical value from the t–distribution with $n_1+n_2+n_3 - 3$ degrees of freedom.

e. Under H_0, the quantity $t = \dfrac{(\hat{\theta} - \theta_0)}{s_p \sqrt{\frac{a_1^2}{n_1} + \frac{a_2^2}{n_1} + \frac{a_3^2}{n_3}}}$ has a t–distribution with $n_1+n_2+n_3 - 3$ degrees of freedom. Thus, the rejection region is: $|t| > t_{\alpha/2}$.

10.127 Let $P = X + Y - W$. Then, P has a normal distribution with mean $\mu_1 + \mu_2 - \mu_3$ and variance $(1 + a + b)\sigma^2$. Further, $\overline{P} = \overline{X} + \overline{Y} - \overline{W}$ is normal with mean $\mu_1 + \mu_2 - \mu_3$ and variance $(1 + a + b)\sigma^2/n$. Therefore,

$$Z = \frac{\overline{P} - (\mu_1 + \mu_2 - \mu_3)}{\sigma\sqrt{(1 + a + b)/n}}$$

is standard normal. Next, the quantities

$$\frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{\sigma^2}, \quad \frac{\sum_{i=1}^{n}(Y_i - \overline{Y})^2}{a\sigma^2}, \quad \frac{\sum_{i=1}^{n}(W_i - \overline{W})^2}{b\sigma^2}$$

have independent chi–square distributions, each with $n - 1$ degrees of freedom. So, their sum is chi–square with $3n - 3$ degrees of freedom. Therefore, by Definition 7.2, we can build a random variable that follows a t–distribution (under H_0) by

$$T = \frac{\overline{P} - k}{S_p\sqrt{(1 + a + b)/n}},$$

where $S_P^2 = \left(\sum_{i=1}^{n}(X_i - \overline{X})^2 + \frac{1}{a}\sum_{i=1}^{n}(Y_i - \overline{Y})^2 + \frac{1}{b}\sum_{i=1}^{n}(W_i - \overline{W})^2\right)\Big/(3n - 3)$. For the test, we reject if $|t| > t_{.025}$, where $t_{.025}$ is the upper .024 critical value from the t–distribution with $3n - 3$ degrees of freedom.

10.128 The point of this exercise is to perform a "two–sample" test for means, but information will be garnered from three samples – that is, the common variance will be estimated using three samples. From Section 10.3, we have the standard normal quantity

$$Z = \frac{\overline{X} - \overline{Y} - (\mu_1 - \mu_2)}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}.$$

As in Ex. 10.127, $\left(\sum_{i=1}^{n_1}(X_i - \overline{X})^2 + \sum_{i=1}^{n_2}(Y_i - \overline{Y})^2 + \sum_{i=1}^{n_3}(W_i - \overline{W})^2\right)\Big/\sigma^2$ has a chi–square distribution with $n_1+n_2+n_3 - 3$ degrees of freedom. So, define the statistic

$$S_P^2 = \left(\sum_{i=1}^{n_1}(X_i - \overline{X})^2 + \sum_{i=1}^{n_2}(Y_i - \overline{Y})^2 + \sum_{i=1}^{n_3}(W_i - \overline{W})^2\right)\Big/(n_1 + n_2 + n_3 - 3)$$

and thus the quantity $T = \dfrac{\overline{X} - \overline{Y} - (\mu_1 - \mu_2)}{S_P\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$ has a t–distribution with $n_1+n_2+n_3 - 3$ degrees of freedom.

For the data given in this exercise, we have H_0: $\mu_1 - \mu_2 = 0$ vs. H_a: $\mu_1 - \mu_2 \neq 0$ and with $s_P = 10$, the computed test statistic is $|t| = \frac{|60-50|}{10\sqrt{\frac{2}{10}}} = 2.326$ with 27 degrees of freedom. Since $t_{.025} = 2.052$, the null hypothesis is rejected.

10.129 The likelihood function is $L(\Theta) = \theta_1^{-n}\exp[-\sum_{i=1}^{n}(y_i - \theta_2)/\theta_1]$. The MLE for θ_2 is $\hat{\theta}_2 = Y_{(1)}$. To find the MLE of θ_1, we maximize the log–likelihood function to obtain $\hat{\theta}_1 = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{\theta}_2)$. Under H_0, the MLEs for θ_1 and θ_2 are (respectively) $\theta_{1,0}$ and $\hat{\theta}_2 = Y_{(1)}$ as before. Thus, the LRT is

$$\lambda = \frac{L(\hat{\Omega}_0)}{L(\hat{\Omega})} = \left(\frac{\hat{\theta}_1}{\theta_{1,0}}\right)^n \exp\left[-\frac{\sum_{i=1}^{n}(y_i - y_{(1)})}{\theta_{1,0}} + \frac{\sum_{i=1}^{n}(y_i - y_{(1)})}{\hat{\theta}_1}\right]$$

$$= \left(\frac{\sum_{i=1}^{n}(y_i - y_{(1)})}{n\theta_{1,0}}\right)^n \exp\left[-\frac{\sum_{i=1}^{n}(y_i - y_{(1)})}{\theta_{1,0}} + n\right].$$

Values of $\lambda \leq k$ reject the null hypothesis.

10.130 Following Ex. 10.129, the MLEs are $\hat{\theta}_1 = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{\theta}_2)$ and $\hat{\theta}_2 = Y_{(1)}$. Under H_0, the MLEs for θ_2 and θ_1 are (respectively) $\theta_{2,0}$ and $\hat{\theta}_{1,0} = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \theta_{2,0})$. Thus, the LRT is given by

$$\lambda = \frac{L(\hat{\Omega}_0)}{L(\hat{\Omega})} = \left(\frac{\hat{\theta}_1}{\hat{\theta}_{1,0}}\right)^n \exp\left[-\frac{\sum_{i=1}^{n}(y_i - \theta_{2,0})}{\hat{\theta}_{1,0}} + \frac{\sum_{i=1}^{n}(y_i - y_{(1)})}{\hat{\theta}_1}\right] = \left[\frac{\sum_{i=1}^{n}(y_i - y_{(1)})}{\sum_{i=1}^{n}(y_i - \theta_{2,0})}\right]^n.$$

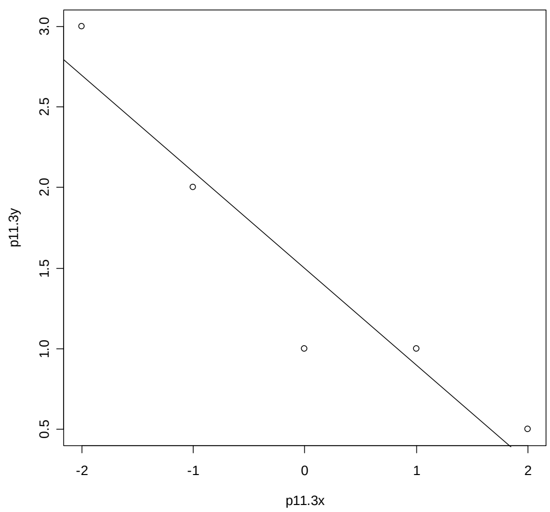
Values of $\lambda \leq k$ reject the null hypothesis.

Chapter 11: Linear Models and Estimation by Least Squares

11.1 Using the hint, $\hat{y}(\bar{x}) = \hat{\beta}_0 + \hat{\beta}_1\bar{x} = (\bar{y} - \hat{\beta}_1\bar{x}) + \hat{\beta}_1\bar{x} = \bar{y}$.

11.2 **a.** slope = 0, intercept = 1. SSE = 6.
b. The line with a negative slope should exhibit a better fit.
c. SSE decreases when the slope changes from .8 to .7. The line is pivoting around the point (0, 1), and this is consistent with ($\bar{x}$, $\bar{y}$) from part Ex. 11.1.
d. The best fit is: $y = 1.000 + 0.700x$.

11.3 The summary statistics are: $\bar{x} = 0$, $\bar{y} = 1.5$, $S_{xy} = -6$, $S_{xx} = 10$. Thus, $\hat{y} = 1.5 - .6x$.



The graph is above.

11.4 The summary statistics are: $\bar{x} = 72$, $\bar{y} = 72.1$, $S_{xy} = 54{,}243$, $S_{xx} = 54{,}714$. Thus, $\hat{y} = 0.72 + 0.99x$. When $x = 100$, the best estimate of y is $\hat{y} = 0.72 + 0.99(100) = 99.72$.

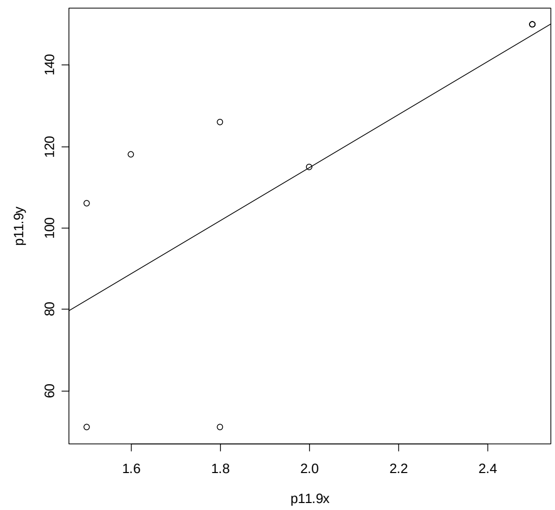
11.5 The summary statistics are: $\bar{x} = 4.5$, $\bar{y} = 43.3625$, $S_{xy} = 203.35$, $S_{xx} = 42$. Thus, $\hat{y} = 21.575 + 4.842x$. Since the slope is positive, this suggests an increase in median prices over time. Also, the expected annual increase is \$4,842.

11.6 **a.** intercept = 43.362, SSE = 1002.839.
b. the data show an increasing trend, so a line with a negative slope would not fit well.
c. Answers vary.
d. Answers vary.
e. (4.5, 43.3625)
f. The sum of the areas is the SSE.

11.7 **a.** The relationship appears to be proportional to x^2.
b. No.
c. No, it is the best *linear* model.

11.8 The summary statistics are: $\bar{x} = 15.505$, $\bar{y} = 9.448$, $S_{xy} = 1546.459$, $S_{xx} = 2359.929$. Thus, $\hat{y} = -0.712 + 0.655x$. When $x = 12$, the best estimate of y is $\hat{y} = -.712 + 0.655(12) = 7.148$.

11.9 **a.** See part **c**.
b. $\hat{y} = -15.45 + 65.17x$.

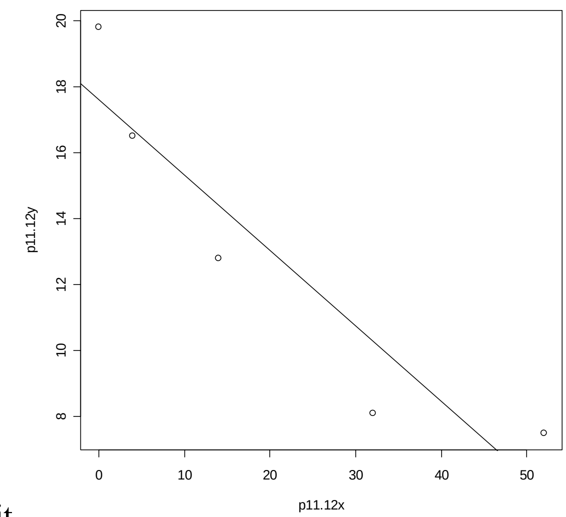


c. The graph is above.
d. When $x = 1.9$, the best estimate of y is $\hat{y} = -15.45 + 65.17(1.9) = 108.373$.

11.10 $\frac{d\text{SSE}}{d\beta_1} = -2\sum_{i=1}^{n}(y_i - \beta_1 x_i)x_i = -2\sum_{i=1}^{n}(x_i y_i - \beta_1 x_i^2) = 0$, so $\hat{\beta}_1 = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2}$.

11.11 Since $\sum_{i=1}^{n} x_i y_i = 134{,}542$ and $\sum_{i=1}^{n} x_i^2 = 53{,}514$, $\hat{\beta}_1 = 2.514$.

11.12 The summary statistics are: $\bar{x} = 20.4$, $\bar{y} = 12.94$, $S_{xy} = -425.571$, $S_{xx} = 1859.2$.
a. The least squares line is: $\hat{y} = 17.609 - 0.229x$.

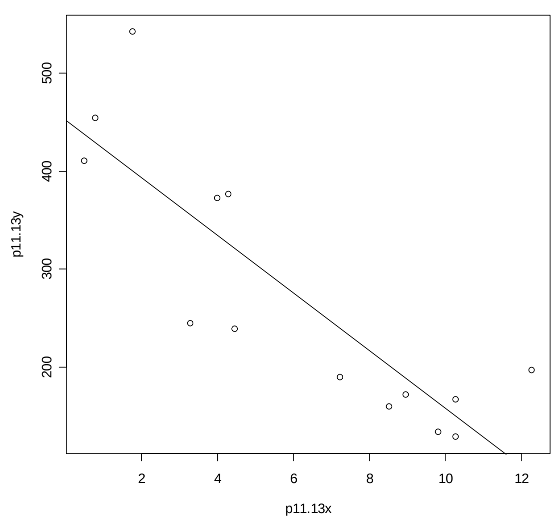


b. The line provides a reasonable fit.

c. When $x = 20$, the best estimate of y is $\hat{y} = 17.609 - 0.229(20) = 13.029$ lbs.

11.13 The summary statistics are: $\bar{x} = 6.177$, $\bar{y} = 270.5$, $S_{xy} = -5830.04$, $S_{xx} = 198.29$.

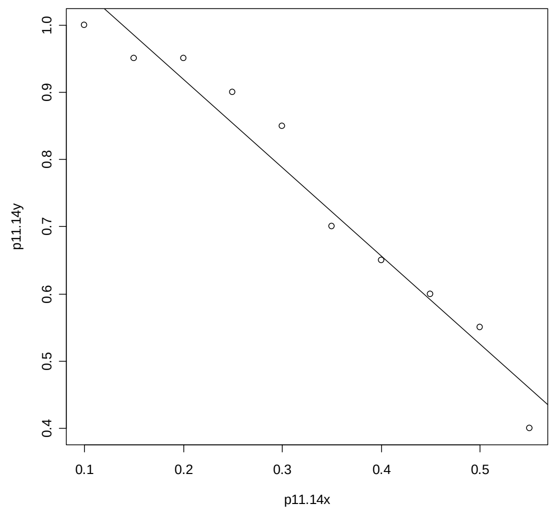
a. The least squares line is: $\hat{y} = 452.119 - 29.402x$.



b. The graph is above.

11.14 The summary statistics are: $\bar{x} = .325$, $\bar{y} = .755$, $S_{xy} = -.27125$, $S_{xx} = .20625$

a. The least squares line is: $\hat{y} = 1.182 - 1.315x$.



b. The graph is above. The line provides a reasonable fit to the data.

11.15 **a.** SSE $= \sum_{i=1}^{n}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2 = \sum_{i=1}^{n}[y_i - \bar{y} - \hat{\beta}_1(x_i - \bar{x})]^2 = \sum_{i=1}^{n}(y_i - \bar{y})^2)$

$$+\ \hat{\beta}_1^2 \sum_{i=1}^{n}(x_i - \bar{x})^2 - 2\hat{\beta}_1 \sum_{i=1}^{n}(y_i - \bar{y})(x_i - \bar{x})$$

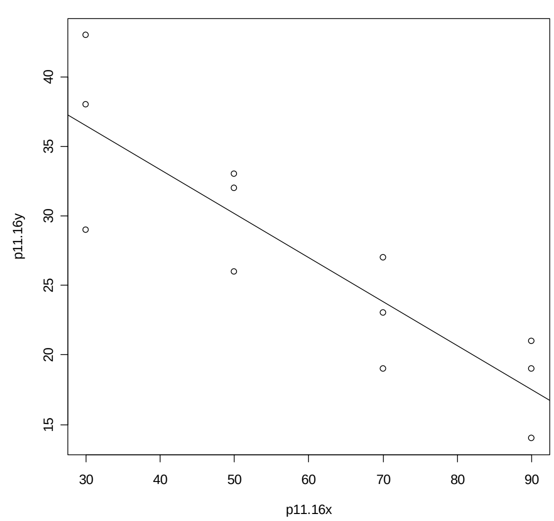
$$= \sum_{i=1}^{n}(y_i - \bar{y})^2 + \hat{\beta}_1 S_{xy} - 2\hat{\beta}_1 S_{xy} = S_{yy} - \hat{\beta}_1 S_{xy}.$$

b. Since $\text{SSE} = S_{yy} - \hat{\beta}_1 S_{xy}$,

$S_{yy} = \text{SSE} + \hat{\beta}_1 S_{xy} = \text{SSE} + (S_{xy})^2 / S_{xx}$. But, $S_{xx} > 0$ and $(S_{xy})^2 \geq 0$. So, $S_{yy} \geq \text{SSE}$.

11.16 The summary statistics are: $\bar{x} = 60$, $\bar{y} = 27$, $S_{xy} = -1900$, $S_{xx} = 6000$.

a. The least squares line is: $\hat{y} = 46.0 - .31667x$.



b. The graph is above.

c. Using the result in Ex. 11.15(a), SSE = $S_{yy} - \hat{\beta}_1 S_{xy}$ = 792 – (–.31667)(–1900) = 190.327. So, s^2 = 190.327/10 = 19.033.

11.17 **a.** With S_{yy} = 1002.8388 and S_{xy} = 203.35, SSE = 1002.8388 – 4.842(203.35) = 18.286. So, s^2 = 18.286/6 = 3.048.

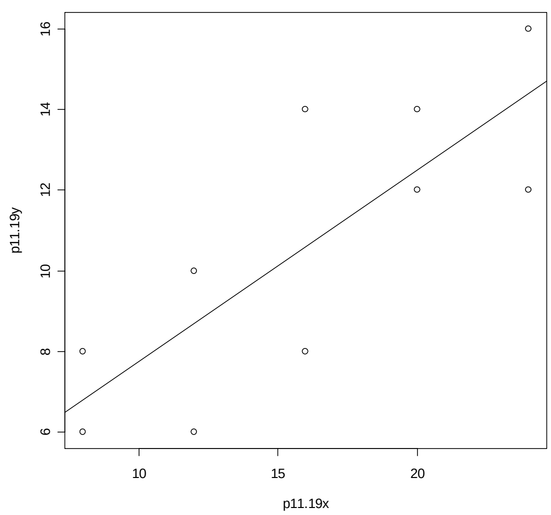
b. The fitted line is $\hat{y}$ = 43.35 + 2.42x^*. The same answer for SSE (and thus s^2) is found.

11.18 **a.** For Ex. 11.8, S_{yy} = 1101.1686 and S_{xy} = 1546.459, SSE = 1101.1686 – .6552528(1546.459) = 87.84701. So, s^2 = 87.84701/8 = 10.98.

b. Using the coding $x_i^* = x_i - \bar{x}$, the fitted line is $\hat{y}$ = 9.448 + .655x^*. The same answer for s^2 is found.

11.19 The summary statistics are: $\bar{x}$ = 16, $\bar{y}$ = 10.6, S_{xy} = 152.0, S_{xx} = 320.

a. The least squares line is: $\hat{y}$ = 3.00 + 4.75x.



b. The graph is above.

c. s^2 = 5.025.

11.20 The likelihood function is given by, $K = \left(\sigma\sqrt{2\pi}\right)^n$,

$$L(\beta_0, \beta_1) = K \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)^2\right], \text{ so that}$$

$$\ln L(\beta_0, \beta_1) = \ln K - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)^2 .$$

Note that maximizing the likelihood (or equivalently the log–likelihood) with respect to β_0 and β_1 is identical to minimizing the positive quantity $\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)^2$. This is the least–squares criterion, so the estimators will be the same.

11.21 Using the results of this section and Theorem 5.12,

$$\text{Cov}(\hat{\beta}_0, \hat{\beta}_1) = \text{Cov}(\overline{Y} - \hat{\beta}_1 \overline{x}, \hat{\beta}_1) = \text{Cov}(\overline{Y}, \hat{\beta}_1) - \text{Cov}(\hat{\beta}_1 \overline{x}, \hat{\beta}_1) = 0 - \overline{x} V(\hat{\beta}_1).$$

Thus, $\text{Cov}(\hat{\beta}_0, \hat{\beta}_1) = -\overline{x}\sigma^2 / S_{xx}$. Note that if $\sum_{i=1}^n x_i = 0$, $\overline{x} = 0$ so $\text{Cov}(\hat{\beta}_0, \hat{\beta}_1) = 0$.

11.22 From Ex. 11.20, let θ = σ^2 so that the log–likelihood is

$$\ln L(\theta) = -\tfrac{n}{2}\ln(2\pi) - \tfrac{n}{2}\ln\theta - \frac{1}{2\theta}\sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2.$$

Thus,

$$\tfrac{d}{d\theta}\ln L(\theta) = -\frac{n}{2\theta} + \frac{1}{2\theta^2}\sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2.$$

The MLE is $\hat{\theta} = \hat{\sigma}^2 = \tfrac{1}{n}\sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2$, but since β_0 and β_1 are unknown, we can insert their MLEs from Ex. 11.20 to obtain:

$$\hat{\sigma}^2 = \tfrac{1}{n}\sum_{i=1}^n \left(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i\right)^2 = \tfrac{1}{n}\text{SSE}.$$

11.23 From Ex. 11.3, it is found that $S_{yy} = 4.0$

a. Since SSE = 4 – (–.6)(–6) = .4, s^2 = .4/3 = .1333. To test H_0: $\beta_1 = 0$ vs. H_a: $\beta_1 \neq 0$, $|t| = \frac{|-.6|}{\sqrt{.1333(.1)}}$ = 5.20 with 3 degrees of freedom. Since $t_{.025}$ = 3.182, we can reject H_0.

b. Since $t_{.005}$ = 5.841 and $t_{.01}$ = 4.541, .01 < p–value < .02.
Using the Applet, $2P(T > 5.20) = 2(.00691) = .01382$.

c. $-.6 \pm 3.182\sqrt{.1333}\sqrt{.1} = -.6 \pm .367$ or (–.967, –.233).

11.24 To test H_0: $\beta_1 = 0$ vs. H_a: $\beta_1 \neq 0$, SSE = 61,667.66 and s^2 =5138.97. Then,
$|t| = \frac{|-29.402|}{\sqrt{5139.97(.005043)}}$ = 5.775 with 12 degrees of freedom.

a. From Table 5, $P(|T| > 3.055) = 2(.005) = .01 > p$–value.
b. Using the Applet, $2P(T > 5.775) = .00008$.
c. Reject H_0.

11.25 From Ex. 11.19, to test H_0: $\beta_1 = 0$ vs. H_a: $\beta_1 \neq 0$, s^2 = 5.025 and S_{xx} = 320. Then,
$|t| = \frac{|.475|}{\sqrt{5.025/320}}$ = 3.791 with 8 degrees of freedom.

a. From Table 5, $P(|T| > 3.355) = 2(.005) = .01 > p$–value.
b. Using the Applet, 2P(T > 3.791) = 2(.00265) = .0053.
c. Reject H_0.
d. We cannot assume the linear trend continues – the number of errors could level off at some point.
e. A 95% CI for β_1: $.475 \pm 2.306\sqrt{5.025/320} = .475 \pm .289$ or (.186, .764). We are 95% confident that the expected change in number of errors for an hour increase of lost sleep is between (.186, .764).

11.26 The summary statistics are: $\overline{x}$ = 53.9, $\overline{y}$ = 7.1, S_{xy} = 198.94, S_{xx} = 1680.69, S_{yy} = 23.6.

a. The least squares line is: $\hat{y} = 0.72 + 0.118x$.

b. SSE = 23.6 – .118(198.94) = .125 so s^2 = .013. A 95% CI for β_1 is 0.118 ± $2.776\sqrt{.013}\sqrt{.00059}$ = 0.118 ± .008.

c. When $x = 0$, $E(Y) = \beta_0 + \beta_1(0) = \beta_0$. So, to test H_0: $\beta_0 = 0$ vs. H_a: $\beta_0 \neq 0$, the test statistic is $|t| = \frac{|.72|}{\sqrt{.013}\sqrt{1.895}}$ = 4.587 with 4 degrees of freedom. Since $t_{.005}$ = 4.604 and $t_{.01}$ = 3.747, we know that .01 < p–value < .02.

d. Using the Applet, $2P(T > 4.587) = 2(.00506) = .01012$.

e. Reject H_0.

11.27 Assuming that the error terms are independent and normally distributed with 0 mean and constant variance σ^2:

a. We know that $Z = \dfrac{\hat{\beta}_i - \beta_{i,0}}{\sigma\sqrt{c_{ii}}}$ has a standard normal distribution under H0.

Furthermore, $V = (n-2)S^2/\sigma^2$ has a chi–square distribution with n – 2 degrees of freedom. Therefore, by Definition 7.2,

$$\frac{Z}{\sqrt{V/(n-2)}} = \frac{\hat{\beta}_i - \beta_{i,0}}{S\sqrt{c_{ii}}}$$

has a t–distribution with n – 2 degrees of freedom under H_0 for i = 1, 2.

b. Using the pivotal quantity expressed above, the result follows from the material in Section 8.8.

11.28 Restricting to Ω_0, the likelihood function is

$$L(\Omega_0) = \frac{1}{(2\pi)^{n/2}\sigma^n}\exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^n (y_i - \beta_0)^2\right].$$

It is not difficult to verify that the MLEs for β_0 and σ^2 under the restricted space are $\overline{Y}$ and $\frac{1}{n}\sum_{i=1}^n (Y_i - \overline{Y})^2$ (respectively). The MLEs have already been found for the unrestricted space so that the LRT simplifies to

$$\lambda = \frac{L(\hat{\Omega}_0)}{L(\hat{\Omega})} = \left(\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \overline{y})^2}\right)^{n/2} = \left(\frac{\text{SSE}}{S_{yy}}\right)^{n/2}.$$

So, we reject if $\lambda \leq k$, or equivalently if

$$\frac{S_{yy}}{\text{SSE}} \geq k^{-2/n} = k'.$$

Using the result from 11.15,

$$\frac{\text{SSE} + \hat{\beta}_1 S_{xy}}{\text{SSE}} = 1 + \frac{\hat{\beta}_1 S_{xy}}{\text{SSE}} = 1 + \frac{\hat{\beta}_1^2 S_{xx}}{(n-2)S} = 1 + \frac{T^2}{(n-2)}.$$

So, we see that λ is small whenever $T = \dfrac{\hat{\beta}_1}{S\sqrt{c_{11}}}$ is <u>large in magnitude</u>, where $c_{11} = \dfrac{1}{S_{xx}}$.

This is the usual t–test statistic, so the result has been proven.

11.29 Let $\hat{\beta}_1$ and $\hat{\gamma}_1$ be the least–squares estimators for the linear models $Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$ and $W_i = \gamma_0 + \gamma_1 c_i + \varepsilon_i$ as defined in the problem Then, we have that:

- $E(\hat{\beta}_1 - \hat{\gamma}_1) = \beta_1 - \gamma_1$
- $V(\hat{\beta}_1 - \hat{\gamma}_1) = \sigma^2\left(\frac{1}{S_{xx}} + \frac{1}{S_{cc}}\right)$, where $S_{cc} = \sum_{i=1}^{m}(c_i - \bar{c})^2$
- $\hat{\beta}_1 - \hat{\gamma}_1$ follows a normal distribution, so that under H_0, $\beta_1 - \gamma_1 = 0$ so that

$$Z = \frac{\hat{\beta}_1 - \hat{\gamma}_1}{\sigma\sqrt{\left(\frac{1}{S_{xx}} + \frac{1}{S_{cc}}\right)}} \text{ is standard normal}$$

- Let $V = \text{SSE}_Y + \text{SSE}_W = \sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2 + \sum_{i=1}^{m}(W_i - \hat{W}_i)^2$. Then, V/σ^2 has a chi–square distribution with $n + m - 4$ degrees of freedom
- By Definition 7.2 we can build a random variable with a t–distribution (under H_0):

$$T = \frac{Z}{\sqrt{V/(n+m-4)}} = \frac{\hat{\beta}_1 - \hat{\gamma}_1}{S\sqrt{\left(\frac{1}{S_{xx}} + \frac{1}{S_{cc}}\right)}}, \text{ where } S = (\text{SSE}_Y + \text{SSE}_W)/(n + m - 4).$$

H_0 is rejected in favor of H_a for large values of $|T|$.

11.30 **a.** For the first experiment, the computed test statistic for H_0: $\beta_1 = 0$ vs. H_a: $\beta_1 \neq 0$ is $t_1 = (.155)/(.0202) = 7.67$ with 29 degrees of freedom. For the second experiment, the computed test statistic is $t_2 = (.190)/(.0193) = 9.84$ with 9 degrees of freedom. Both of these values reject the null hypothesis at $\alpha = .05$, so we can conclude that the slopes are significantly different from 0.

b. Using the result from Ex. 11.29, $S = (2.04 + 1.86)/(31 + 11 - 4) = .1026$. We can extract the values of S_{xx} and S_{cc} from the given values of $V(\hat{\beta}_1)$:

$$S_{xx} = \frac{SSE_Y/(n-2)}{V(\hat{\beta}_1)} = \frac{2.04/29}{(.0202)^2} = 172.397,$$

so similarly $S_{cc} = 554.825$. So, to test equality for the slope parameters, the computed test statistic is

$$|t| = \frac{|.155 - .190|}{\sqrt{.1024\left(\frac{1}{172.397} + \frac{1}{554.825}\right)}} = 1.25$$

with 38 degrees of freedom. Since $t_{.025} \approx z_{.025} = 1.96$, we fail to reject H_0: we cannot conclude that the slopes are different.

11.31 Here, R is used to fit the regression model:

```
> x <- c(19.1, 38.2, 57.3, 76.2, 95, 114, 131, 150, 170)
> y <- c(.095, .174, .256, .348, .429, .500, .580, .651, .722)
> summary(lm(y~x))

Call:
lm(formula = y ~ x)
```

```
Residuals:
       Min         1Q     Median         3Q        Max
-1.333e-02 -4.278e-03 -2.314e-05  8.056e-03  9.811e-03

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 1.875e-02  6.129e-03   3.059   0.0183 *
x           4.215e-03  5.771e-05  73.040 2.37e-11 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.008376 on 7 degrees of freedom
Multiple R-Squared: 0.9987,     Adjusted R-squared: 0.9985
F-statistic:  5335 on 1 and 7 DF,  p-value: 2.372e-11
```

From the output, the fitted model is $\hat{y}$ = .01875 + .004215x. To test H_0: $\beta_1 = 0$ against H_a: $\beta_1 \neq 0$, note that the p–value is quite small indicating a very significant test statistic. Thus, H_0 is rejected and we can conclude that peak current increases as nickel concentrations increase (note that this is a one–sided alternative, so the p–value is actually `2.37e-11` divided by 2).

11.32 **a.** From Ex. 11.5, $\hat{\beta}_1$ = 4.8417 and S_{xx} = 42. From Ex. 11.15, s^2 = 3.0476 so to test H_0: $\beta_1 = 0$ vs. H_a: $\beta_1 > 0$, the required test statistic is t = 17.97 with 6 degrees of freedom. Since $t_{.01}$ = 3.143, H_0 is rejected: there is evidence of an increase.

b. The 99% CI for β_1 is 4.84 ± 1.00 or (3.84, 5.84).

11.33 Using the coded x's from 11.18, $\hat{\beta}_1^*$ = .655 and s^2 = 10.97. Since $S_{xx} = \sum_{i=1}^{10}\left(x_i^*\right)^2$ = 2360.2388, the computed test statistic is $|t| = \frac{|.655|}{\sqrt{\frac{10.97}{2360.2388}}}$ = 9.62 with 8 degrees of freedom. Since $t_{.025}$ = 2.306, we can conclude that there is evidence of a linear relationship.

11.34 **a.** Since $t_{.005}$ = 3.355, we have that p–value < 2(.005) = .01.
b. Using the Applet, $2P(T < 9.61)$ = 2(.00001) = .00002.

11.35 With $a_0 = 1$ and $a_1 = x^*$, the result follows since

$$V(\hat{\beta}_0 + \hat{\beta}_1 x^*) = \frac{1\cdot\frac{1}{n}\sum_{i=1}^n x_i^2 + (x^*)^2 - 2x^*\bar{x}}{S_{xx}}\sigma^2 = \frac{\frac{1}{n}\left(\sum_{i=1}^n x_i^2 - n\bar{x}^2\right) + (x^*)^2 - 2x^*\bar{x} + \bar{x}^2}{S_{xx}}\sigma^2$$

$$= \frac{\frac{1}{n}S_{xx} + (x^* - \bar{x})^2}{S_{xx}}\sigma^2 = \left[\frac{1}{n} + \frac{(x^*-\bar{x})^2}{S_{xx}}\right]\sigma^2 .$$

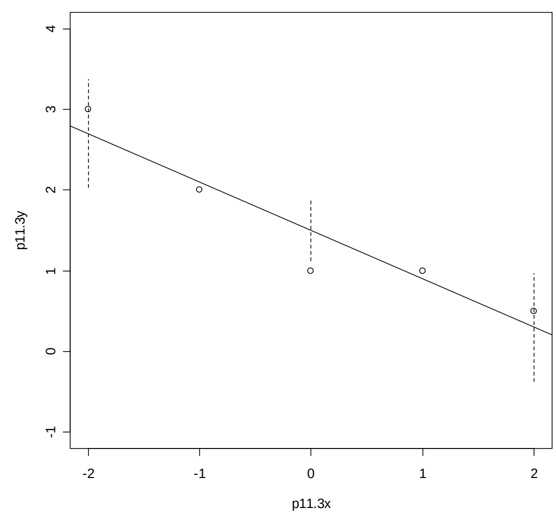
This is minimized when $(x^* - \bar{x})^2 = 0$, so $x^* = \bar{x}$.

11.36 From Ex. 11.13 and 11.24, when $x^* = 5$, $\hat{y} = 452.119 - 29.402(5) = 305.11$ so that $V(\hat{Y})$ is estimated to be 402.98. Thus, a 90% CI for $E(Y)$ is $305.11 \pm 1.782\sqrt{402.98} = 305.11 \pm 35.773$.

11.37 From Ex. 11.8 and 11.18, when $x^* = 12$, $\hat{y} = 7.15$ so that $V(\hat{Y})$ is estimated to be $10.97\left[.1 + \frac{(12-15.504)^2}{2359.929}\right] = 1.154$ Thus, a 95% CI for $E(Y)$ is $7.15 \pm 2.306\sqrt{1.154} = 7.15 \pm 2.477$ or (4.67, 9.63).

11.38 Refer to Ex. 11.3 and 11.23, where $s^2 = .1333$, $\hat{y} = 1.5 - .6x$, $S_{xx} = 10$ and $\bar{x} = 0$.

- When $x^* = 0$, the 90% CI for $E(Y)$ is $1.5 \pm 2.353\sqrt{.1333(\frac{1}{5})}$ or (1.12, 1.88).
- When $x^* = -2$, the 90% CI for $E(Y)$ is $2.7 \pm 2.353\sqrt{.1333(\frac{1}{5} + \frac{4}{10})}$ or (2.03, 3.37).
- When $x^* = 2$, the 90% CI for $E(Y)$ is $.3 \pm 2.353\sqrt{.1333(\frac{1}{5} + \frac{4}{10})}$ or (–.37, .97).



On the graph, note the interval lengths.

11.39 Refer to Ex. 11.16. When $x^* = 65$, $\hat{y} = 25.395$ and a 95% CI for $E(Y)$ is

$$25.395 \pm 2.228\sqrt{19.033\left[\frac{1}{12} + \frac{(65-60)^2}{6000}\right]} \text{ or } 25.395 \pm 2.875.$$

11.40 Refer to Ex. 11.14. When $x^* = .3$, $\hat{y} = .7878$ and with SSE = .0155, $S_{xx} = .20625$, and $\bar{x} = .325$, the 90% CI for $E(Y)$ is $.7878 \pm 1.86\sqrt{\frac{.0155}{8}\left[\frac{1}{10} + \frac{(.3-.325)^2}{.20625}\right]}$ or (.76, .81).

11.41 **a.** Using $\hat{\beta}_0 = \overline{Y} - \hat{\beta}_1\overline{x}$ and $\hat{\beta}_1$ as estimators, we have $\hat{\mu}_y = \overline{Y} - \hat{\beta}_1\overline{x} + \hat{\beta}_1\mu_x$ so that

$$\hat{\mu}_y = \overline{Y} - \hat{\beta}_1(\overline{x} - \mu_x).$$

b. Calculate $V(\hat{\mu}_y) = V(\overline{Y}) + (\overline{x} - \mu_x)^2 V(\hat{\beta}_1) = \frac{\sigma^2}{n} + (\overline{x} - \mu_x)^2 \frac{\sigma^2}{S_{xx}} = \sigma^2\left(\frac{1}{n} + \frac{(\overline{x}-\mu_x)^2}{S_{xx}}\right).$
From Ex. 11.4, $s^2 = 7.1057$ and $S_{xx} = 54{,}714$ so that $\hat{\mu}_y = 72.1 + .99(74 - 72) = 74.08$ and the variance of this estimate is calculated to be $7.1057\left[\frac{1}{10} + \frac{(74-72)^2}{54{,}714}\right] = .711$. The two–standard deviation error bound is $2\sqrt{.711} = 1.69$.

11.42 Similar to Ex. 11.35, the variance is minimized when $x^* = \overline{x}$.

11.43 Refer to Ex. 11.5 and 11.17. When $x = 9$ (year 1980), $\hat{y} = 65.15$ and the 95% PI is

$$65.15 \pm 2.447\sqrt{3.05\left(1 + \tfrac{1}{8} + \tfrac{(9-4.5)^2}{42}\right)} = 65.15 \pm 5.42 \text{ or } (59.73, 70.57).$$

11.44 For the year 1981, $x = 10$. So, $\hat{y} = 69.99$ and the 95% PI is

$$69.99 \pm 2.447\sqrt{3.05\left(1 + \tfrac{1}{8} + \tfrac{(10-4.5)^2}{42}\right)} = 69.99 \pm 5.80.$$

For the year 1982, $x = 11$. So, $\hat{y} = 74.83$ and the 95% PI is

$$74.83 \pm 2.447\sqrt{3.05\left(1 + \tfrac{1}{8} + \tfrac{(11-4.5)^2}{42}\right)} = 74.83 \pm 6.24.$$

Notice how the intervals get wider the further the prediction is from the mean. For the year 1988, this is far beyond the limits of experimentation. So, the linear relationship may not hold (note that the intervals for 1980, 1981 and 1982 are also outside of the limits, so caveat emptor).

11.45 From Ex. 11.8 and 11.18 (also see 11.37), when $x^* = 12$, $\hat{y} = 7.15$ so that the 95% PI is

$$7.15 \pm 2.306\sqrt{10.97\left[1 + \frac{1}{10} + \frac{(12 - 15.504)^2}{2359.929}\right]} = 7.15 \pm 8.03 \text{ or } (-.86, 15.18).$$

11.46 From 11.16 and 11.39, when $x^* = 65$, $\hat{y} = 25.395$ so that the 95% PI is given by

$$25.395 \pm 2.228\sqrt{19.033\left[1 + \frac{1}{12} + \frac{(65 - 60)^2}{6000}\right]} = 25.395 \pm 10.136.$$

11.47 From Ex. 11.14, when $x^* = .6$, $\hat{y} = .3933$ so that the 95% PI is given by

$$.3933 \pm 2.306\sqrt{.00194\left[1 + \frac{1}{10} + \frac{(.6 - .325)^2}{.20625}\right]} = .3933 \pm .12 \text{ or } (.27, .51).$$

11.48 The summary statistics are $S_{xx} = 380.5$, $S_{xy} = 2556.0$, and $S_{yy} = 19{,}263.6$. Thus, $r = .944$. To test H_0: $\rho = 0$ vs. H_a: $\rho > 0$, $t = 8.0923$ with 8 degrees of freedom. From Table 7, we find that p–value < .005.

11.49 **a.** r^2 behaves inversely to SSE, since $r^2 = 1 - \text{SSE}/S_{yy}$.
b. The best model has $r^2 = .817$, so $r = .90388$ (since the slope is positive, r is as well).

11.50 **a.** r^2 increases as the fit improves.
b. For the best model, $r^2 = .982$ and so $r = .99096$.
c. The scatterplot in this example exhibits a smaller error variance about the line.

11.51 The summary statistics are $S_{xx} = 2359.929$, $S_{xy} = 1546.459$, and $S_{yy} = 1101.1686$. Thus, $r = .9593$. To test H_0: $\rho = 0$ vs. H_a: $\rho \neq 0$, $|t| = 9.608$ with 8 degrees of freedom. From Table 7, we see that p–value < 2(.005) = .01 so we can reject the null hypothesis that the correlation is 0.

11.52 **a.** Since the slope of the line is negative, $r = -\sqrt{r^2} = -\sqrt{.61} = -.781$.
b. This is given by r^2, so 61%.
c. To test H_0: $\rho = 0$ vs. H_a: $\rho < 0$, $t = \frac{-.781\sqrt{12}}{\sqrt{1-(-.781)^2}} = -4.33$ with 12 degrees of freedom.

Since $-t_{.05} = -1.782$, we can reject H_0 and conclude that plant density decreases with increasing altitude.

11.53 **a.** This is given by $r^2 = (.8261)^2 = .68244$, or 68.244%.
b. Same answer as part **a**.
c. To test H_0: $\rho = 0$ vs. H_a: $\rho > 0$, $t = \frac{.8261\sqrt{8}}{\sqrt{1-(.8261)^2}} = 4.146$ with 8 degrees of freedom. Since $t_{.01} = 2.896$, we can reject H_0 and conclude that heights and weights are positively correlated for the football players.

d. p–value = $P(T > 4.146) = .00161$.

11.54 **a.** The MOM estimators for σ_X^2 and σ_Y^2 were given in Ex. 9.72.
b. By substituting the MOM estimators, the MOM estimator for ρ is identical to r, the MLE.

11.55 Since $\hat{\beta}_1 = S_{xy}/S_{xx}$ and $r = \hat{\beta}_1\sqrt{S_{xx}/S_{yy}}$, we have that the usual t–test statistic is:

$$T = \frac{\hat{\beta}_1}{\sqrt{S/S_{xx}}} = \frac{\sqrt{S_{xx}}\hat{\beta}_1\sqrt{n-2}}{\sqrt{S_{yy} - \hat{\beta}_1 S_{xy}}} = \frac{\sqrt{S_{xx}/S_{yy}}\hat{\beta}_1\sqrt{n-2}}{\sqrt{1-\hat{\beta}_1 S_{xy}/S_{yy}}} = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}.$$

11.56 Here, $r = .8$.
a. For $n = 5$, $t = 2.309$ with 3 degrees of freedom. Since $t_{.05} = 2.353$, fail to reject H_0.
b. For $n = 12$, $t = 4.2164$ with 10 degrees of freedom. Here, $t_{.05} = 1.812$, reject H_0.
c. For part a, p–value = $P(T > 2.309) = .05209$. For part (b), p–value = .00089.
d. Different conclusions: note the $\sqrt{n-2}$ term in the numerator of the test statistic.
e. The larger sample size in part b caused the computed test statistic to be more extreme. Also, the degrees of freedom were larger.

11.57 **a.** The sample correlation r determines the sign.
b. Both r and n determine the magnitude of $|t|$.

11.58 For the test H_0: $\rho = 0$ vs. H_a: $\rho > 0$, we reject if $t = \frac{r\sqrt{2}}{\sqrt{1-r^2}} \geq t_{.05} = 2.92$. The smallest value of r that would lead to a rejection of H_0 is the solution to the equation

$$r = \tfrac{2.92}{\sqrt{2}}\sqrt{1-r^2}\ .$$

Numerically, this is found to be $r = .9000$.

11.59 For the test H_0: $\rho = 0$ vs. H_a: $\rho < 0$, we reject if $t = \frac{r\sqrt{18}}{\sqrt{1-r^2}} \leq -t_{.05} = -1.734$. The largest value of r that would lead to a rejection of H_0 is the solution to the equation

$$r = \tfrac{-1.734}{\sqrt{18}}\sqrt{1-r^2}\ .$$

Numerically, this is found to be $r = -.3783$.

11.60 Recall the approximate normal distribution of $\frac{1}{2}\ln\left(\frac{1+r}{1-r}\right)$ given on page 606. Therefore, for sample correlations r_1 and r_2, each being calculated from independent samples of size n_1 and n_2 (respectively) and drawn from bivariate normal populations with correlations coefficients ρ_1 and ρ_2 (respectively), we have that

$$Z = \frac{\frac{1}{2}\ln\left(\frac{1+r_1}{1-r_1}\right) - \frac{1}{2}\ln\left(\frac{1+r_2}{1-r_2}\right) - \left[\frac{1}{2}\ln\left(\frac{1+\rho_1}{1-\rho_1}\right) - \frac{1}{2}\ln\left(\frac{1+\rho_2}{1-\rho_2}\right)\right]}{\frac{1}{\sqrt{n_1-3}} + \frac{1}{\sqrt{n_2-3}}}$$

is approximately standard normal for large n_1 and n_2.
Thus, to test H_0: $\rho_1 = \rho_2$ vs. H_a: $\rho_1 \neq \rho_2$ with $r_1 = .9593$, $n_1 = 10$, $r_2 = .85$, $n_2 = 20$, the computed test statistic is

$$z = \frac{\frac{1}{2}\ln\left(\frac{1.9593}{.0407}\right) - \frac{1}{2}\ln\left(\frac{1.85}{.15}\right)}{\frac{1}{\sqrt{7}} + \frac{1}{\sqrt{17}}} = 1.52\ .$$

Since the rejection region is all values $|z| > 1.96$ for $\alpha = .05$, we fail to reject H_0.

11.61 Refer to Example 11.10 and the results given there. The 90% PI is

$$.979 \pm 2.132(.045)\sqrt{1 + \tfrac{1}{6} + \tfrac{(1.5-1.457)^2}{.234}} = .979 \pm .104 \text{ or } (.875, 1.083).$$

11.62 Using the calculations from Example 11.11, we have $r = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}} = .9904$. The proportion of variation described is $r^2 = (.9904)^2 = .9809$.

11.63 **a.** Observe that $\ln E(Y) = \ln\alpha_0 - \alpha_1 x$. Thus, the logarithm of the expected value of Y is linearly related to x. So, we can use the linear model

$$w_i = \beta_0 + \beta_1 x_i + \varepsilon_i,$$

where $w_i = \ln y_i$, $\beta_0 = \ln\alpha_0$ and $\beta_1 = -\alpha_1$. In the above, note that we are assuming an additive error term that is in effect after the transformation. Using the method of least squares, the summary statistics are:

$\bar{x} = 5.5$, $\Sigma x^2 = 385$, $\bar{w} = 3.5505$, $S_{xw} = -.7825$, $S_{xx} = 82.5$, and $S_{ww} = .008448$.
Thus, $\hat{\beta}_1 = -.0095$, $\hat{\beta}_0 = 3.603$ and $\hat{\alpha}_1 = -(-.0095) = .0095$, $\hat{\alpha}_0 = \exp(3.603) = 36.70$.
Therefore, the prediction equation is $\hat{y} = 36.70e^{-.0095x}$.

b. To find a CI for α_0, we first must find a CI for β_0 and then merely transform the endpoints of the interval. First, we calculate the SSE using SSE = $S_{ww} - \hat{\beta}_1 S_{xw}$ = $.008448 - (-.0095)(-.782481) = .0010265$ and so $s^2 = (.0010265)/8 = .0001283$ Using the methods given in Section 11.5, the 90% CI for β_0 is
$3.6027 \pm 1.86\sqrt{.0001283\left(\frac{385}{10(82.5)}\right)}$ or (3.5883, 3.6171). So the 90% CI for α_0 is given by

$$\left(e^{3.5883},\ e^{3.6171}\right) = (36.17, 37.23).$$

11.64 This is similar to Ex. 11.63. Note that $\ln E(Y) = -\alpha_0 x^{\alpha_1}$ and $\ln[-\ln E(Y)] = \ln\alpha_0 + \alpha_1 \ln x$. So, we would expect that $\ln(-\ln y)$ to be linear in $\ln x$. Define $w_i = \ln(-\ln y_i)$, $t_i = \ln x_i$, $\beta_0 = \ln\alpha_0$, $\beta_1 = \alpha_1$. So, we now have the familiar linear model

$$w_i = \beta_0 + \beta_1 t_i + \varepsilon_i$$

(again, we are assuming an additive error term that is in effect after the transformation). The methods of least squares can be used to estimate the parameters. The summary statistics are

$$\bar{t} = -1.12805,\ \bar{w} = -1.4616,\ S_{tw} = 3.6828, \text{ and } S_{tt} = 1.51548$$

So, $\hat{\beta}_1 = 2.4142$, $\hat{\beta}_0 = 1.2617$ and thus $\hat{\alpha}_1 = 2.4142$ and $\hat{\alpha}_0 = \exp(1.2617) = 3.5315$.
This fitted model is $\hat{y} = \exp\left(-3.5315x^{2.4142}\right)$.

11.65 If y is related to t according to $y = 1 - e^{-\beta t}$, then $-\ln(1 - y) = \beta t$. Thus, let $w_i = -\ln(1 - y_i)$ and we have the linear model

$$w_i = \beta t_i + \varepsilon_i$$

(again assuming an additive error term). This is the "no–intercept" model described in Ex. 11.10 and the least squares estimator for β is given to be $\hat{\beta} = \dfrac{\sum_{i=1}^n t_i w_i}{\sum_{i=1}^n t_i^2}$. Now, using similar methods from Section 11.4, note that $V(\hat{\beta}) = \dfrac{\sigma^2}{\sum_{i=1}^n t_i^2}$ and $\dfrac{\text{SSE}}{\sigma^2} = \dfrac{\sum_{i=1}^n (w_i - \hat{w})^2}{\sigma^2}$ is chi–square with $n - 1$ degrees of freedom. So, by Definition 7.2, the quantity

$$T = \frac{\hat{\beta} - \beta}{S / \sum_{i=1}^n t_i^2},$$

where $S = \text{SSE}/(n - 1)$, has a t–distribution with $n - 1$ degrees of freedom.

A $100(1 - \alpha)$% CI for β is

$$\hat{\beta} \pm t_{\alpha/2} S \sqrt{\frac{1}{\sum_{i=1}^{n} t_i^2}} ,$$

and $t_{\alpha/2}$ is the upper– $\alpha/2$ critical value from the t–distribution with $n-1$ degrees of freedom.

11.66 Using the matrix notation from this section,

$$\boldsymbol{X} = \begin{bmatrix} 1 & -2 \\ 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \qquad \boldsymbol{Y} = \begin{bmatrix} 3 \\ 2 \\ 1 \\ 1 \\ .5 \end{bmatrix} \qquad \boldsymbol{X'Y} = \begin{bmatrix} 7.5 \\ -6 \end{bmatrix} \qquad \boldsymbol{X'X} = \begin{bmatrix} 5 & 0 \\ 0 & 10 \end{bmatrix}.$$

Thus, $\hat{\boldsymbol{\beta}} = \begin{bmatrix} 5 & 0 \\ 0 & 10 \end{bmatrix}^{-1} \begin{bmatrix} 7.5 \\ -6 \end{bmatrix} = \begin{bmatrix} .2 & 0 \\ 0 & .1 \end{bmatrix} \begin{bmatrix} 7.5 \\ -6 \end{bmatrix} = \begin{bmatrix} 1.5 \\ -.6 \end{bmatrix}$ so that $\hat{y} = 1.5 - .6x$.

11.67 $$\boldsymbol{X} = \begin{bmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \qquad \boldsymbol{Y} = \begin{bmatrix} 3 \\ 2 \\ 1 \\ 1 \\ .5 \end{bmatrix} \qquad \boldsymbol{X'Y} = \begin{bmatrix} 7.5 \\ 1.5 \end{bmatrix} \qquad \boldsymbol{X'X} = \begin{bmatrix} 5 & 5 \\ 5 & 15 \end{bmatrix}$$

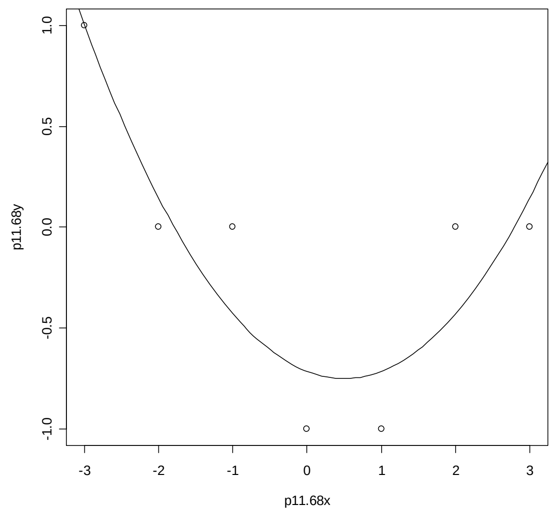
The student should verify that $(\boldsymbol{X'X})^{-1} = \begin{bmatrix} .3 & -.1 \\ -.1 & .1 \end{bmatrix}$ so that $\hat{\boldsymbol{\beta}} = \begin{bmatrix} 2.1 \\ -.6 \end{bmatrix}$. Not that the slope is the same as in Ex. 11.66, but the y–intercept is different. Since $\boldsymbol{X'X}$ is not a diagonal matrix (as in Ex. 11.66), computing the inverse is a bit more tedious.

11.68 $$\boldsymbol{X} = \begin{bmatrix} 1 & -3 & 9 \\ 1 & -2 & 4 \\ 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{bmatrix} \qquad \boldsymbol{Y} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \\ -1 \\ 0 \\ 0 \end{bmatrix} \qquad \boldsymbol{X'Y} = \begin{bmatrix} -1 \\ 4 \\ 8 \end{bmatrix} \qquad \boldsymbol{X'X} = \begin{bmatrix} 7 & 0 & 28 \\ 0 & 28 & 0 \\ 28 & 0 & 196 \end{bmatrix}.$$

The student should verify (either using Appendix I or a computer),

$$(\boldsymbol{X'X})^{-1} = \begin{bmatrix} .3333 & 0 & -.04762 \\ 0 & .035714 & 0 \\ -.04762 & 0 & .011905 \end{bmatrix} \text{ so that } \hat{\boldsymbol{\beta}} = \begin{bmatrix} -.714285 \\ -.142857 \\ .142859 \end{bmatrix} \text{ and the fitted}$$

model is $\hat{y} = -.714285 - .142857x + .142859x^2$.



The graphed curve is above.

11.69 For this problem, R will be used.

```
> x <- c(-7, -5, -3, -1, 1, 3, 5, 7)
> y <- c(18.5,22.6,27.2,31.2,33.0,44.9,49.4,35.0)
```

a. <u>Linear model</u>:

```
> lm(y~x)

Call:
lm(formula = y ~ x)

Coefficients:
(Intercept)            x
     32.725        1.812
```

$\leftarrow \hat{y} = 32.725 + 1.812x$

b. <u>Quadratic model</u>

```
> lm(y~x+I(x^2))

Call:
lm(formula = y ~ x + I(x^2))

Coefficients:
(Intercept)            x       I(x^2)
    35.5625       1.8119      -0.1351
```

$\leftarrow \hat{y} = 35.5625 + 1.8119x - .1351x^2$

11.70 **a.** The student should verify that $\boldsymbol{Y'Y} = 105{,}817$, $\boldsymbol{X'Y} = \begin{bmatrix} 721 \\ 106155 \end{bmatrix}$, and $\hat{\boldsymbol{\beta}} = \begin{bmatrix} .719805 \\ .991392 \end{bmatrix}$.

So, SSE = 105,817 – 105,760.155 = 56.845 and s^2 = 56.845/8 = 7.105625.

b. Using the coding as specified, the data are:

x_i^*	–62	–60	–63	–45	–25	40	–36	169	–13	95
y_i	9	14	7	29	45	109	40	238	60	70

The student should verify that $\boldsymbol{X}^{*\prime}\boldsymbol{Y} = \begin{bmatrix} 721 \\ 54243 \end{bmatrix}$, $\boldsymbol{X}^{*\prime}\boldsymbol{X}^* = \begin{bmatrix} 10 & 0 \\ 0 & 54{,}714 \end{bmatrix}$ and

$\hat{\boldsymbol{\beta}} = \begin{bmatrix} .72.1 \\ .991392 \end{bmatrix}$. So, SSE = 105,817 – 105,760.155 = 56.845 (same answer as part a).

11.71 Note that the vector **a** is composed of k 0's and one 1. Thus,

$$E(\hat{\beta}_i) = E(\mathbf{a}'\hat{\boldsymbol{\beta}}) = \mathbf{a}'E(\hat{\boldsymbol{\beta}}) = \mathbf{a}'\boldsymbol{\beta} = \beta_i$$

$$V(\hat{\beta}_i) = V(\mathbf{a}'\hat{\boldsymbol{\beta}}) = \mathbf{a}'E(\hat{\boldsymbol{\beta}})\mathbf{a} = \mathbf{a}'\sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-1}\mathbf{a} = \sigma^2\mathbf{a}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\mathbf{a} = c_{ii}\sigma^2$$

11.72 Following Ex. 11.69, more detail with the R output is given by:

```
> summary(lm(y~x+I(x^2)))

Call:
lm(formula = y ~ x + I(x^2))

Residuals:
     1      2      3      4      5      6      7      8
 2.242 -0.525 -1.711 -2.415 -4.239  5.118  8.156 -6.625

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  35.5625     3.1224  11.390 9.13e-05 ***
x             1.8119     0.4481   4.044  0.00988 **
I(x^2)       -0.1351     0.1120  -1.206  0.28167
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5.808 on 5 degrees of freedom
Multiple R-Squared: 0.7808,     Adjusted R-squared: 0.6931
F-statistic: 8.904 on 2 and 5 DF,  p-value: 0.0225
```

a. To test H_0: $\beta_2 = 0$ vs. H_a: $\beta_2 \neq 0$, the computed test statistic is $t = –1.206$ and p–value = .28167. Thus, H_0 would not be rejected (no quadratic effect).

b. From the output, it is presented that $\sqrt{V(\hat{\beta}_2)}$ = .1120. So, with 5 degrees of freedom, $t_{.05} = 3.365$ so a 90% for β_2 is –.1351 ± (3.365)(.1120) = –.1351 ± .3769 or (–.512, .2418). Note that this interval contains 0, agreeing with part a.

11.73 If the minimum value is to occur at $x_0 = 1$, then this implies $\beta_1 + 2\beta_2 = 0$. To test this claim, let $\mathbf{a}' = [0\ \ 1\ \ 2]$ for the hypothesis H_0: $\beta_1 + 2\beta_2 = 0$ vs. H_a: $\beta_1 + 2\beta_2 \neq 0$. From

Ex. 11.68, we have that $\hat{\beta}_1 + 2\hat{\beta}_2 = .142861$, $s^2 = .14285$ and we calculate $\mathbf{a}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\mathbf{a} =$ .083334. So, the computed value of the test statistic is $|t| = \frac{|.142861|}{\sqrt{.14285(.083334)}} = 1.31$ with 4 degrees of freedom. Since $t_{.025} = 2.776$, H_0 is not rejected.

11.74 **a.** Each transformation is defined, for each factor, by subtracting the midpoint (the mean) and dividing by one–half the range.

b. Using the matrix definitions of $\boldsymbol{X}$ and $\boldsymbol{Y}$, we have that

$$\boldsymbol{X}'\boldsymbol{Y} = \begin{bmatrix} 338 \\ -50.2 \\ -19.4 \\ -2.6 \\ -20.4 \end{bmatrix} \quad \boldsymbol{X}'\boldsymbol{X} = \begin{bmatrix} 16 & 0 & 0 & 0 & 0 \\ 0 & 16 & 0 & 0 & 0 \\ 0 & 0 & 16 & 0 & 0 \\ 0 & 0 & 0 & 16 & 0 \\ 0 & 0 & 0 & 0 & 16 \end{bmatrix} \text{ so that } \hat{\boldsymbol{\beta}} = \begin{bmatrix} 21.125 \\ -3.1375 \\ -1.2125 \\ -.1625 \\ -1.275 \end{bmatrix}.$$

The fitted model is $\hat{y} = 21.125 - 3.1375x_1 - 1.2125x_2 - .1625x_3 - 1.275x_4$.

c. First, note that SSE = $\boldsymbol{Y}'\boldsymbol{Y} - \hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{Y}$ = 7446.52 – 7347.7075 = 98.8125 so that s^2 = 98.8125/(16 – 5) = 8.98. Further, tests of H_0: $\beta_i = 0$ vs. H_0: $\beta_i \neq 0$ for i = 1, 2, 3, 4, are based on the statistic $t_i = \dfrac{\hat{\beta}_i}{s\sqrt{c_{ii}}} = \dfrac{4\hat{\beta}_i}{\sqrt{8.98}}$ and H_0 is rejected if $|t_i| > t_{.005} = 3.106$. The four computed test statistics are $t_1 = -4.19$, $t_2 = -1.62$, $t_3 = -.22$ and $t_4 = -1.70$. Thus, only the first hypothesis involving the first temperature factor is significant.

11.75 With the four given factor levels, we have $\mathbf{a}' = [1\ \ -1\ \ 1\ \ -1\ \ 1]$ and so $\mathbf{a}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\mathbf{a} =$ 5/16. The estimate of the mean of Y at this setting is

$$\hat{y} = 21.125 + 3.1375 - 1.2125 + .1625 - 1.275 = 21.9375$$

and the 90% confidence interval (based on 11 degrees of freedom) is

$$21.9375 \pm 1.796\sqrt{8.96}\sqrt{5/16} = 21.9375 \pm 3.01 \text{ or } (18.93, 24.95).$$

11.76 First, we calculate $s^2 = \text{SSE}/(n - k - 1) = 1107.01/11 = 100.637$.

a. To test H_0: $\beta_2 = 0$ vs. H_0: $\beta_2 < 0$, we use the t–test with $c_{22} = 8.1\cdot 10^{-4}$:

$$t = \frac{-.92}{\sqrt{100.637(.00081)}} = -3.222.$$

With 11 degrees of freedom, $-t_{.05} = -1.796$ so we reject H_0: there is sufficient evidence that $\beta_2 < 0$.

b. (Similar to Ex. 11.75) With the three given levels, we have $\mathbf{a}' = [1\ \ 914\ \ 65\ \ 6]$ and so $\mathbf{a}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\mathbf{a} = 92.76617$. The estimate of the mean of Y at this setting is

$$\hat{y} = 38.83 - .0092(914) - .92(65) + 11.56(6) = 39.9812$$

and the 95% CI based on 11 degrees of freedom is

$$39.9812 \pm 2.201\sqrt{100.637}\sqrt{92.76617} = 39.9812 \pm 212.664.$$

11.77 Following Ex. 11.76, the 95% PI is $39.9812 \pm 2.201\sqrt{100.637}\sqrt{93.76617} = 39.9812 \pm 213.807$.

11.78 From Ex. 11.69, the fitted model is $\hat{y} = 35.5625 + 1.8119x - .1351x^2$. For the year 2004, $x = 9$ and the predicted sales is $\hat{y} = 35.5625 + 1.8119(9) - .135(9^2) = 40.9346$. With we have $\mathbf{a}' = [1 \quad 9 \quad 81]$ and so $\mathbf{a}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{a} = 1.94643$. The 98% PI for *Lexus* sales in 2004 is then

$$40.9346 \pm 3.365(5.808)\sqrt{1 + 1.94643} = 40.9346 \pm 33.5475.$$

11.79 For the given levels, $\hat{y} = 21.9375$, $\mathbf{a}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{a} = .3135$, and $s^2 = 8.98$. The 90% PI based on 11 degrees of freedom is $21.9375 \pm 1.796\sqrt{8.98(1 + .3135)} = 21.9375 \pm 6.17$ or (15.77, 28.11).

11.80 Following Ex. 11.31, $S_{yy} = .3748$ and SSE $= S_{yy} - \hat{\beta}_1 S_{xy} = .3748 - (.004215)(88.8) = .000508$. Therefore, the F–test is given by $F = \frac{(.3748 - .000508)/1}{.000508/7} = 5157.57$ with 1 numerator and 7 denominator degrees of freedom. Clearly, p–value < .005 so reject H_0.

11.81 From Definition 7.2, let $Z \sim \text{Nor}(0, 1)$ and $W \sim \chi^2_\nu$, and let Z and W be independent. Then, $T = \dfrac{Z}{\sqrt{W/\nu}}$ has the t–distribution with ν degrees of freedom. But, since $Z^2 \sim \chi^2_1$, by Definition 7.3, $F = T^2$ has a F–distribution with 1 numerator and ν denominator degrees of freedom. Now, specific to this problem, note that if $k = 1$, $SSE_R = Syy$. So, the reduced model F–test simplifies to

$$F = \frac{S_{yy} - (S_{yy} - \hat{\beta}_1 S_{xy})}{SSE_C/(n-2)} = \frac{\hat{\beta}_1^2}{s^2/S_{xx}} = T^2.$$

11.82 **a.** To test H_0: $\beta_1 = \beta_2 = \beta_3 = 0$ vs. H_a: at least one $\beta_i \neq 0$, the F–statistic is

$$F = \frac{(10965.46 - 1107.01)/3}{1107.01/11} = 32.653,$$

with 3 numerator and 11 denominator degrees of freedom. From Table 7, we see that p–value < .005, so there is evidence that at least one predictor variable contributes.

b. The coefficient of determination is $R^2 = 1 - \dfrac{1107.01}{10965.46} = .899$, so 89.9% of the variation in percent yield (Y) is explained by the model.

11.83 **a.** To test H_0: $\beta_2 = \beta_3 = 0$ vs. H_a: at least one $\beta_i \neq 0$, the reduced model F–test is

$$F = \frac{(5470.07 - 1107.01)/2}{1107.01/11} = 21.677,$$

with 2 numerator and 11 denominator degrees of freedom. Since $F_{.05}$ = 3.98, we can reject H_0.

b. We must find the value of SSE_R such that $\dfrac{(\text{SSE}_R - 1107.01)/2}{1107.01/11} = 3.98$. The solution is SSE_R = 1908.08

11.84 **a.** The result follows from

$$\frac{n-(k+1)}{k}\left(\frac{R^2}{1-R^2}\right) = \frac{n-(k+1)}{k}\left(\frac{1-\text{SSE}/S_{yy}}{\text{SSE}/S_{yy}}\right) = \frac{(S_{yy} - \text{SSE})/k}{\text{SSE}/[n-(k+1)]} = F.$$

b. The form is $F = T^2$.

11.85 Here, n = 15, k = 4.

a. Using the result from Ex. 11.84, $F = \dfrac{10}{4}\left(\dfrac{.942}{1-.942}\right)$ = 40.603 with 4 numerator and 10 denominator degrees of freedom. From Table 7, it is clear that p–value < .005, so we can safely conclude that at least one of the variables contributes to predicting the selling price.

b. Since $R^2 = 1 - \text{SSE}/S_{yy}$, SSE = 16382.2(1 – .942) = 950.1676.

11.86 To test H_0: $\beta_2 = \beta_3 = \beta_4 = 0$ vs. H_a: at least one $\beta_i \neq 0$, the reduced–model F–test is

$$F = \frac{(1553 - 950.16)/3}{950.1676/10} = 2.115,$$

with 3 numerator and 10 denominator degrees of freedom. Since $F_{.05}$ = 3.71, we fail to reject H_0 and conclude that these variables should be dropped from the model.

11.87 **a.** The F–statistic, using the result in Ex. 11.84, is $F = \dfrac{2}{4}\left(\dfrac{.9}{.1}\right)$ = 4.5 with 4 numerator and 2 denominator degrees of freedom. Since $F_{.1}$ = 9.24, we fail to reject H_0.

b. Since k is large with respect to n, this makes the computed F–statistic small.

c. The F–statistic, using the result in Ex. 11.84, is $F = \dfrac{40}{3}\left(\dfrac{.15}{.85}\right)$ = 2.353 with 3 numerator and 40 denominator degrees of freedom. Since $F_{.1}$ = 2.23, we can reject H_0.

d. Since k is small with respect to n, this makes the computed F–statistic large.

11.88 **a.** False; there are 15 degrees of freedom for SSE.
b. False; the fit (R^2) cannot improve when independent variables are removed.
c. True

d. False; not necessarily, since the degrees of freedom associated with each SSE is different.
e. True.
f. False; Model III is not a reduction of Model I (note the x_1x_2 term).

11.89 **a.** True.
b. False; not necessarily, since Model III is not a reduction of Model I (note the x_1x_2 term).
c. False; for the same reason in part (b).

11.90 Refer to Ex. 11.69 and 11.72.

a. We have that $\text{SSE}_R = 217.7112$ and $\text{SSE}_C = 168.636$. For H_0: $\beta_2 = 0$ vs. H_a: $\beta_2 \neq 0$, the reduced model F–test is $F = \dfrac{217.7112 - 168.636}{168.636/5} = 1.455$ with 1 numerator and 5 denominator degrees of freedom. With $F_{.05} = 6.61$, we fail to reject H_0.

b. Referring to the R output given in Ex. 11.72, the F–statistic is $F = 8.904$ and the p–value for the test is .0225. This leads to a rejection at the $\alpha = .05$ level.

11.91 The hypothesis of interest is H_0: $\beta_1 = \beta_4 = 0$ vs. H_a: at least one $\beta_i \neq 0$, $i = 1, 4$. From Ex. 11.74, we have $\text{SSE}_C = 98.8125$. To find SSE_R, we fit the linear regression model with just x_2 and x_3 so that

$$\boldsymbol{X'Y} = \begin{bmatrix} 338 \\ -19.4 \\ -2.6 \end{bmatrix} \qquad (\boldsymbol{X'X})^{-1} = \begin{bmatrix} 1/16 & 0 & 0 \\ 0 & 1/16 & 0 \\ 0 & 0 & 1/16 \end{bmatrix}$$

and so $\text{SSE}_R = 7446.52 - 7164.195 = 282.325$. The reduced–model F–test is

$$F = \frac{(282.325 - 98.8125)/2}{98.8125/11} = 10.21,$$

with 2 numerator and 11 denominator degrees of freedom. Thus, since $F_{.05} = 3.98$, we can reject H_0 can conclude that either T_1 or T_2 (or both) affect the yield.

11.92 To test H_0: $\beta_3 = \beta_4 = \beta_5 = 0$ vs. H_a: at least one $\beta_i \neq 0$, the reduced–model F–test is

$$F = \frac{(465.134 - 152.177)/3}{152.177/18} = 12.34,$$

with 3 numerator and 18 denominator degrees of freedom. Since $F_{.005} = 5.92$, we have that p–value < .005.

11.93 Refer to Example. 11.19. For the reduced model, $s^2 = 326.623/8 = 40.83$. Then,

$$(\boldsymbol{X'X})^{-1} = \begin{bmatrix} 1/11 & 0 & 0 \\ 0 & 2/17 & 0 \\ 0 & 0 & 2/17 \end{bmatrix}, \ \mathbf{a'} = [1 \quad 1 \quad -1].$$

So, $\hat{y} = \mathbf{a}'\hat{\boldsymbol{\beta}}$ = 93.73 + 4 – 7.35 = 90.38 and $\mathbf{a}'(\mathbf{X'X})^{-1}\mathbf{a}$ = .3262. The 95% CI for $E(Y)$ is $90.38 \pm 2.306\sqrt{40.83(.3262)}$ = 90.38 ± 8.42 or (81.96, 98.80).

11.94 From Example 11.19, tests of H_0: $\beta_i = 0$ vs. H_0: $\beta_i \neq 0$ for i = 3, 4, 5, are based on the statistic $t_i = \dfrac{\hat{\beta}_i}{s\sqrt{c_{ii}}}$ with 5 degrees of freedom and H_0 is rejected if $|t_i| > t_{.01} = 4.032$.

The three computed test statistics are $|t_3|$ = .58, $|t_4|$ = 3.05, $|t_5|$ = 2.53. Therefore, none of the three parameters are significantly different from 0.

11.95 **a.** The summary statistics are: $\bar{x}$ = –268.28, $\bar{y}$ = .6826, S_{xy} = –15.728, S_{xx} = 297.716, and S_{yy} = .9732. Thus, $\hat{y}$ = –13.54 – 0.053x.

b. First, SSE = .9732 – (–.053)(–15.728) = .14225, so s^2 = .14225/8 = .01778. The test statistic is $t = \dfrac{-.053}{\sqrt{\frac{.01778}{297.716}}}$ = –6.86 and H_0 is rejected at the α = .01 level.

c. With x = –273, $\hat{y}$ = –13.54 – .053(–273) = .929. The 95% PI is $.929 \pm 2.306\sqrt{.01778}\sqrt{1+\frac{1}{10}+\frac{(273-268.28)^2}{297.716}}$ = .929 ± .33.

11.96 Here, R will be used to fit the model:

```
> x <- c(.499, .558, .604, .441, .550, .528, .418, .480, .406, .467)
> y <- c(11.14,12.74,13.13,11.51,12.38,12.60,11.13,11.70,11.02,11.41)
> summary(lm(y~x))

Call:
lm(formula = y ~ x)

Residuals:
     Min       1Q   Median       3Q      Max
-0.77823 -0.07102  0.08181  0.16435  0.36771

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   6.5143     0.8528   7.639 6.08e-05 ***
x            10.8294     1.7093   6.336 0.000224 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.3321 on 8 degrees of freedom
Multiple R-Squared: 0.8338,     Adjusted R-squared: 0.813
F-statistic: 40.14 on 1 and 8 DF,  p-value: 0.0002241
```

a. The fitted model is $\hat{y}$ = 6.5143 + 10.8294x.

b. The test H_0: $\beta_1 = 0$ vs. H_a: $\beta_1 \neq 0$ has a p–value of .000224, so H_0 is rejected.

c. It is found that s = .3321 and S_{xx} = .0378. So, with x = .59, $\hat{y}$ = 6.5143 + 10.8294(.59) = 12.902. The 90% CI for $E(Y)$ is $12.902 \pm 1.860(.3321)\sqrt{\frac{1}{10}+\frac{(.59-.4951)^2}{.0378}}$ = 12.902 ± .36.

11.97 **a.** Using the matrix notation,

$$\boldsymbol{X} = \begin{bmatrix} 1 & -3 & 5 & -1 \\ 1 & -2 & 0 & 1 \\ 1 & -1 & -3 & 1 \\ 1 & 0 & -4 & 0 \\ 1 & 1 & -3 & -1 \\ 1 & 2 & 0 & -1 \\ 1 & 3 & 5 & 1 \end{bmatrix}, \ \boldsymbol{Y} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 2 \\ 3 \\ 3 \end{bmatrix}, \ \boldsymbol{X'Y} = \begin{bmatrix} 10 \\ 14 \\ 10 \\ -3 \end{bmatrix}, \ (\boldsymbol{X'X})^{-1} = \begin{bmatrix} 1/7 & 0 & 0 & 0 \\ 0 & 1/28 & 0 & 0 \\ 0 & 0 & 1/84 & 0 \\ 0 & 0 & 0 & 1/6 \end{bmatrix}.$$

So, the fitted model is found to be $\hat{y} = 1.4825 + .5x_1 + .1190x_2 - .5x_3$.

b. The predicted value is $\hat{y} = 1.4825 + .5 - .357 + .5 = 2.0715$. The observed value at these levels was $y = 2$. The predicted value was based on a model fit (using all of the data) and the latter is an observed response.

c. First, note that SSE = 24 – 23.9757 = .0243 so $s^2 = .0243/3 = .008$. The test statistic is $t = \frac{\hat{\beta}_3}{s\sqrt{c_{ii}}} = \frac{-.5}{\sqrt{.008(1/6)}} = -13.7$ which leads to a rejection of the null hypothesis.

d. Here, $\mathbf{a}' = [1 \quad 1 \quad -3 \quad -1]$ and so $\mathbf{a}'(\boldsymbol{X'X})^{-1}\mathbf{a} = .45238$. So, the 95% CI for $E(Y)$ is $2.0715 \pm 3.182\sqrt{.008}\sqrt{.45238} = 2.0715 \pm .19$ or (1.88, 2.26).

e. The prediction interval is $2.0715 \pm 3.182\sqrt{.008}\sqrt{1+.45238} = 2.0715 \pm .34$ or (1.73, 2.41).

11.98 Symmetric spacing about the origin creates a diagonal **X′X** matrix which is very easy to invert.

11.99 Since $V(\hat{\beta}_1) = \dfrac{\sigma^2}{S_{xx}}$, this will be minimized when $S_{xx} = \sum_{i=1}^{n}(x_i - \bar{x})^2$ is as large as possible. This occurs when the x_i are as far away from $\bar{x}$ as possible. If $-9 \le x \le 9$, chose $n/2$ at $x = -9$ and $n/2$ at $x = 9$.

11.100 Based on the minimization strategy in Ex. 11.99, the values of x are: –9, –9, –9, –9, –9, 9, 9, 9, 9, 9. Thus $S_{xx} = \sum_{i=1}^{10}(x_i - \bar{x})^2 = \sum_{i=1}^{10} x_i^2 = 810$. If equal spacing is employed, the values of x are: –9, –7, –5, –3, –1, 1, 3, 5, 7, 9. Thus, $S_{xx} = \sum_{i=1}^{10}(x_i - \bar{x})^2 = \sum_{i=1}^{10} x_i^2$ = 330. The relative efficiency is the ratio of the variances, or 330/810 = 11/27.

11.101 Here, R will be used to fit the model:

```
> x1 <- c(0,0,0,0,0,1,1,1,1,1)
> x2 <- c(-2,-1,0,1,2,-2,-1,0,1,2)
> y <- c(8,9,9.1,10.2,10.4,10,10.3,12.2,12.6,13.9)
```

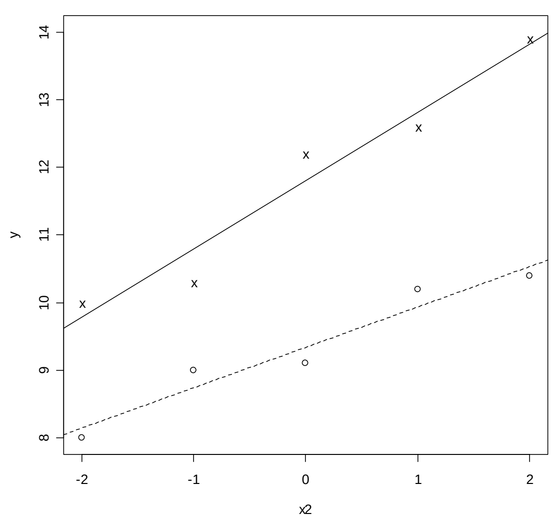
```
> summary(lm(y~x1+x2+I(x1*x2)))
Call:
lm(formula = y ~ x1 + x2 + I(x1 * x2))

Residuals:
    Min      1Q  Median      3Q     Max
-0.4900 -0.1925 -0.0300  0.2500  0.4000

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   9.3400     0.1561  59.834 1.46e-09 ***
x1            2.4600     0.2208  11.144 3.11e-05 ***
x2            0.6000     0.1104   5.436  0.00161 **
I(x1 * x2)    0.4100     0.1561   2.627  0.03924 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 0.349 on 6 degrees of freedom
Multiple R-Squared: 0.9754,     Adjusted R-squared: 0.963
F-statistic: 79.15 on 3 and 6 DF,  p-value: 3.244e-05
```

a. The fitted model is $\hat{y} = 9.34 + 2.46x_1 + .6x_2 + .41x_1x_2$.

b. For bacteria type A, $x_1 = 0$ so $\hat{y} = 9.34 + .6x_2$ (dotted line)
For bacteria type B, $x_1 = 1$ so $\hat{y} = 11.80 + 1.01\,x_2$ (solid line)



c. For bacteria A, $x_1 = 0$, $x_2 = 0$, so $\hat{y} = 9.34$. For bacteria B, $x_1 = 1$, $x_2 = 0$, so $\hat{y} = 11.80$. The observed growths were 9.1 and 12.2, respectively.

d. The rates are different if the parameter β_3 is nonzero. So, H_0: $\beta_3 = 0$ vs. H_a: $\beta_3 \neq 0$ has a p–value = .03924 (R output above) and H_0 is rejected.

e. With $x_1 = 1$, $x_2 = 1$, so $\hat{y} = 12.81$. With $s = .349$ and $\mathbf{a}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\mathbf{a} = .3$, the 90% CI is $12.81 \pm .37$.

f. The 90% PI is $12.81 \pm .78$.

11.102 The reduced model F statistic is $F = \dfrac{(795.23 - 783.9)/2}{783.9/195} = 1.41$ with 2 numerator and 195 denominator degrees of freedom. Since $F_{.05} \approx 3.00$, we fail to reject H_0: salary is not dependent on gender.

11.103 Define $\boldsymbol{1}$ as a column vector of n 1's. Then $\bar{y} = \frac{1}{n}\boldsymbol{1}'\boldsymbol{Y}$. We must solve for the vector $\boldsymbol{x}$ such that $\bar{y} = \boldsymbol{x}'\hat{\boldsymbol{\beta}}$. Using the matrix definition of $\hat{\boldsymbol{\beta}}$, we have

$$\bar{y} = \boldsymbol{x}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y} = \tfrac{1}{n}\boldsymbol{1}'\boldsymbol{Y}$$
$$\boldsymbol{x}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y}\boldsymbol{Y}' = \tfrac{1}{n}\boldsymbol{1}'\boldsymbol{Y}\boldsymbol{Y}'$$

which implies

$$\boldsymbol{x}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}' = \tfrac{1}{n}\boldsymbol{1}'$$
$$\boldsymbol{x}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{X} = \tfrac{1}{n}\boldsymbol{1}'\boldsymbol{X}$$

so that

$$\boldsymbol{x}' = \tfrac{1}{n}\boldsymbol{1}'\boldsymbol{X}.$$

That is, $\boldsymbol{x}' = [1 \;\; \bar{x}_1 \;\; \bar{x}_2 \;\; \ldots \;\; \bar{x}_k]$.

11.104 Here, we will use the coding $x_1 = \frac{P-65}{15}$ and $x_2 = \frac{T-200}{100}$. Then, the levels are $x_1 = -1, 1$ and $x_2 = -1, 0, 1$.

a. $\boldsymbol{Y} = \begin{bmatrix} 21 \\ 23 \\ 26 \\ 22 \\ 23 \\ 28 \end{bmatrix}$ $\boldsymbol{X} = \begin{bmatrix} 1 & -1 & -1 & 1 \\ 1 & -1 & 0 & 0 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & -1 & 1 \end{bmatrix}$ $\boldsymbol{X}'\boldsymbol{Y} = \begin{bmatrix} 143 \\ 3 \\ 11 \\ 97 \end{bmatrix}$ $(\boldsymbol{X}'\boldsymbol{X})^{-1} = \begin{bmatrix} .5 & 0 & 0 & -.5 \\ 0 & .1667 & 0 & 0 \\ 0 & 0 & .25 & 0 \\ -.5 & 0 & 0 & .75 \end{bmatrix}$

So, the fitted model is $\hat{y} = 23 + .5x_1 + 2.75x_2 + 1.25x_2^2$.

b. The hypothesis of interest is H_0: $\beta_3 = 0$ vs. H_a: $\beta_3 \neq 0$ and the test statistic is (verify that SSE = 1 so that $s^2 = .5$) $|t| = \frac{|1.25|}{\sqrt{.5(.75)}} = 2.040$ with 2 degrees of freedom. Since $t_{.025} = 4.303$, we fail to reject H_0.

c. To test H_0: $\beta_2 = \beta_3 = 0$ vs. H_a: at least one $\beta_i \neq 0$, $i = 2, 3$, the reduced model must be fitted. It can be verified that $SSE_R = 33.33$ so that the reduced model F–test is $F = 32.33$ with 2 numerator and 2 denominator degrees of freedom. It is easily seen that H_0 should be rejected; temperature does affect yield.

11.105 **a.** $\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}}\sqrt{\frac{S_{yy}}{S_{xx}}} = r\sqrt{\frac{S_{yy}}{S_{xx}}}$.

b. The conditional distribution of Y_i, given $X_i = x_i$, is (see Chapter 5) normal with mean $\mu_y + \rho\frac{\sigma_y}{\sigma_x}(x_i - \mu_x)$ and variance $\sigma_y^2(1-\rho^2)$. Redefine $\beta_1 = \rho\frac{\sigma_y}{\sigma_x}$, $\beta_0 = \mu_y - \beta_1\mu_x$. So, if $\rho = 0$, $\beta_1 = 0$. So, using the usual t–statistic to test $\beta_1 = 0$, we have

$$T = \frac{\hat{\beta}_1}{S/\sqrt{S_{xx}}} = \frac{\hat{\beta}_1}{\sqrt{\frac{\text{SSE}}{n-2}}\sqrt{\frac{1}{S_{xx}}}} = \frac{\hat{\beta}_1\sqrt{(n-2)S_{xx}}}{\sqrt{(1-r^2)S_{yy}}}.$$

c. By part a, $\hat{\beta}_1 = r\sqrt{\frac{S_{yy}}{S_{xx}}}$ and the statistic has the form as shown. Note that the distribution only depends on $n - 2$ and not the particular value x_i. So, the distribution is the same unconditionally.

11.106 The summary statistics are $S_{xx} = 66.54$, $S_{xy} = 71.12$, and $S_{yy} = 93.979$. Thus, $r = .8994$. To test H_0: $\rho = 0$ vs. H_a: $\rho \neq 0$, $|t| = 5.04$ with 6 degrees of freedom. From Table 7, we see that p–value < 2(.005) = .01 so we can reject the null hypothesis that the correlation is 0.

11.107 The summary statistics are $S_{xx} = 153.875$, $S_{xy} = 12.8$, and $S_{yy} = 1.34$.

a. Thus, $r = .89$.

b. To test H_0: $\rho = 0$ vs. H_a: $\rho \neq 0$, $|t| = 4.78$ with 6 degrees of freedom. From Table 7, we see that p–value < 2(.005) = .01 so we can reject the null hypothesis that the correlation is 0.

11.108 **a.-c.** Answers vary.

Chapter 12: Considerations in Designing Experiments

12.1 (See Example 12.1) Let $n_1 = \left(\frac{\sigma_1}{\sigma_1+\sigma_2}\right)n = \left(\frac{3}{3+5}\right)90$ = 33.75 or 34 and n_2 = 90 – 34 = 56.

12.2 (See Ex. 12.1). If n_1 = 34 and n_2 = 56, then

$$\sigma_{\bar{Y}_1-\bar{Y}_2} = \sqrt{\tfrac{9}{34}+\tfrac{25}{56}} = \sqrt{.7111}$$

In order to achieve this same bound with equal sample sizes, we must have

$$\sqrt{\tfrac{9}{n}+\tfrac{25}{n}} = \sqrt{.7111}$$

The solution is n = 47.8 or 48. Thus, it is necessary to have $n_1 = n_2$ = 48 so that the same amount of information is implied.

12.3 The length of a 95% CI is twice the margin of error:

$$2(1.96)\sqrt{\tfrac{9}{n_1}+\tfrac{25}{n_2}},$$

and this is required to be equal to two. In Ex. 12.1, we found $n_1 = (3/8)n$ and $n_1 = (5/8)n$, so substituting these values into the above and equating it to two, the solution is found to be n = 245.9. Thus, n_1 = 93 and n_2 = 154.

12.4 (Similar to Ex. 12.3) Here, the equation to solve is

$$2(1.96)\sqrt{\tfrac{9}{n_1}+\tfrac{25}{n_1}} = 2.$$

The solution is n_1 = 130.6 or 131, and the total sample size required is 131 + 131 = 262.

12.5 Refer to Section 12.2. The variance of the slope estimate is minimized (maximum information) when S_{xx} is as large as possible. This occurs when the data are as far away from $\bar{x}$ as possible. So, with n = 6, three rats should receive x = 2 units and three rats should receive x = 5 units.

12.6 When σ is known, a 95% CI for β is given by

$$\hat{\beta}_1 \pm z_{\alpha/2}\frac{\sigma}{\sqrt{S_{xx}}}.$$

Under the two methods, we calculate that S_{xx} = 13.5 for Method 1 and S_{xx} = 6.3 for Method 2. Thus, Method 2 will produce the longer interval. By computing the ratio of the margins of error for the two methods (Method 2 to Method 1), we obtain $\sqrt{\frac{13.5}{6.3}}$ = 1.464; thus Method 2 produces an interval that is 1.464 times as large as Method 1.

Under Method 2, suppose we take n measurements at each of the six dose levels. It is not difficult to show that now $S_{xx} = 6.3n$. So, in order for the intervals to be equivalent, we must have that $6.3n = 13.5$, and so n = 2.14. So, roughly twice as many observations are required.

12.7 Although it was assumed that the response variable Y is truly linear over the range of x, the experimenter has no way to verify this using Method 2. By assigning a few points at x = 3.5, the experimenter could check for curvature in the response function.

12.8 Checking for true linearity and constant error variance cannot be performed if the data points are spread out as far as possible.

12.9 **a.** Each half of the iron ore sample should be reasonably similar, and assuming the two methods are similar, the data pairs should be positively correlated.

b. Either analysis compares means. However, the paired analysis requires fewer ore samples and reduces the sample–to–sample variability.

12.10 The sample statistics are: $\bar{d} = -.0217$, $s_D^2 = .0008967$.

a. To test H_0: $\mu_D = 0$ vs. H_a: $\mu_D \neq 0$, the test statistic is $|t| = \frac{|-.0217|}{\sqrt{.0008967/6}} = 1.773$ with 5 degrees of freedom. Since $t_{.025} = 2.571$, H_0 is not rejected.

b. From Table 5, $.10 < p\text{–value} < .20$.

c. The 95% CI is $-.0217 \pm 2.571\sqrt{\frac{.0008967}{6}} = -.0217 \pm .0314$.

12.11 Recall that $Var(\bar{D}) = \frac{1}{n}\left(\sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2\right)$ given in this section.

a. This occurs when $\rho > 0$.

b. This occurs when $\rho = 0$.

c. This occurs when $\rho < 0$.

d. If the samples are negatively correlated, a matched–pairs experiment should not be performed. Otherwise, if it is possible, the matched–pairs experiment will have an associated variance that is equal or less than the variance associated with the independent samples experiment.

12.12 **a.** There are $2n - 2$ degrees of freedom for error.

b. There are $n - 1$ degrees of freedom for error.

c.

n	Independent samples	Matched–pairs
5	d.f. = 8, $t_{.025} = 2.306$	d.f. = 4, $t_{.025} = 2.776$
10	d.f. = 18, $t_{.025} = 2.101$	d.f. = 9, $t_{.025} = 2.262$
30	d.f. = 58, $t_{.025} = 1.96$	d.f. = 29, $t_{.025} = 2.045$

d. Since more observations are required for the independent samples design, this increases the degrees of freedom for error and thus shrinks the critical values used in confidence intervals and hypothesis tests.

12.13 A matched–pairs experiment is preferred since there could exist sample–to–sample variability when using independent samples (one person could be more prone to plaque buildup than another).

12.14 The sample statistics are: $\bar{d} = -.333$, $s_D^2 = 5.466$. To test H_0: $\mu_D = 0$ vs. H_a: $\mu_D < 0$, the test statistic is $t = \frac{-.333}{\sqrt{5.466/6}} = -.35$ with 5 degrees of freedom. From Table 5, p–value > .1 so H_0 is not rejected.

12.15 **a.** The sample statistics are: $\overline{d}$ = 1.5, s_D^2 = 2.571. To test H_0: $\mu_D = 0$ vs. H_a: $\mu_D \neq 0$, the test statistic is $|t| = \frac{|1.5|}{\sqrt{2.571/8}}$ = 2.65 with 7 degrees of freedom. Since $t_{.025}$ = 2.365, H_0 is rejected.

b. Notice that each technician's score is similar under both design A and B, but the technician's scores are not similar in general (some are high and some are low). Thus, pairing is important to screen out the variability among technicians.

c. We assumed that the population of differences follows a normal distribution, and that the sample used in the analysis was randomly selected.

12.16 The sample statistics are: $\overline{d}$ = –3.88, s_D^2 = 8.427.

a. To test H_0: $\mu_D = 0$ vs. H_a: $\mu_D < 0$, the test statistic is $t = \frac{-3.88}{\sqrt{8.427/15}}$ = –5.176 with 14 degrees of freedom. From Table 5, it is seen that p–value < .005, so H_0 is rejected when $\alpha = .01$.

b. A 95% CI is $-3.88 \pm 2.145\sqrt{8.427/15}$ = –3.88 ± 1.608.

c. Using the Initial Reading data, $\overline{y}$ = 36.926 and s^2 = 40.889. A 95% CI for the mean muck depth is $36.926 \pm 2.145\sqrt{40.889/15}$ = 36.926 ± 3.541.

d. Using the Later Reading data, $\overline{y}$ = 33.046 and s^2 = 35.517. A 95% CI for the mean much depth is $33.046 \pm 2.145\sqrt{35.517/15}$ = 33.046 ± 3.301.

e. For parts a and b, we assumed that the population of differences follows a normal distribution, and that the sample used in the analysis was randomly selected. For parts c and d, we assumed that the individual samples were randomly selected from two normal populations.

12.17 **a.** $E(Y_{ij}) = \mu_i + E(U_i) + E(\varepsilon_{ij}) = \mu_i$.

b. Each Y_{1j} involves the sum of a uniform and a normal random variable, and this convolution does not result in a normal random variable.

c. $\text{Cov}(Y_{1j}, Y_{2j}) = \text{Cov}(\mu_1 + U_j + \varepsilon_{1j}, \mu_2 + U_j + \varepsilon_{2j})$ = $\text{Cov}(\mu_1, \mu_2)$ + $\text{Cov}(U_j, U_j)$ + $\text{Cov}(\varepsilon_{1j}, \varepsilon_{2j})$ = 0 + $V(U_j)$ + 0 = 1/3.

d. Observe that $D_j = Y_{1j} - Y_{2j} = \mu_1 - \mu_2 + \varepsilon_{1j} - \varepsilon_{2j}$. Since the random errors are independent and follow a normal distribution, D_j is a normal random variable. Further, for $j \neq j'$, $\text{Cov}(D_j, D_{j'}) = 0$ since the two random variables are comprised of constants and independent normal variables. Thus, D_j and $D_{j'}$ are independent (recall that if two normal random variables are uncorrelated, they are also independent – but this is not true in general).

e. Provided that the distribution of U_j has a mean of zero and finite variance, the result will hold.

12.18 Use Table 12 and see Section 12.4 of the text.

12.19 Use Table 12 and see Section 12.4 of the text.

12.20 **a.** There are six treatments. One example would be the first catalyst and the first temperature setting.
b. After assigning the n experimental units to the treatments, the experimental units are numbered from 1 to n. Then, a random number table is used to select numbers until all experimental units have been selected.

12.21 Randomization avoids the possibility of bias introduced by a nonrandom selection of sample elements. Also, it provides a probabilistic basis for the selection of a sample.

12.22 Factors are independent experimental variables that the experimenter can control.

12.23 A treatment is a specific combination of factor levels used in an experiment.

12.24 Yes. Suppose that a plant biologist is comparing three soil types used for planting, where the response is the yield of a crop planted in the different soil types. Then, "soil type" is a factor variable. But, if the biologist is comparing the yields of different greenhouses, but each greenhouse used different soil types, then "soil type" is a nuisance variable.

12.25 <u>Increases accuracy of the experiment</u>: 1) selection of treatments, 2) choice of number of experimental units assigned to each treatment.
<u>Decreases the impact of extraneous sources of variability</u>: randomization; assigning treatments to experimental units.

12.26 There is a possibility of significant rat–to–rat variation. By applying all four dosages to tissue samples extracted from the same rat, the experimental error is reduced. This design is an example of a randomized block design.

12.27 In the Latin square design, each treatment appears in each row and each column exactly once. So, the design is:

B	*A*	*C*
C	*B*	*A*
A	*C*	*B*

12.28 A CI could be constructed for the specific population parameter, and the width of the CI gives the quantity of information.

12.29 A random sample of size n is a sample that was randomly selected from all possible (unique) samples of size n (constructed of observations from the population of interest) and each sample had an equal chance of being selected.

12.30 From Section 12.5, the choice of factor levels and the allocation of the experimental units to the treatments, as well as the total number of experimental units being used, affect the total quantity of information. Randomization and blocking can control these factors.

12.31 Given the model proposed in this exercise, we have the following:

a. $E(Y_{ij}) = \mu_i + E(P_i) + E(\varepsilon_{ij}) = \mu_i + 0 + 0 = \mu_i$.

b. Obviously, $E(\overline{Y}_i) = \mu_i$. Also, $V(\overline{Y}_i) = \frac{1}{n}V(Y_{ij}) = \frac{1}{n}\left[V(P_i) + V(\varepsilon_{ij})\right] = \frac{1}{n}\left[\sigma_P^2 + \sigma^2\right]$, since P_i and ε_{ij} are independent for all i, j.

c. From part b, $E(\overline{D}) = E(\overline{Y}_1) - E(\overline{Y}_2) = \mu_1 - \mu_2$. Now, to find $V(\overline{D})$, note that

$$\overline{D} = \frac{1}{n}\sum_{j=1}^{n} D_j = \mu_1 - \mu_2 + \frac{1}{n}\left[\sum_{j=1}^{n} \varepsilon_{1j} + \sum_{j=1}^{n} \varepsilon_{2j}\right].$$

Thus, since the ε_{ij} are independent, $V(\overline{D}) = \frac{1}{n^2}\left[\sum_{j=1}^{n} V(\varepsilon_{1j}) + \sum_{j=1}^{n} V(\varepsilon_{2j})\right] = 2\sigma^2/n$.

Further, since $\overline{D}$ is a linear combination of normal random variables, it is also normally distributed.

12.32 From Exercise 12.31, clearly $\dfrac{\overline{D} - (\mu_1 - \mu_2)}{\sqrt{2\sigma^2/n}}$ has a standard normal distribution. In addition, since $D_1, \ldots, D_n$ are independent normal random variables with mean $\mu_1 - \mu_2$ and variance $2\sigma^2$, the quantity

$$W = \frac{(n-1)S_D^2}{2\sigma^2} = \frac{\sum_{i=1}^{n}(D_i - \overline{D})^2}{2\sigma^2}$$

is chi–square with $\nu = n - 1$ degrees of freedom. Therefore, by Definition 7.2 and under H_0: $\mu_1 - \mu_2 = 0$,

$$\frac{Z}{\sqrt{W/\nu}} = \frac{\overline{D}}{S_D/\sqrt{n}}$$

has a t–distribution with $n - 1$ degrees of freedom.

12.33 Using similar methods as in Ex. 12.31, we find that for this model,

$$V(\overline{D}) = \frac{1}{n^2}\sum_{j=1}^{n}\left[V(P_{1j}) + V(P_{2j}) + V(\varepsilon_{1j}) + V(\varepsilon_{2j})\right] = \frac{1}{n}\left[2\sigma_P^2 + 2\sigma^2\right] > \frac{1}{n}2\sigma^2.$$

Thus, the variance is larger with the completely randomized design, since the unwanted variation due to pairing is not eliminated.

12.34 The sample statistics are: $\overline{d} = -.062727$, $s_D^2 = .012862$.

a. We expect the observations to be positively correlated since (assuming the people are honest) jobs that are estimated to take a long time actually take a long time when processed. Similar for jobs that are estimated to take a small amount of processor time.

b. To test H_0: $\mu_D = 0$ vs. H_a: $\mu_D < 0$, the test statistic is $t = \frac{-.0627272}{\sqrt{.012862/11}} = -1.834$ with 10 degrees of freedom. Since $-t_{.10} = -1.362$, H_0 is rejected: there is evidence that the customers tend to underestimate the processor time.

c. From Table 5, we have that $.025 < p$–value $< .05$.

d. A 90% CI for $\mu_D = \mu_1 - \mu_2$, is $-.062727 \pm 1.812\sqrt{.012862/11} = -.063 \pm .062$ or $(-.125, -.001)$.

12.35 The sample statistics are: $\overline{d} = -1.58$, $s_D^2 = .667$.

a. To test H_0: $\mu_D = 0$ vs. H_a: $\mu_D \neq 0$, the test statistic is $|t| = \frac{|-1.58|}{\sqrt{.667/5}} = 4.326$ with 4 degrees of freedom. From Table 5, we can see that $.01 < p$–value $< .025$, so H_0 would be rejected for any $\alpha \geq .025$.

b. A 95% CI is given by $-1.58 \pm 2.776\sqrt{.667/5} = -1.58 \pm 1.014$ or $(-2.594, -.566)$.

c. We will use the estimate of the variance of paired differences. Also, since the required sample will (probably) be large, we will use the critical value from the standard normal distribution. Our requirement is then:

$$.2 = z_{.025}\sqrt{\frac{\sigma_D^2}{n}} \approx 1.96\sqrt{\frac{.667}{n}}.$$

The solution is $n = 64.059$, or 65 observations (pairs) are necessary.

12.36 The sample statistics are: $\overline{d} = 106.9$, $s_D^2 = 1364.989$.

a. Each subject is presented each sign in random order. If the subject's reaction time is (in general) high, both responses should be high. If the subject's reaction time is (in general) low. both responses should be low. Because of the subject–to–subject variability, the matched pairs design can eliminate this extraneous source of variation.

b. To test H_0: $\mu_D = 0$ vs. H_a: $\mu_D \neq 0$, the test statistic is $|t| = \frac{|106.0|}{\sqrt{1364.989/10}} = 9.15$ with 9 degrees of freedom. Since t.025 = 2.262, H_0 is rejected.

c. From Table 5, we see that p–value $< 2(.005) = .01$.

d. The 95% CI is given by $106.9 \pm 2.262\sqrt{1364.989/10} = 106.9 \pm 26.428$ or (80.472, 133.328).

12.37 There are nk_1 points at $x = -1$, nk_2 at $x = 0$, and nk_3 points at $x = 1$. The design matrix $\boldsymbol{X}$ can be expressed as

$$\boldsymbol{X} = \begin{bmatrix} 1 & -1 & 1 \\ 1 & -1 & 1 \\ \vdots & \vdots & \vdots \\ 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ \vdots & \vdots & \vdots \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ \vdots & \vdots & \vdots \\ 1 & 1 & 1 \end{bmatrix}; \text{ thus } \boldsymbol{X'X} = \begin{bmatrix} n & n(k_3 - k_1) & n(k_1 + k_3) \\ n(k_3 - k_1) & n(k_1 + k_3) & n(k_3 - k_1) \\ n(k_1 + k_3) & n(k_3 - k_1) & n(k_1 + k_3) \end{bmatrix} = n\begin{bmatrix} 1 & b & a \\ b & a & b \\ a & b & a \end{bmatrix} = n\boldsymbol{A},$$

where $a = k_1 + k_3$ and $b = k_3 - k_1$.

Now, the goal is to minimize $V(\hat{\beta}_2) = \sigma^2 c_{22}$, where c_{22} is the (3, 3) element of $(\boldsymbol{X'X})^{-1}$. To calculate $(\boldsymbol{X'X})^{-1}$, note that it can be expressed as

$$\boldsymbol{A}^{-1} = \frac{1}{n\det(\boldsymbol{A})}\begin{bmatrix} a^2-b^2 & 0 & b^2-a^2 \\ 0 & a-a^2 & ab-b \\ b^2-a^2 & ab-b & a-b^2 \end{bmatrix}$$

, and (the student should verify) the determinant of $\boldsymbol{A}$ simplifies to $\det(\boldsymbol{A}) = 4k_1k_2k_3$. Hence,

$$V(\hat{\beta}_2) = \sigma^2 \frac{a-b^2}{4nk_1k_2k_3} = \frac{\sigma^2}{n}\left(\frac{k_1+k_3-(k_3-k_1)^2}{4k_1k_2k_3}\right).$$

We must minimize

$$Q = \frac{k_1+k_3-(k_3-k_1)^2}{4k_1k_2k_3} = \frac{k_1+k_3-\left[(k_3+k_1)^2-4k_1k_3\right]}{4k_1k_2k_3} = \frac{(k_1+k_3)\left[1-k_1-k_3\right]}{4k_1k_2k_3} - \frac{4k_1k_3}{4k_1k_2k_3}$$

$$= \frac{k_1+k_3}{4k_1k_3} - \frac{1}{k_2} = \frac{k_1+k_3}{4k_1k_3} - \frac{1}{1-k_1-k_3}.$$

So, with $Q = \frac{k_1+k_3}{4k_1k_3} - \frac{1}{1-k_1-k_3}$, we can differentiate this with respect to k_1 and k_3 and set these equal to zero. The two equations are:

$$\begin{aligned} 4k_1^2 &= (1-k_1-k_3)^2 \qquad (*) \\ 4k_3^2 &= (1-k_1-k_3)^2 \end{aligned}.$$

Since k_1, k_2, and k_3 are all positive, $k_1 = k_3$ by symmetry of the above equations and therefore by (*), $4k_1^2 = (1-2k_1)^2$ so that $k_1 = k_3 = .25$. Thus, $k_2 = .50$.

Chapter 13: The Analysis of Variance

13.1 The summary statistics are: $\bar{y}_1 = 1.875$, $s_1^2 = .6964286$, $\bar{y}_2 = 2.625$, $s_2^2 = .8392857$, and $n_1 = n_2 = 8$. The desired test is: H_0: $\mu_1 = \mu_2$ vs. H_a: $\mu_1 \neq \mu_2$, where μ_1, μ_2 represent the mean reaction times for Stimulus 1 and 2 respectively.

a. SST = $4(1.875 - 2.625)^2 = 2.25$, SSE = $7(.696428) + 7(.8392857) = 10.75$. Thus, MST = 2.25/1 = 2.25 and MSE = 10.75/14 = .7679. The test statistic $F = 2.25/.7679 = 2.93$ with 1 numerator and 14 denominator degrees of freedom. Since $F_{.05} = 4.60$, we fail to reject H_0: the stimuli are not significantly different.

b. Using the Applet, p–value = $P(F > 2.93) = .109$.

c. Note that s_p^2 = MSE = .7679. So, the two–sample t–test statistic is $|t| = \frac{|1.875-2.625|}{\sqrt{.7679\left(\frac{2}{8}\right)}}$ = 1.712 with 14 degrees of freedom. Since $t_{.025} = 2.145$, we fail to reject H_0. The two tests are equivalent, and since $F = T^2$, note that $2.93 \approx (1.712)^2$ (roundoff error).

d. We assumed that the two random samples were selected independently from normal populations with equal variances.

13.2 Refer to Ex. 10.77. The summary statistics are: $\bar{y}_1 = 446$, $s_1^2 = 42$, $\bar{y}_2 = 534$, $s_2^2 = 45$, and $n_1 = n_2 = 15$.

a. SST = $7.5(446 - 534)^2 = 58{,}080$, SSE = $14(42) + 14(45) = 1218$. So, MST = 58,080 and MSE = 1218/28 = 1894.5. The test statistic $F = 58{,}080/1894.5 = 30.64$ with 1 numerator and 28 denominator degrees of freedom. Clearly, p–value < .005.

b. Using the Applet, p–value = $P(F > 30.64) = .00001$.

c. In Ex. 10.77, $t = -5.54$. Observe that $(-5.54)^2 \approx 30.64$ (roundoff error).

d. We assumed that the two random samples were selected independently from normal populations with equal variances.

13.3 See Section 13.3 of the text.

13.4 For the four groups of students, the sample variances are: $s_1^2 = 66.6667$, $s_2^2 = 50.6192$, $s_3^2 = 91.7667$, $s_4^2 = 33.5833$ with $n_1 = 6$, $n_2 = 7$, $n_3 = 6$, $n_4 = 4$. Then, SSE = 5(66.6667) + 6(50.6192) + 5(91.7667) + 3(33.5833) = 1196.6321, which is identical to the prior result.

13.5 Since W has a chi–square distribution with r degrees of freedom, the mgf is given by

$$m_W(t) = E(e^{tW}) = (1-2t)^{-r/2}.$$

Now, $W = U + V$, where U and V are independent random variables and V is chi–square with s degrees of freedom. So,

$$m_W(t) = E(e^{tW}) = E(e^{t(U+V)}) = E(e^{tU})E(e^{tV}) = E(e^{tU})(1-2t)^{-s/2} = (1-2t)^{-r/2}.$$

Therefore, $m_U(t) = E(e^{tU}) = \frac{(1-2t)^{-r/2}}{(1-2t)^{-s/2}} = (1-2t)^{-(r-s)/2}$. Since this is the mgf for a chi–square random variable with $r - s$ degrees of freedom, where $r > s$, by the Uniqueness Property for mgfs U has this distribution.

13.6 **a.** Recall that by Theorem 7.3, $(n_i - 1)S_i^2 / \sigma^2$ is chi–square with $n_i - 1$ degrees of freedom. Since the samples are independent, by Ex. 6.59, $\text{SSE}/\sigma^2 = \sum_{i=1}^{k} (n_i - 1)S_i^2 / \sigma^2$ is chi–square with $n - k$ degrees of freedom.

b. If H_0 is true, all of the observations are identically distributed since it was already assumed that the samples were drawn independently from normal populations with common variance. Thus, under H_0, we can combine all of the samples to form an estimator for the common mean, $\overline{Y}$, and an estimator for the common variance, given by TSS/(n – 1). By Theorem 7.3, TSS/σ^2 is chi–square with n – 1 degrees of freedom.

c. The result follows from Ex. 13.5: let W = TSS/σ^2 where $r = n - 1$ and let V = SSE/σ^2 where $s = n - k$. Now, SSE/σ^2 is distributed as chi–square with $n - k$ degrees of freedom and TSS/σ^2 is distributed as chi–square <u>under H_0</u>. Thus, U = SST/σ^2 is chi–square <u>under H_0</u> with $n - 1 - (n - k) = k - 1$ degrees of freedom.

d. Since SSE and TSS are independent, by Definition 7.3

$$F = \frac{\text{SST}/\left(\sigma^2(k-1)\right)}{\text{SSE}/\left(\sigma^2(n-k)\right)} = \frac{\text{MST}}{\text{MSE}}$$

has an F–distribution with $k - 1$ numerator and $n - k$ denominator degrees of freedom.

13.7 We will use R to solve this problem:

```
> waste <- c(1.65, 1.72, 1.5, 1.37, 1.6, 1.7, 1.85, 1.46, 2.05, 1.8,
1.4, 1.75, 1.38, 1.65, 1.55, 2.1, 1.95, 1.65, 1.88, 2)
> plant <- c(rep("A",5), rep("B",5), rep("C",5), rep("D",5))
> plant <- factor(plant)      # change plant to a factor variable
> summary(aov(waste~plant))
            Df  Sum Sq Mean Sq F value  Pr(>F)
plant        3 0.46489 0.15496  5.2002 0.01068 *
Residuals   16 0.47680 0.02980
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

a. The F statistic is given by F = MST/MSE = .15496/.0298 = 5.2002 (given in the ANOVA table above) with 3 numerator and 16 denominator degrees of freedom. Since $F_{.05} = 3.24$, we can reject H_0: $\mu_1 = \mu_2 = \mu_3 = \mu_4$ and conclude that at least one of the plant means are different.
b. The p–value is given in the ANOVA table: p–value = .01068.

13.8 Similar to Ex. 13.7, R will be used to solve the problem:

```
> salary <- c(49.3, 49.9, 48.5, 68.5, 54.0, 81.8, 71.2, 62.9, 69.0,
69.0, 66.9, 57.3, 57.7, 46.2, 52.2)
> type <- factor(c(rep("public",5), rep("private",5), rep("church",5)))
```

a. This is a completely randomized, one–way layout (this is sampled data, not a designed experiment).
b. To test H_0: $\mu_1 = \mu_2 = \mu_3$, the ANOVA table is given below (using R):

```
> summary(aov(salary~type))
            Df Sum Sq Mean Sq F value   Pr(>F)
type         2 834.98  417.49  7.1234 0.009133 **
Residuals   12 703.29   58.61
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

From the output, F = MST/MSE = 7.1234 with 3 numerator and 12 denominator degrees of freedom. From Table 7, .005 < p–value < .01.

c. From the output, p-value = .009133.

13.9 The test to be conducted is H_0: $\mu_1 = \mu_2 = \mu_3 = \mu_4$, where μ_i is the mean strength for the i^{th} mix of concrete, i = 1, 2, 3, 4. The alternative hypothesis at least one of the equalities does not hold.

a. The summary statistics are: TSS = .035, SST = .015, and so SSE = .035 – .015 = .020. The mean squares are MST = .015/3 = .005 and MSE = .020/8 = .0025, so the F statistic is given by F = .005/.0025 = 2.00, with 3 numerator and 8 denominator degrees of freedom. Since $F_{.05}$ = 4.07, we fail to reject H_0: there is not enough evidence to reject the claim that the concrete mixes have equal mean strengths.

b. Using the Applet, p–value = $P(F > 2)$ = .19266. The ANOVA table is below.

Source	d.f	SS	MS	F	p–value
Treatments	3	.015	.005	2.00	.19266
Error	8	.020	.0025		
Total	11	.035			

13.10 The test to be conducted is H_0: $\mu_1 = \mu_2 = \mu_3$, where μ_i is the mean score where the i^{th} method was applied, i = 1, 2, 3. The alternative hypothesis at least one of the equalities does not hold

a. The summary statistics are: TSS = 1140.5455, SST = 641.8788, and so SSE = 1140.5455 – 641.8788 = 498.6667. The mean squares are MST = 641.8788/2 = 320.939 and MSE = 498.6667/8 = 62.333, so the F statistic is given by F = 320.939/62.333 = 5.148, with 2 numerator and 8 denominator degrees of freedom. By Table 7, .025 < p–value < .05.

b. Using the Applet, p–value = $P(F > 5.148)$ = .03655. The ANOVA table is below.

Source	d.f	SS	MS	F	p–value
Treatments	2	641.8788	320.939	5.148	.03655
Error	8	498.6667	62.333		
Total	10	1140.5455			

c. With α = .05, we would reject H_0: at least one of the methods has a different mean score.

13.11 Since the three sample sizes are equal, $\bar{y} = \frac{1}{3}(\bar{y}_1 + \bar{y}_2 + \bar{y}_3) = \frac{1}{3}(.93 + 1.21 + .92) = 1.02$. Thus, SST = $n_1 \sum_{i=1}^{3} (\bar{y}_i - \bar{y})^2 = 14 \sum_{i=1}^{3} (\bar{y}_i - 1.02)^2 = .7588$. Now, recall that the "standard error of the mean" is given by $s/\sqrt{n}$, so SSE can be found by

$$\text{SSE} = 13[14(.04)^2 + 14(.03)^2 + 14(.04)^2] = .7462.$$

Thus, the mean squares are MST = .7588/2 = .3794 and MSE = .7462/39 = .019133, so that the F statistic is F = .3794/.019133 = 19.83 with 2 numerator and 39 denominator degrees of freedom. From Table 7, it is seen that p–value < .005, so at the .05 significance level we reject the null hypothesis that the mean bone densities are equal.

13.12 The test to be conducted is H_0: $\mu_1 = \mu_2 = \mu_3$, where μ_i is the mean percentage of Carbon 14 where the i^{th} concentration of acetonitrile was applied, i = 1, 2, 3. The alternative hypothesis at least one of the equalities does not hold

a. The summary statistics are: TSS = 235.219, SST = 174.106, and so SSE = 235.219 – 174.106 = 61.113. The mean squares are MST = 174.106/2 = 87.053 and MSE = 235.219/33 = 1.852, so the F statistic is given by F = 87.053/1.852 = 47.007, with 2 numerator and 33 denominator degrees of freedom. Since $F_{.01} \approx 5.39$, we reject H0: at least one of the mean percentages is different and p–value < .005. The ANOVA table is below.

Source	d.f	SS	MS	F	p–value
Treatments	2	174.106	87.053	47.007	< .005
Error	33	61.113	1.852		
Total	35	235.219			

b. We must assume that the independent measurements from low, medium, and high concentrations of acetonitrile are normally distributed with common variance.

13.13 The grand mean is $\bar{y} = \frac{45(4.59)+102(4.88)+18(6.24)}{165} = 4.949$. So,

$$\text{SST} = 45(4.59 - 4.949)^2 + 102(4.88 - 4.949)^2 + 18(6.24 - 4.949)^2 = 36.286.$$

$$\text{SSE} = \sum_{i=1}^{3} (n-1)s_i^2 = 44(.70)^2 + 101(.64)^2 + 17(.90)^2 = 76.6996.$$

The F statistic is $F = \frac{\text{MST}}{\text{MSE}} = \frac{36.286/2}{76.6996/162} = 38.316$ with 2 numerator and 162 denominator degrees of freedom. From Table 7, p–value < .005 so we can reject the null hypothesis of equal mean maneuver times. The ANOVA table is below.

Source	d.f	SS	MS	F	p–value
Treatments	2	36.286	18.143	38.316	< .005
Error	162	76.6996	.4735		
Total	164	112.9856			

13.14 The grand mean is $\bar{y} = \frac{.032+.022+.041}{3} = 0.0317$. So,

$$\text{SST} = 10[(.032 - .0317)^2 + (.022 - .0317)^2 + (.041 - .0317)^2 = .001867.$$

$$\text{SSE} = \sum_{i=1}^{3} (n-1)s_i^2 = 9[(.014)^2 + (.008)^2 + (.017)^2] = .004941.$$

The F statistic is $F = 4.94$ with 2 numerator and 27 denominator degrees of freedom. Since $F_{.05} = 3.35$, we can reject H_0 and conclude that the mean chemical levels are different.

13.15 We will use R to solve this problem:

```
> oxygen <- c(5.9, 6.1, 6.3, 6.1, 6.0, 6.3, 6.6, 6.4, 6.4, 6.5, 4.8,
4.3, 5.0, 4.7, 5.1, 6.0, 6.2, 6.1, 5.8)
> location <- factor(c(1,1,1,1,1,2,2,2,2,2,3,3,3,3,3,4,4,4,4))
> summary(aov(oxygen~location))
            Df Sum Sq Mean Sq F value    Pr(>F)
location     3 7.8361  2.6120  63.656 9.195e-09 ***
Residuals   15 0.6155  0.0410
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
>
```

The null hypothesis is H_0: $\mu_1 = \mu_2 = \mu_3 = \mu_4$, where μ_i is the mean dissolved O_2 in location i, $i = 1, 2, 3, 4$. Since the p–value is quite small, we can reject H_0 and conclude the mean dissolved O_2 levels differ.

13.16 The ANOVA table is below:

Source	d.f	SS	MS	F	p–value
Treatments	3	67.475	22.4917	.87	> .1
Error	36	935.5	25.9861		
Total	39	1002.975			

With 3 numerator and 36 denominator degrees of freedom, we fail to reject with $\alpha = .05$: there is not enough evidence to conclude a difference in the four age groups.

13.17 $E(\overline{Y}_{i\bullet}) = \frac{1}{n_i}\sum_{j=1}^{n_i} E(Y_{ij}) = \frac{1}{n_i}\sum_{j=1}^{n_i}(\mu + \tau_i) = \frac{1}{n_i}\sum_{j=1}^{n_i}\mu_i = \mu_i$

$V(\overline{Y}_{i\bullet}) = \frac{1}{n_i^2}\sum_{j=1}^{n_i} V(Y_{ij}) = \frac{1}{n_i^2}\sum_{j=1}^{n_i} V(\varepsilon_{ij}) = \frac{1}{n_i}\sigma^2$.

13.18 Using the results from Ex. 13.17,

$$E(\overline{Y}_{i\bullet} - \overline{Y}_{i'\bullet}) = \mu_i - \mu_{i'} = \mu + \tau_i - (\mu + \tau_{i'}) = \tau_i - \tau_{i'}$$
$$V(\overline{Y}_{i\bullet} - \overline{Y}_{i'\bullet}) = V(\overline{Y}_{i\bullet}) + V(\overline{Y}_{i'\bullet}) = \left(\tfrac{1}{n_i} + \tfrac{1}{n_{i'}}\right)\sigma^2$$

13.19 **a.** Recall that $\mu_i = \mu + \tau_i$ for $i = 1, \ldots, k$. If all τ_i's = 0, then all μ_i's = μ. Conversely, if $\mu_1 = \mu_2 = \ldots = \mu_k$, we have that $\mu + \tau_1 = \mu + \tau_2 = \ldots = \mu + \tau_k$ and $\tau_1 = \tau_2 = \ldots = \tau_k$. Since it was assumed that $\sum_{i=1}^{k}\tau_i = 0$, all τ_i's = 0. Thus, the null hypotheses are equivalent.

b. Consider $\mu_i = \mu + \tau_i$ and $\mu_{i'} = \mu + \tau_{i'}$. If $\mu_i \neq \mu_{i'}$, then $\mu + \tau_i \neq \mu + \tau_{i'}$ and thus $\tau_i \neq \tau_{i'}$. Since $\sum_{i=1}^{k}\tau_i = 0$, at least one $\tau_i \neq 0$ (actually, there must be at least two). Conversely, let

$\tau_i \neq 0$. Since $\sum_{i=1}^{k} \tau_i = 0$, there must be at least one i' such that $\tau_i \neq \tau_{i'}$. With $\mu_i = \mu + \tau_i$ and $\mu_{i'} = \mu + \tau_{i'}$, it must be so that $\mu_i \neq \mu_{i'}$. Thus, the alternative hypotheses are equivalent.

13.20 **a.** First, note that $\bar{y}_1 = 75.67$ and $s_1^2 = 66.67$. Then, with $n_1 = 6$, a 95% CI is given by

$$75.67 \pm 2.571\sqrt{66.67/6} = 75.67 \pm 8.57 \text{ or } (67.10, 84.24).$$

b. The interval computed above is longer than the one in Example 13.3.
c. When only the first sample was used to estimate σ^2, there were only 5 degrees of freedom for error. However, when all four samples were used, there were 14 degrees of freedom for error. Since the critical value $t_{.025}$ is larger in the above, the CI is wider.

13.21 **a.** The 95% CI would be given by

$$\bar{y}_1 - \bar{y}_4 \pm t_{.025} s_{14} \sqrt{\tfrac{1}{n_1} + \tfrac{1}{n_4}},$$

where $s_{14} = \sqrt{\frac{(n_1-1)s_1^2 + (n_4-1)s_4^2}{n_1+n_4-2}} = 7.366$. Since $t_{.025} = 2.306$ based on 8 degrees of freedom, the 95% CI is $-12.08 \pm 2.306(7.366)\sqrt{\tfrac{1}{6} + \tfrac{1}{4}} = -12.08 \pm 10.96$ or $(-23.04, -1.12)$.

b. The CI computed above is longer.

c. The interval computed in Example 13.4 was based on 19 degrees of freedom, and the critical value $t_{.025}$ was smaller.

13.22 **a.** Based on Ex. 13.20 and 13.21, we would expect the CIs to be shorter when all of the data in the one–way layout is used.

b. If the estimate of σ^2 using only one sample is much smaller than the pooled estimate (MSE) – so that the difference in degrees of freedom is offset – the CI width using just one sample could be shorter.

13.23 From Ex. 13.7, the four sample means are (again, using R):

```
> tapply(waste,plant,mean)
    A     B     C     D
1.568 1.772 1.546 1.916
>
```

a. In the above, the sample mean for plant A is 1.568 and from Ex. 13.7, MSE = .0298 with 16 degrees of freedom. Thus, a 95% CI for the mean amount of polluting effluent per gallon for plant A is

$$1.568 \pm 2.12\sqrt{.0298/5} = 1.568 \pm .164 \text{ or } (1.404, 1.732).$$

There is evidence that the plant is exceeding the limit since values larger than 1.5 lb/gal are contained in the CI.

b. A 95% CI for the difference in mean polluting effluent for plants A and D is

$$1.568 - 1.916 \pm 2.12\sqrt{.0298\left(\tfrac{2}{5}\right)} = -.348 \pm .231 \text{ or } (-.579, -.117).$$

Since 0 is not contained in the CI, there is evidence that the means differ for the two plants.

13.24 From Ex. 13.8, the three sample means are (again, using R):

```
> tapply(salary,type,mean)
 church private  public
  56.06   70.78   54.04
```

Also, MSE = 58.61 based on 12 degrees of freedom. A 98% CI for the difference in mean starting salaries for assistant professors at public and private/independent universities is

$$54.04 - 70.78 \pm 2.681\sqrt{58.61\left(\tfrac{2}{5}\right)} = -16.74 \pm 12.98 \text{ or } (-29.72, -3.76).$$

13.25 The 95% CI is given by $.93 - 1.21 \pm 1.96(.1383)\sqrt{2/14} = -.28 \pm .102$ or $(-.382, -.178)$ (note that the degrees of freedom for error is large, so 1.96 is used). There is evidence that the mean densities for the two groups are different since the CI does not contain 0.

13.26 Refer to Ex. 13.9. MSE = .0025 with 8 degrees of freedom.

a. 90% CI for μ_A: $2.25 \pm 1.86\sqrt{.0025/3} = 2.25 \pm .05$ or (2.20, 2.30).

b. 95% CI for $\mu_A - \mu_B$: $2.25 - 2.166 \pm 2.306\sqrt{.0025\left(\tfrac{2}{3}\right)} = .084 \pm .091$ or $(-.007, .175)$.

13.27 Refer to Ex. 13.10. MSE = 62.233 with 8 degrees of freedom.

a. 95% CI for μ_A: $76 \pm 2.306\sqrt{62.333/5} = 76 \pm 8.142$ or (67.868, 84.142).

b. 95% CI for μ_B: $66.33 \pm 2.306\sqrt{62.333/3} = 66.33 \pm 10.51$ or (55.82, 76.84).

c. 95% CI for $\mu_A - \mu_B$: $76 - 66.33 \pm 2.306\sqrt{62.333\left(\tfrac{1}{5}+\tfrac{1}{3}\right)} = 9.667 \pm 13.295$.

13.28 Refer to Ex. 13.12. MSE = 1.852 with 33 degrees of freedom

a. $23.965 \pm 1.96\sqrt{1.852/12} = 23.962 \pm .77$.

b. $23.965 - 20.463 \pm 1.645\sqrt{1.852\left(\tfrac{2}{12}\right)} = 3.502 \pm .914$.

13.29 Refer to Ex. 13.13. MSE = .4735 with 162 degrees of freedom.

a. $6.24 \pm 1.96\sqrt{.4735/18} = 6.24 \pm .318$.

b. $4.59 - 4.58 \pm 1.96\sqrt{.4735\left(\tfrac{1}{45}+\tfrac{1}{102}\right)} = -.29 \pm .241$.

c. Probably not, since the sample was only selected from one town and driving habits can vary from town to town.

13.30 The ANOVA table for these data is below.

Source	d.f	SS	MS	F	p–value
Treatments	3	36.7497	12.2499	4.88	$< .05$
Error	24	60.2822	2.5118		
Total	27	97.0319			

a. Since $F_{.05} = 3.01$ with 3 numerator and 24 denominator degrees of freedom, we reject the hypothesis that the mean wear levels are equal for the four treatments.

b. With $\bar{y}_2 = 14.093$ and $\bar{y}_3 = 12.429$, a 99% CI for the difference in the means is

$$14.093 - 12.429 \pm 2.797\sqrt{2.5118\left(\tfrac{2}{7}\right)} = 1.664 \pm 2.3695.$$

c. A 90% CI for the mean wear with treatment A is
$11.986 \pm 1.711\sqrt{2.5118\left(\tfrac{1}{7}\right)} = 11.986 \pm 1.025$ or (10.961, 13.011).

13.31 The ANOVA table for these data is below.

Source	d.f	SS	MS	F	p–value
Treatments	3	18.1875	2.7292	1.32	> .1
Error	12	24.75	2.0625		
Total	15	32.9375			

a. Since $F_{.05} = 3.49$ with 3 numerator and 12 denominator degrees of freedom, we fail to reject the hypothesis that the mean amounts are equal.

b. The methods of interest are 1 and 4. So, with $\bar{y}_1 = 2$ and $\bar{y}_4 = 4$, a 95% CI for the difference in the mean levels is
$2 - 4 \pm 2.052\sqrt{2.0625\left(\tfrac{2}{4}\right)} = -2 \pm 2.21$ or (–.21, 4.21).

13.32 Refer to Ex. 13.14. MSE = .000183 with 27 degrees of freedom. A 95% CI for the mean residue from DDT is $.041 \pm 2.052\sqrt{.000183/10} = .041 \pm .009$ or (.032, .050).

13.33 Refer to Ex. 13.15. MSE = .041 with 15 degrees of freedom. A 95% CI for the difference in mean O2 content for midstream and adjacent locations is

$$6.44 - 4.78 \pm 2.131\sqrt{.041\left(\tfrac{2}{5}\right)} = 1.66 \pm .273 \text{ or } (1.39, 1.93).$$

13.34 The estimator for $\theta = \frac{1}{2}(\mu_1 + \mu_2) - \mu_4$ is $\hat{\theta} = \frac{1}{2}(\bar{y}_1 + \bar{y}_2) - \bar{y}_4$. So, $V(\hat{\theta}) = \frac{1}{4}\left(\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}\right) + \frac{\sigma^2}{n_4}$.
A 95% CI for θ is given by $\frac{1}{2}(\bar{y}_1 + \bar{y}_2) - \bar{y}_4 \pm t_{.025}\sqrt{\text{MSE}\left(\frac{1}{4n_1} + \frac{1}{4n_2} + \frac{1}{n_4}\right)}$. Using the supplied data, this is found to be .235 ± .255.

13.35 Refer to Ex. 13.16. MSE = 25.986 with 36 degrees of freedom.

a. A 90% CI for the difference in mean heart rate increase for the 1st and 4th groups is

$$30.9 - 28.2 \pm 1.645\sqrt{25.986\left(\tfrac{2}{10}\right)} = 2.7 \pm 3.75.$$

b. A 90% CI for the 2nd group is

$$27.5 \pm 1.645\sqrt{25.986/10} = 27.5 \pm 2.652 \text{ or } (24.85, 30.15).$$

13.36 See Sections 12.3 and 13.7.

13.37 **a.** $\frac{1}{bk}\sum_{i=1}^{k}\sum_{j=1}^{b} E(Y_{ij}) = \frac{1}{bk}\sum_{i=1}^{k}\sum_{j=1}^{b}(\mu + \tau_i + \beta_j) = \frac{1}{bk}(bk\mu + b\sum_{i=1}^{k}\tau_i + k\sum_{j=1}^{b}\beta_j = \mu$.

b. The parameter μ represents the overall mean.

13.38 We have that:

$$\bar{Y}_{i\bullet} = \tfrac{1}{b}\sum_{j=1}^{b} Y_{ij} = \tfrac{1}{b}\sum_{j=1}^{b}(\mu + \tau_i + \beta_j + \varepsilon_{ij})$$
$$= \mu + \tau_i + \tfrac{1}{b}\sum_{j=1}^{b}\beta_j + \tfrac{1}{b}\sum_{j=1}^{b}\varepsilon_{ij} = \mu + \tau_i + \tfrac{1}{b}\sum_{j=1}^{b}\varepsilon_{ij}\ .$$

Thus: $E(\bar{Y}_{i\bullet}) = \mu + \tau_i + \tfrac{1}{b}\sum_{j=1}^{b} E(\varepsilon_{ij}) = \mu + \tau_i = \mu_i$, so $\bar{Y}_{i\bullet}$ is an unbiased estimator.

$$V(\bar{Y}_{i\bullet}) = \tfrac{1}{b^2}\sum_{j=1}^{b} V(\varepsilon_{ij}) = \tfrac{1}{b}\sigma^2\ .$$

13.39 Refer to Ex. 13.38.

a. $E(\bar{Y}_{i\bullet} - \bar{Y}_{i'\bullet}) = \mu + \tau_i - (\mu + \tau_{i'}) = \tau_i - \tau_{i'}$.

b. $V(\bar{Y}_{i\bullet} - \bar{Y}_{i'\bullet}) = V(\bar{Y}_{i\bullet}) + V(\bar{Y}_{i'\bullet}) = \tfrac{2}{b}\sigma^2$, since $\bar{Y}_{i\bullet}$ and $\bar{Y}_{i'\bullet}$ are independent.

13.40 Similar to Ex. 13.38, we have that

$$\bar{Y}_{\bullet j} = \tfrac{1}{k}\sum_{i=1}^{k} Y_{ij} = \tfrac{1}{k}\sum_{i=1}^{k}(\mu + \tau_i + \beta_j + \varepsilon_{ij})$$
$$= \mu + \tfrac{1}{k}\sum_{i=1}^{k}\tau_i + \beta_j + \tfrac{1}{k}\sum_{i=1}^{k}\varepsilon_{ij} = \mu + \beta_j + \tfrac{1}{k}\sum_{i=1}^{k}\varepsilon_{ij}\ .$$

a. $E(\bar{Y}_{\bullet j}) = \mu + \beta_j = \mu_j$, $V(\bar{Y}_{\bullet j}) = \tfrac{1}{k^2}\sum_{i=1}^{k} V(\varepsilon_{ij}) = \tfrac{1}{k}\sigma^2$.

b. $E(\bar{Y}_{\bullet j} - \bar{Y}_{\bullet j'}) = \mu + \beta_j - (\mu + \beta_{j'}) = \beta_j - \beta_{j'}$.

c. $V(\bar{Y}_{\bullet j} - \bar{Y}_{\bullet j'}) = V(\bar{Y}_{\bullet j}) + V(\bar{Y}_{\bullet j'}) = \tfrac{2}{k}\sigma^2$, since $\bar{Y}_{\bullet j}$ and $\bar{Y}_{\bullet j'}$ are independent.

13.41 The sums of squares are Total SS = 1.7419, SST = .0014, SSB = 1.7382, and SSE = .0023. The ANOVA table is given below:

Source	d.f	SS	MS	F
Program	5	1.7382	.3476	772.4
Treatments	1	.0014	.0014	3.11
Error	5	.0023	.00045	
Total	11	1.7419		

a. To test H_0: $\mu_1 = \mu_2$, the F–statistic is $F = 3.11$ with 1 numerator and 5 denominator degrees of freedom. Since $F_{.05} = 6.61$, we fail to reject the hypothesis that the mean CPU times are equal. This is the same result as Ex. 12.10(b).

b. From Table 7, p–value > .10.

c. Using the Applet, p–value = $P(F > 3.11) = .1381$.

d. Ignoring the round–off error, s_D^2 = 2MSE.

13.42 Using the formulas from this section, TSS = 674 – 588 = 86, SSB = $\frac{20^2+36^2+28^2}{4}$ – CM = 32, SST = $\frac{21^2+\ldots+18^2}{3}$ – CM = 42. Thus, SSE = 86 – 32 – 42 = 12. The remaining calculations are given in the ANOVA table below.

Source	d.f	SS	MS	F
Treatments	3	42	14	7
Blocks	2	32	16	
Error	6	12	2	
Total	11	86		

The F–statistic is $F = 7$ with 3 and 6 degrees of freedom. With $\alpha = .05$, $F_{.05} = 4.76$ so we can reject the hypothesis that the mean resistances are equal. Also, $.01 < p\text{–value} < .025$ from Table 7.

13.43 Since the four chemicals (the treatment) were applied to three different materials, the material type could add unwanted variation to the analysis. So, material type was treated as a blocking variable.

13.44 Here, R will be used to analyze the data. We will use the letters A, B, C, and D to denote the location and the numbers 1, 2, 3, 4, and 5 to denote the company.

```
> rate <- c(736, 745, 668, 1065, 1202, 836, 725, 618, 869, 1172, 1492,
1384,1214, 1502, 1682, 996, 884, 802, 1571, 1272)
> location <- factor(c(rep(“A”,5),rep(“B”,5),rep(“C”,5),rep(“D”,5)))
> company <- factor(c(1:5,1:5,1:5,1:5))
> summary(aov(rate ~ company + location))
            Df  Sum Sq Mean Sq F value    Pr(>F)
company      4  731309  182827  12.204 0.0003432 ***
location     3 1176270  392090  26.173 1.499e-05 ***
Residuals   12  179769   14981
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

a. This is a randomized block design (applied to sampled data).
b. The F–statistic is $F = 26.173$ with a p–value of .00001499. Thus, we can safely conclude that there is a difference in mean premiums.
c. The F–statistic is $F = 12.204$ with a p–value of .0003432. Thus, we can safely conclude that there is a difference in the locations.
d. See parts b and c above.

13.45 The treatment of interest is the soil preparation and the location is a blocking variable. The summary statistics are:
$CM = (162)^2/12 = 2187$, $TSS = 2298 - CM = 111$, $SST = 8900/4 - CM = 38$,
$SSB = 6746/3 - CM = 61.67$. The ANOVA table is below.

Source	d.f	SS	MS	F
Treatments	2	38	19	10.05
Blocks	3	61.67	20.56	10.88
Error	6	11.33	1.89	
Total	11	111		

a. The F–statistic for soil preparations is $F = 10.05$ with 2 numerator and 6 denominator degrees of freedom. From Table 7, p–value < .025 so we can reject the null hypothesis that the mean growth is equal for all soil preparations.

b. The F–statistic for the locations is $F = 10.88$ with 3 numerator and 6 denominator degrees of freedom. Here, p–value < .01 so we can reject the null hypothesis that the mean growth is equal for all locations.

13.46 The ANOVA table is below.

Source	d.f	SS	MS	F
Treatments	4	.452	.113	8.37
Blocks	3	1.052	.3507	25.97
Error	12	.162	.0135	
Total	19	1.666		

a. To test for a difference in the varieties, the F–statistic is $F = 8.37$ with 4 numerator and 12 denominator degrees of freedom. From Table 7, p–value < .005 so we would reject the null hypothesis at $\alpha = .05$.

b. The F–statistic for blocks is 25.97 with 3 numerator and 12 denominator degrees of freedom. Since $F_{.05} = 3.49$, we reject the hypothesis of no difference between blocks.

13.47 Using a randomized block design with locations as blocks, the ANOVA table is below.

Source	d.f	SS	MS	F
Treatments	3	8.1875	2.729	1.40
Blocks	3	7.1875	2.396	1.23
Error	9	17.5625	1.95139	
Total	15	32.9375		

With 3 numerator and 9 denominator degrees of freedom, $F_{.05} = 3.86$. Thus, neither the treatment effect nor the blocking effect is significant.

13.48 Note that there are $2bk$ observations. So, let y_{ijl} denote the l^{th} observation in the j^{th} block receiving the i^{th} treatment. Therefore, with CM = $\dfrac{\left(\sum_{i,j,l} y_{ijl}\right)^2}{2bk}$,

TSS = $\sum_{i,j,l} y_{ijl}^2$ – CM with $2bk - 1$ degrees of freedom,

SST = $\dfrac{\sum_i y_{i\bullet\bullet}^2}{2b}$ – CM with $k - 1$ degrees of freedom,

SSB = $\dfrac{\sum_j y_{\bullet j\bullet}^2}{2k}$, with $b - 1$ degrees of freedom, and

SSE = TSS – SST – SSB with $2bk - b - k - 1$ degrees of freedom.

13.49 Using a randomized block design with ingots as blocks, the ANOVA table is below.

Source	d.f	SS	MS	F
Treatments	2	131.901	65.9505	6.36
Blocks	6	268.90	44.8167	
Error	12	124.459	10.3716	
Total	20	524.65		

To test for a difference in the mean pressures for the three bonding agents, the F–statistic is $F = 6.36$ with 2 numerator and 12 denominator degrees of freedom. Since $F_{.05} = 3.89$, we can reject H_0.

13.50 Here, R will be used to analyze the data. The carriers are the treatment levels and the blocking variable is the shipment.

```
> time <- c(15.2,14.3, 14.7, 15.1, 14.0, 16.9, 16.4, 15.9, 16.7, 15.6,
17.1, 16.1, 15.7, 17.0, 15.5)     # data is entered going down columns
> carrier <- factor(c(rep("I",5),rep("II",5),rep("III",5)))
> shipment <- factor(c(1:5,1:5,1:5))
> summary(aov(time ~ carrier + shipment))
            Df Sum Sq Mean Sq F value    Pr(>F)
carrier      2 8.8573  4.4287  83.823 4.303e-06 ***
shipment     4 3.9773  0.9943  18.820  0.000393 ***
Residuals    8 0.4227  0.0528
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
>
```

To test for a difference in mean delivery times for the carriers, from the output we have the F–statistic $F = 83.823$ with 2 numerator and 8 denominator degrees of freedom. Since the p–value is quite small, we can conclude there is a difference in mean delivery times between carriers.

A randomized block design was used because different size/weight shipments can also affect the delivery time. In the experiment, shipment type was blocked.

13.51 Some preliminary results are necessary in order to obtain the solution (see Ex. 13.37–40):

(1) $E(Y_{ij}^2) = V(Y_{ij}) + [E(Y_{ij})]^2 = \sigma^2 + (\mu + \tau_i + \beta_j)^2$

(2) With $\overline{Y}_{\bullet\bullet} = \frac{1}{bk}\sum_{i,j} Y_{ij}$, $E(\overline{Y}_{\bullet\bullet}) = \mu$, $V(\overline{Y}_{\bullet\bullet}) = \frac{1}{bk}\sigma^2$, $E(\overline{Y}_{\bullet\bullet}^2) = \frac{1}{bk}\sigma^2 + \mu^2$

(3) With $\overline{Y}_{\bullet j} = \frac{1}{k}\sum_i Y_{ij}$, $E(\overline{Y}_{\bullet j}) = \mu + \beta_j$, $V(\overline{Y}_{\bullet j}) = \frac{1}{k}\sigma^2$, $E(\overline{Y}_{\bullet j}^2) = \frac{1}{k}\sigma^2 + (\mu + \beta_j)^2$

(4) With $\overline{Y}_{i\bullet} = \frac{1}{b}\sum_j Y_{ij}$, $E(\overline{Y}_{i\bullet}) = \mu + \tau_i$, $V(\overline{Y}_{i\bullet}) = \frac{1}{b}\sigma^2$, $E(\overline{Y}_{i\bullet}^2) = \frac{1}{b}\sigma^2 + (\mu + \tau_i)^2$

a. $$E(\text{MST}) = \frac{b}{k-1}E\left[\sum_i (\overline{Y}_{i\bullet} - \overline{Y}_{\bullet\bullet})^2\right] = \frac{b}{k-1}\left[\sum_i E\left(\overline{Y}_{i\bullet}^2\right) - kE\left(\overline{Y}_{\bullet\bullet}^2\right)\right]$$

$$= \frac{b}{k-1}\left[\sum_i\left(\frac{\sigma^2}{b} + \mu^2 + 2\mu\tau_i + \tau_i^2\right) - k\left(\frac{\sigma^2}{bk} + \mu^2\right)\right] = \sigma^2 + \frac{b}{k-1}\sum_i \tau_i^2 .$$

b. $E(\text{MSB}) = \dfrac{k}{b-1} E\left[\sum_j (\overline{Y}_{\bullet j} - Y_{\bullet\bullet})^2\right] = \dfrac{k}{b-1}\left[\sum_j E\left(\overline{Y}_{\bullet j}^2\right) - bE\left(\overline{Y}_{\bullet\bullet}^2\right)\right]$

$$= \frac{k}{b-1}\left[\sum_j \left(\frac{\sigma^2}{k} + \mu^2 + 2\mu\beta_j + \beta_j^2\right) - b\left(\frac{\sigma^2}{bk} + \mu^2\right)\right] = \sigma^2 + \frac{k}{b-1}\sum_j \beta_j^2 .$$

c. Recall that TSS = $\sum_{i,j} Y_{ij}^2 - bk\overline{Y}_{\bullet\bullet}^2$. Thus,

$$E(\text{TSS}) = \sum_{i,j}\left(\sigma^2 + \mu^2 + \tau_i^2 + \beta_j^2\right) - bk\left(\frac{\sigma^2}{bk} + \mu^2\right) = (bk-1)\sigma^2 + b\sum_i \tau_i^2 + k\sum_j \beta_j^2 .$$

Therefore, since E(SSE) = E(TSS) – E(SST) – E(SSB), we have that

$$E(\text{SSE}) = E(\text{TSS}) - (k-1)E(\text{MST}) - (b-1)E(\text{MSB}) = (bk - k - b + 1)\sigma^2.$$

Finally, since $\text{MST} = \dfrac{\text{SSE}}{bk - k - b + 1}$, $E(\text{MST}) = \sigma^2$.

13.52 From Ex. 13.41, recall that MSE = .00045 with 5 degrees of freedom and $b = 6$. Thus, a 95% CI for the difference in mean CPU times for the two computers is

$$1.553 - 1.575 \pm 2.571\sqrt{.00045\left(\tfrac{2}{6}\right)} = -.022 \pm .031 \text{ or } (-.053, .009).$$

This is the same interval computed in Ex. 12.10(c).

13.53 From Ex. 13.42, MSE = 2 with 6 degrees of freedom and $b = 3$. Thus, the 95% CI is

$$7 - 5 \pm 2.447\sqrt{2\left(\tfrac{2}{3}\right)} = 2 \pm 2.83.$$

13.54 From Ex. 13.45, MSE = 1.89 with 6 degrees of freedom and $b = 4$. Thus, the 90% CI is

$$16 - 12.5 \pm 1.943\sqrt{1.89\left(\tfrac{2}{4}\right)} = 3.5 \pm 1.89 \text{ or } (1.61, 5.39).$$

13.55 From Ex. 13.46, MSE = .0135 with 12 degrees of freedom and $b = 4$. The 95% CI is

$$2.689 - 2.544 \pm 2.179\sqrt{.0135\left(\tfrac{2}{4}\right)} = .145 \pm .179.$$

13.56 From Ex. 13.47, MSE = 1.95139 with 9 degrees of freedom and $b = 4$. The 95% CI is

$$2 \pm 2.262\sqrt{1.95139\left(\tfrac{2}{4}\right)} = 2 \pm 2.23.$$

This differs very little from the CI computed in Ex. 13.31(b) (without blocking).

13.57 From Ex. 13.49, MSE = 10.3716 with 12 degrees of freedom and $b = 7$. The 99% CI is

$$71.1 - 75.9 \pm 3.055\sqrt{10.3716\left(\tfrac{2}{7}\right)} = -4.8 \pm 5.259.$$

13.58 Refer to Ex. 13.9. We require an error bound of no more than .02, so we need n such that

$$2\sqrt{\sigma^2\left(\tfrac{2}{n}\right)} \le .02 ,$$

The best estimate of σ^2 is MSE = .0025, so using this in the above we find that $n \ge 50$. So the entire number of observations needed for the experiment is $4n \ge 4(50) = 200$.

13.59 Following Ex. 13.27(a), we require $2\sqrt{\frac{\sigma^2}{n_A}} \le 10$, where $2 \approx t_{.025}$. Estimating σ^2 with MSE = 62.333, the solution is $n_A \ge 2.49$, so at least 3 observations are necessary.

13.60 Following Ex. 13.27(c), we require $2\sqrt{\sigma^2\left(\frac{2}{n}\right)} \le 20$ where $2 \approx t_{.025}$. Estimating σ^2 with MSE = 62.333, the solution is $n \ge 1.24$, so at least 2 observations are necessary. The total number of observations that are necessary is $3n \ge 6$.

13.61 Following Ex. 13.45, we must find b, the number of locations (blocks), such that

$$2\sqrt{\sigma^2\left(\tfrac{2}{b}\right)} \le 1,$$

where $2 \approx t_{.025}$. Estimating σ^2 with MSE = 1.89, the solution is $b \ge 15.12$, so at least 16 locations must be used. The total number of locations needed in the experiment is at least 3(16) = 48.

13.62 Following Ex. 13.55, we must find b, the number of locations (blocks), such that

$$2\sqrt{\sigma^2\left(\tfrac{2}{b}\right)} \le .5,$$

where $2 \approx t_{.025}$. Estimating σ^2 with MSE = 1.95139, the solution is $b \ge 62.44$, so at least 63 locations are needed.

13.63 The CI lengths also depend on the sample sizes n_i and $n_{i'}$, and since these are not equal, the intervals differ in length.

13.64 **a.** From Example 13.9, $t_{.00417} = 2.9439$. A 99.166% CI for $\mu_1 - \mu_2$ is

$$75.67 - 78.43 \pm 2.9439(7.937)\sqrt{\tfrac{1}{6}+\tfrac{1}{7}} = -2.76 \pm 13.00.$$

b. The ratio is $\dfrac{2(12.63)}{2(13.00)} = .97154$.

c. The ratios are equivalent (save roundoff error).

d. If we divide the CI length for $\mu_1 - \mu_3$ (or equivalently the margin of error) found in Ex. 13.9 by the ratio given in part b above, a 99.166% CI for $\mu_1 - \mu_3$ can be found to be

$$4.84 \pm 13.11/.97154 = 4.84 \pm 13.49.$$

13.65 Refer to Ex. 13.13. Since there are three intervals, each should have confidence coefficient 1 – .05/3 = .9833. Since MSE = .4735 with 162 degrees of freedom, a critical value from the standard normal distribution can be used. So, since $\alpha = 1 - .9833 = .0167$, we require $z_{\alpha/2} = z_{.00833} = 2.39$. Thus, for pairs (i, j) of (1, 2), (1, 3) and (2, 3), the CIs are

$$\begin{aligned}
(1,2)&: \quad -0.29 \pm 2.39\sqrt{.4735\left(\tfrac{1}{45}+\tfrac{1}{102}\right)} \quad \text{or } -0.29 \pm .294 \\
(1,3)&: \quad -1.65 \pm 2.39\sqrt{.4735\left(\tfrac{1}{45}+\tfrac{1}{18}\right)} \quad \text{or } -1.65 \pm .459\,. \\
(2,3)&: \quad -1.36 \pm 2.39\sqrt{.4735\left(\tfrac{1}{102}+\tfrac{1}{18}\right)} \quad \text{or } -1.36 \pm .420
\end{aligned}$$

The simultaneous coverage rate is at least 95%. Note that only the interval for (1, 2) contains 0, suggesting that μ_1 and μ_2 could be equal.

13.66 In this case there are three pairwise comparisons to be made. Thus, the Bonferroni technique should be used with $m = 3$.

13.67 Refer to Ex. 13.45. There are three intervals to construct, so with $\alpha = .10$, each CI should have confidence coefficient $1 - .10/3 = .9667$. Since MSE = 1.89 with 6 degrees of freedom, we require $t_{.0167}$ from this t–distribution. As a conservative approach, we will use $t_{.01} = 3.143$ since $t_{.0167}$ is not available in Table 5 (thus, the simultaneous coverage rate is at least 94%). The intervals all have half width $3.143\sqrt{1.89\left(\frac{2}{4}\right)} = 3.06$ so that the intervals are:

$(1, 2)$: -3.5 ± 3.06 or $(-6.56, -.44)$
$(1, 3)$: $.5 \pm 3.06$ or $(-2.56, 3.56)$
$(2, 3)$: 4.0 ± 3.06 or $(.94, 7.06)$

13.68 Following Ex. 13.47, MSE = 1.95139 with 9 degrees of freedom. For an overall confidence level of 95% with 3 intervals, we require $t_{.025/3} = t_{.0083}$. By approximating this with $t_{.01}$, the half width of each interval is $2.821\sqrt{1.95139\left(\frac{2}{4}\right)} = 2.79$. The intervals are:

$(1, 4)$: -2 ± 2.79 or $(-4.79, .79)$
$(2, 4)$: -1 ± 2.79 or $(-3.79, 1.79)$
$(3, 4)$: $-.75 \pm 2.79$ or $(-3.54, 2.04)$

13.69 **a.** $\beta_0 + \beta_3$ is the mean response to treatment A in block III.
b. β_3 is the difference in mean responses to chemicals A and D in block III.

13.70 **a.** The complete model is $Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$, where

$$x_1 = \begin{cases} 1 & \text{if method } A \\ 0 & \text{otherwise} \end{cases}, \qquad x_2 = \begin{cases} 1 & \text{if method } B \\ 0 & \text{otherwise} \end{cases}.$$

Then, we have

$$\boldsymbol{Y} = \begin{bmatrix} 73 \\ 83 \\ 76 \\ 68 \\ 80 \\ 54 \\ 74 \\ 71 \\ 79 \\ 95 \\ 87 \end{bmatrix} \qquad \boldsymbol{X} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \qquad \boldsymbol{X'X} = \begin{bmatrix} 11 & 5 & 3 \\ 5 & 5 & 0 \\ 3 & 0 & 3 \end{bmatrix} \qquad \hat{\boldsymbol{\beta}} = \begin{bmatrix} 87 \\ -11 \\ 20.67 \end{bmatrix}$$

Thus, $\text{SSE}_c = \boldsymbol{Y'Y} - \hat{\boldsymbol{\beta}}'\boldsymbol{X'Y}$ = 65,286 – 54,787.33 = 498.67 with 11 – 3 = 8 degrees of freedom. The reduced model is $Y = \beta_0 + \varepsilon$, so that $\boldsymbol{X}$ is simply a column vector of eleven 1's and $(\boldsymbol{X'X})^{-1} = \frac{1}{11}$. Thus, $\hat{\beta} = \bar{y}$ = 76.3636. Thus, SSE_R = 65,286 – 64,145.455 = 1140.5455. Thus, to test H_0: $\beta_1 = \beta_2 = 0$, the reduced model F–test statistic is

$$F = \frac{(1140.5455 - 498.67)/2}{498.67/8} = 5.15$$

with 2 numerator and 8 denominator degrees of freedom. Since $F_{.05} = 4.46$, we reject H_0.

b. The hypotheses of interest are H_0: $\mu_A - \mu_B = 0$ versus a two–tailed alternative. Since MSE = $\text{SSE}_c/8$ = 62.333, the test statistic is

$$|t| = \frac{|76-66.33|}{\sqrt{62.333\left(\frac{1}{5}+\frac{1}{3}\right)}} = 1.68.$$

Since $t_{.025} = 2.306$, the null hypothesis is not rejected: there is not a significant difference between the two mean levels.

c. For part **a**, from Table 7 we have .025 < p–value < .05. For part **b**, from Table 5 we have 2(.05) < p–value < 2(.10) or .10 < p–value < .20.

13.71 The complete model is $Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \varepsilon$, where x_1 and x_2 are dummy variables for blocks and x_3, x_4, x_5 are dummy variables for treatments. Then,

$$\boldsymbol{Y} = \begin{bmatrix} 5 \\ 3 \\ 8 \\ 4 \\ 9 \\ 8 \\ 13 \\ 6 \\ 7 \\ 4 \\ 9 \\ 8 \end{bmatrix} \quad \boldsymbol{X} = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \boldsymbol{X'X} = \begin{bmatrix} 12 & 4 & 4 & 3 & 3 & 3 \\ 4 & 4 & 0 & 1 & 1 & 1 \\ 4 & 0 & 4 & 1 & 1 & 1 \\ 3 & 1 & 1 & 3 & 0 & 0 \\ 3 & 1 & 1 & 0 & 3 & 0 \\ 3 & 1 & 1 & 0 & 0 & 3 \end{bmatrix} \quad \hat{\boldsymbol{\beta}} = \begin{bmatrix} 6 \\ -2 \\ 2 \\ 1 \\ -1 \\ 4 \end{bmatrix}$$

Thus, SSE_c = 674 – 662 = 12 with 12 – 6 = 6 degrees of freedom. The reduced model is $Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$, where x_1 and x_2 are as defined in the complete model. Then,

$$(\boldsymbol{X'X})^{-1} = \begin{bmatrix} .25 & -.25 & -.25 \\ -.25 & .5 & .25 \\ -.25 & .25 & .5 \end{bmatrix}, \quad \hat{\boldsymbol{\beta}} = \begin{bmatrix} 7 \\ 2 \\ -2 \end{bmatrix}$$

so that SSE_R = 674 – 620 = 54 with 12 – 3 = 9 degrees of freedom. The reduced model F–test statistic is $F = \frac{(54-12)/3}{12/6} = 7$ with 3 numerator and 6 denominator degrees of freedom. Since $F_{.05} = 4.76$, H_0 is rejected: the treatment means are different.

13.72 (Similar to Ex. 13.71). The full model is $Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \varepsilon$, where x_1, x_2, and x_3 are dummy variables for blocks and x_4 and x_5 are dummy variables for treatments. It can be shown that $SSE_c = 2298 - 2286.6667 = 11.3333$ with $12 - 6 = 6$ degrees of freedom. The reduced model is $Y = \beta_0 + \beta_4 x_4 + \beta_5 x_5 + \varepsilon$, and $SSE_R = 2298 - 2225 = 73$ with $12 - 3 = 9$ degrees of freedom. Then, the reduced model F–test statistic is $F = \frac{(73-11.3333)/3}{11.3333/6} = 10.88$ with 3 numerator and 6 denominator degrees of freedom. Since Since $F_{.05} = 4.76$, H_0 is rejected: there is a difference due to location.

13.73 See Section 13.8. The experimental units within each block should be as homogenous as possible.

13.74 **a.** For the CRD, experimental units are randomly assigned to treatments.
b. For the RBD, experimental units are randomly assigned the k treatments *within each block*.

13.75 **a.** Experimental units are the patches of skin, while the three people act as blocks.
b. Here, MST = 1.18/2 = .59 and MSE = 2.24/4 = .56. Thus, to test for a difference in treatment means, calculate $F = .59/.56 = 1.05$ with 2 numerator and 4 denominator degrees of freedom. Since $F_{.05} = 6.94$, we cannot conclude there is a difference.

13.76 Refer to Ex. 13.9. We have that CM = 58.08, TSS = .035, and SST = .015. Then, SSB = $\frac{(8.9)^2+(8.6)^2+(8.9)^2}{4} - \text{CM} = .015$ with 2 degrees of freedom. The ANOVA table is below:

Source	d.f	SS	MS	F
Treatments	3	.015	.00500	6.00
Blocks	2	.015	.00750	9.00
Error	6	.005	.000833	
Total	11	.035		

a. To test for a "sand" effect, this is determined by an F–test for blocks. From the ANOVA table $F = 9.00$ with 2 numerator and 6 denominator degrees of freedom. Since $F_{.05} = 5.14$, we can conclude that the type of sand is important.
b. To test for a "concrete type" effect, from the ANOVA table $F = 6.00$ with 3 numerator and 6 denominator degrees of freedom. Since $F_{.05} = 4.76$, we can conclude that the type of concrete mix used is important.
c. Compare the sizes of SSE from Ex. 13.9 and what was calculated here. Since the experimental error was estimated to be much larger in Ex. 13.9 (by ignoring a block effect), the test for treatment effect was not significant.

13.77 Refer to Ex. 13.76
a. A 95% CI is given by $2.25 - 2.166 \pm 2.447\sqrt{.000833\left(\frac{2}{3}\right)} = .084 \pm .06$ or (.024, .144).
b. Since the SSE has been reduced by accounting for a block effect, the precision has been improved.

13.78 **a.** This is not a randomized block design. There are 9 treatments (one level of drug 1 and one level of drug 2). Since both drugs are factors, there could be interaction present.

b. The second design is similar to the first, except that there are two patients assigned to each treatment in a completely randomized design.

13.79 **a.** We require $2\sigma\frac{1}{\sqrt{n}} \leq 10$, so that $n \geq 16$.

b. With 16 patients assigned to each of the 9 treatments, there are 16(9) – 9 = 135 degrees of freedom left for error.

c. The half width, using $t_{.025} \approx 2$, is given by $2(20)\sqrt{\frac{1}{16}+\frac{1}{16}}$ = 14.14.

13.80 In this experiment, the car model is the treatment and the gasoline brand is the block. Here, we will use R to analyze the data:

```
> distance <- c(22.4, 20.8, 21.5, 17.0, 19.4, 18.7, 19.2, 20.2, 21.2)
> model <- factor(c("A","A","A","B","B","B","C","C","C"))
> gasoline <- factor(c("X","Y","Z","X","Y","Z","X","Y","Z"))
> summary(aov(distance ~ model + gasoline))
            Df  Sum Sq Mean Sq F value  Pr(>F)
model        2 15.4689  7.7344  6.1986 0.05951 .
gasoline     2  1.3422  0.6711  0.5378 0.62105
Residuals    4  4.9911  1.2478
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

a. To test for a car model effect, the *F*–test statistic is *F* = 6.1986 and by the *p*–value this is not significant at the α = .05 level.

b. To test for a gasoline brand effect, the *F*–test statistic is *F* = .5378. With a *p*–value of .62105, this is not significant and so gasoline brand does not affect gas mileage.

13.81 Following Ex. 13.81, the R output is

```
> summary(aov(distance~model))
            Df  Sum Sq Mean Sq F value  Pr(>F)
model        2 15.4689  7.7344  7.3274 0.02451 *
Residuals    6  6.3333  1.0556
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

a. To test for a car model effect, the *F*–test statistic is *F* = 6.1986 with *p*–value = .02451. Thus, with α = .05, we can conclude that the car model has an effect on gas mileage.

b. In the RBD, SSE was reduced (somewhat) but 2 degrees of freedom were lost. Thus MSE is larger in the RBD than in the CRD.

c. The CRD randomly assigns treatments to experimental units. In the RBD, treatments are randomly assigned to experimental units within each block, and this is not the same randomization procedure as a CRD.

13.82 **a.** This is a completely randomized design.

b. The sums of squares are: TSS = 183.059, SST = 117.642, and SSE = 183.059 – 117.642 = 65.417. The ANOVA table is given below

Source	d.f	SS	MS	F
Treatments	3	117.642	39.214	7.79
Error	13	65.417	5.032	
Total	16	183.059		

To test for equality in mean travel times, the F–test statistic is F = 7.79 with 3 numerator and 13 denominator degrees of freedom. With $F_{.01}$ = 5.74, we can reject the hypothesis that the mean travel times are equal.

c. With $\bar{y}_1$ = 26.75 and $\bar{y}_3$ = 32.4, a 95% CI for the difference in means is

$$26.75 - 32.4 \pm 2.160\sqrt{5.032\left(\tfrac{1}{4}+\tfrac{1}{5}\right)} = -5.65 \pm 3.25 \text{ or } (-8.90, -2.40).$$

13.83 This is a RBD with digitalis as the treatment and dogs are blocks.

a. TSS = 703,681.667, SST = 524,177.167, SSB = 173,415, and SSE = 6089.5. The ANOVA table is below.

Source	d.f	SS	MS	F
Treatments	2	524,177.167	262,088.58	258.237
Blocks	3	173,415	57,805.00	56.95
Error	6	6,089.5	1,014.9167	
Total	11	703,681.667		

b. There are 6 degrees of freedom for SSE.

c. To test for a digitalis effect, the F–test has F = 258.237 with 2 numerator and 6 denominator degrees of freedom. From Table 7, p–value < .005 so this is significant.

d. To test for a dog effect, the F–test has F = 56.95 with 3 numerator and 6 denominator degrees of freedom. From Table 7, p–value < .005 so this is significant.

e. The standard deviation of the difference between the mean calcium uptake for two levels of digitalis is $s\sqrt{\frac{1}{n_i}+\frac{1}{n_j}} = \sqrt{1014.9167\left(\tfrac{1}{4}+\tfrac{1}{4}\right)} = 22.527$.

f. The CI is given by $1165.25 - 1402.5 \pm 2.447(22.53) = -237.25 \pm 55.13$.

13.84 We require $2\sqrt{\sigma^2\left(\tfrac{2}{b}\right)} \le 20$. From Ex. 13.83, we can estimate σ^2 with MSE = 1014.9167 so that the solution is $b \ge 20.3$. Thus, at least 21 replications are required.

13.85 The design is completely randomized with five treatments, containing 4, 7, 6, 5, and 5 measurements respectively.

a. The analysis is as follows:

CM = $(20.6)^2/27$ = 15.717

TSS = 17,500 – CM = 1.783

SST = $\frac{(2.5)^2}{4} + \ldots + \frac{(2.4)^2}{5}$ – CM = 1.212, d.f. = 4

SSE = 1.783 – 1.212 = .571, d.f. = 22.

To test for difference in mean reaction times, $F = \frac{1.212/4}{.571/22}$ = 11.68 with 4 numerator and 22 denominator degrees of freedom. From Table 7, p–value < .005.

b. The hypothesis is H_0: $\mu_A - \mu_D = 0$ versus a two–tailed alternative. The test statistic is

$$|t| = \frac{|.625-.920|}{\sqrt{.02596\left(\frac{1}{4}+\frac{1}{5}\right)}} = 2.73.$$

The critical value (based on 22 degrees of freedom) is $t_{.025} = 2.074$. Thus, H_0 is rejected. From Table 5, $2(.005) < p$–value $< 2(.01)$.

13.86 This is a RBD with people as blocks and stimuli as treatments. The ANOVA table is below.

Source	d.f	SS	MS	F
Treatments	4	.787	.197	27.7
Blocks	3	.140	.047	
Error	12	.085	.0071	
Total	19	1.012		

To test for a difference in the mean reaction times, the test statistic is $F = 27.7$ with 4 numerator and 12 denominator degrees of freedom. With $F_{.05} = 3.25$, we can reject the null hypothesis that the mean reaction times are equal.

13.87 Each interval should have confidence coefficient $1 - .05/4 = .9875 \approx .99$. Thus, with 12 degrees of freedom, we will use the critical value $t_{.005} = 3.055$ so that the intervals have a half width given by $3.055\sqrt{.0135\left(\frac{2}{4}\right)} = .251$. Thus, the intervals for the differences in means for the varieties are

$\mu_A - \mu_D$: $.320 \pm .251$ $\quad\quad$ $\mu_B - \mu_D$: $.145 \pm .251$

$\mu_C - \mu_D$: $.023 \pm .251$ $\quad\quad$ $\mu_E - \mu_D$: $-.124 \pm .251$

13.88 $\text{TSS} = \sum_{j=1}^{b}\sum_{i=1}^{k}(Y_{ij} - \overline{Y})^2 = \sum_{j=1}^{b}\sum_{i=1}^{k}(Y_{ij} - Y_{i\bullet} + Y_{i\bullet} - Y_{\bullet j} + Y_{\bullet j} - \overline{Y} + \overline{Y} - \overline{Y})^2$

$= \sum_{j=1}^{b}\sum_{i=1}^{k}(\underline{Y_{i\bullet} - \overline{Y}} + \underline{Y_{\bullet j} - \overline{Y}} + \underline{Y_{ij} - Y_{i\bullet} - Y_{\bullet j} + \overline{Y}})^2$ $\quad\leftarrow$ expand as shown

$= \sum_{j=1}^{b}\sum_{i=1}^{k}(Y_{i\bullet} - \overline{Y})^2 + \sum_{j=1}^{b}\sum_{i=1}^{k}(Y_{\bullet j} - \overline{Y})^2 + \sum_{j=1}^{b}\sum_{i=1}^{k}(Y_{ij} - Y_{i\bullet} - Y_{\bullet j} + \overline{Y})^2$

$\quad$ + cross terms (= C)

$= b\sum_{i=1}^{k}(Y_{i\bullet} - \overline{Y})^2 + k\sum_{j=1}^{b}(Y_{\bullet j} - \overline{Y})^2 + \sum_{j=1}^{b}\sum_{i=1}^{k}(Y_{ij} - Y_{i\bullet} - Y_{\bullet j} + \overline{Y})^2 + C$

$= \text{SST} + \text{SSB} + \text{SSE} + C.$

So, it is only left to show that the cross terms are 0. They are expressed as

$$C = 2\sum_{j=1}^{b}(\overline{Y}_{\bullet j} - \overline{Y})\sum_{i=1}^{k}(\overline{Y}_{i\bullet} - \overline{Y}) \qquad (1)$$

$$+\,2\sum_{j=1}^{b}(\overline{Y}_{\bullet j} - \overline{Y})\sum_{i=1}^{k}(Y_{ij} - \overline{Y}_{i\bullet} - \overline{Y}_{\bullet j} + \overline{Y}) \qquad (2)$$

$$+\,2\sum_{i=1}^{k}(\overline{Y}_{i\bullet} - \overline{Y})\sum_{j=1}^{b}(Y_{ij} - \overline{Y}_{i\bullet} - \overline{Y}_{\bullet j} + \overline{Y})\,. \qquad (3)$$

Part (1) is equal to zero since

$$\sum_{j=1}^{b}(\overline{Y}_{\bullet j} - \overline{Y}) = \sum_{j=1}^{b}\left(\tfrac{1}{k}\Sigma_i Y_{ij} - \tfrac{1}{bk}\Sigma_{ij}Y_{ij}\right) = \tfrac{1}{k}\Sigma_{ij}Y_{ij} - \tfrac{b}{bk}\Sigma_{ij}Y_{ij} = 0.$$

Part (2) is equal to zero since

$$\sum_{i=1}^{k}(Y_{ij}-\overline{Y}_{i\bullet}-\overline{Y}_{\bullet j}+\overline{Y}) = \sum_{i=1}^{k}\left(Y_{ij}-\tfrac{1}{b}\Sigma_j Y_{ij}-\tfrac{1}{k}\Sigma_i Y_{ij}+\tfrac{1}{bk}\Sigma_{ij}Y_{ij}\right)$$
$$= \Sigma_i Y_{ij}-\tfrac{1}{b}\Sigma_{ij}Y_{ij}-\Sigma_i Y_{ij}+\tfrac{1}{b}\Sigma_{ij}Y_{ij}=0.$$

A similar expansion will shown that part (3) is also equal to 0, proving the result.

13.89 **a.** We have that Y_{ij} and $Y_{ij'}$ are normally distributed. Thus, they are independent if their covariance is equal to 0 (recall that this only holds for the normal distribution). Thus,

$$\text{Cov}(Y_{ij},Y_{ij'}) = \text{Cov}(\mu+\tau_i+\beta_j+\varepsilon_{ij},\mu+\tau_i+\beta_{j'}+\varepsilon_{ij'}) = \text{Cov}(\beta_j+\varepsilon_{ij},\beta_{j'}+\varepsilon_{ij'})$$
$$= \text{Cov}(\beta_j,\beta_{j'})+\text{Cov}(\beta_j,\varepsilon_{ij'})+\text{Cov}(\varepsilon_{ij},\beta_{j'})+\text{Cov}(\varepsilon_{ij},\varepsilon_{ij'}) = 0,$$

by independence specified in the model. The result is similar for Y_{ij} and $Y_{i'j'}$.

b. $\text{Cov}(Y_{ij},Y_{i'j}) = \text{Cov}(\mu+\tau_i+\beta_j+\varepsilon_{ij},\mu+\tau_{i'}+\beta_j+\varepsilon_{i'j}) = \text{Cov}(\beta_j+\varepsilon_{ij},\beta_j+\varepsilon_{i'j})$
$= V(\beta_j) = \sigma_B^2$, by independence of the other terms.

c. When $\sigma_B^2 = 0$, $\text{Cov}(Y_{ij},Y_{i'j}) = 0$.

13.90 **a.** From the model description, it is clear that $E(Y_{ij}) = \mu + \tau_i$ and $V(Y_{ij}) = \sigma_B^2+\sigma_\varepsilon^2$.

b. Note that $\overline{Y}_{i\bullet}$ is the mean of b independent observations in a block. Thus,

$$E(\overline{Y}_{i\bullet}) = E(Y_{ij}) = \mu+\tau_i \text{ (unbiased) and } V(\overline{Y}_{i\bullet}) = \tfrac{1}{b}V(Y_{ij}) = \tfrac{1}{b}(\sigma_B^2+\sigma_\varepsilon^2).$$

c. From part b above, $E(\overline{Y}_{i\bullet}-\overline{Y}_{i'\bullet}) = \mu+\tau_i-(\mu+\tau_{i'}) = \tau_i-\tau_{i'}$.

d. $$V(\overline{Y}_{i\bullet}-\overline{Y}_{i'\bullet}) = V\left[\mu+\tau_i+\frac{1}{b}\sum_{j=1}^{b}\beta_j+\frac{1}{b}\sum_{i=1}^{b}\varepsilon_{ij}-\left(\mu+\tau_{i'}+\frac{1}{b}\sum_{j=1}^{b}\beta_j+\frac{1}{b}\sum_{i=1}^{b}\varepsilon_{i'j}\right)\right]$$
$$= V\left[\frac{1}{b}\sum_{i=1}^{b}\varepsilon_{ij}-\frac{1}{b}\sum_{i=1}^{b}\varepsilon_{i'j}\right] = \frac{1}{b^2}V\left[\sum_{i=1}^{b}\varepsilon_{ij}\right]+\frac{1}{b^2}V\left[\sum_{i=1}^{b}\varepsilon_{i'j}\right] = \frac{2\sigma_\varepsilon^2}{b}.$$

13.91 First, $\overline{Y}_{\bullet j} = \frac{1}{k}\sum_{i=1}^{k}(\mu+\tau_i+\beta_j+\varepsilon_{ij}) = \mu+\frac{1}{k}\sum_{i=1}^{k}\tau_i+\beta_j+\frac{1}{k}\sum_{i=1}^{k}\varepsilon_{ij} = \mu+\beta_j+\frac{1}{k}\sum_{i=1}^{k}\varepsilon_{ij}$.

a. Using the above, $E(\overline{Y}_{\bullet j}) = \mu$ and $V(\overline{Y}_{\bullet j}) = V(\beta_j)+\frac{1}{k^2}\sum_{i=1}^{k}V(\varepsilon_{ij}) = \sigma_B^2+\frac{1}{k}\sigma_\varepsilon^2$.

b. $E(\text{MST}) = \sigma_\varepsilon^2+\left(\frac{b}{k-1}\right)\sum_{i=1}^{k}\tau_i^2$ as calculated in Ex. 13.51, since the block effects cancel here as well.

c. $$E(\text{MSB}) = kE\left[\frac{\sum_{j=1}^{b}(\overline{Y}_{\bullet j}-\overline{Y})^2}{b-1}\right] = \sigma_\varepsilon^2+k\sigma_B^2$$

d. $E(\text{MSE}) = \sigma_\varepsilon^2$, using a similar derivation in Ex. 13.51(c).

13.92 **a.** $\hat{\sigma}^2_\varepsilon = \text{MSE}$.

b. $\hat{\sigma}^2_B = \dfrac{\text{MSB - MSE}}{k}$. By Ex. 13.91, this estimator is unbiased.

13.93 **a.** The vector $\boldsymbol{AY}$ can be displayed as

$$\boldsymbol{AY} = \begin{bmatrix} \dfrac{\sum_i Y_i}{\sqrt{n}} \\ \dfrac{Y_1 - Y_2}{\sqrt{2}} \\ \dfrac{Y_1 + Y_2 - 2Y_3}{\sqrt{2\cdot 3}} \\ \vdots \\ \dfrac{(Y_1 + Y_2 + \ldots + Y_{n-1} - (n-1)Y_n}{\sqrt{n(n-1)}} \end{bmatrix} = \begin{bmatrix} \sqrt{n}\overline{Y} \\ U_1 \\ U_2 \\ \vdots \\ U_{n-1} \end{bmatrix}$$

Then, $\sum_{i=1}^{n} Y_i^2 = \boldsymbol{Y'Y} = \boldsymbol{Y'A'AY} = n\overline{Y}^2 + \sum_{i=1}^{n-1} U_i$.

b. Write $L_i = \sum_{j=1}^{n} a_{ij} Y_j$, a linear function of $Y_1, \ldots, Y_n$. Two such linear functions, say L_i and L_k are pairwise orthogonal if and only if $\sum_{j=1}^{n} a_{ij} a_{kj} = 0$ and so L_i and L_k are independent (see Chapter 5). Let $L_1, L_2, \ldots, L_n$ be the n linear functions in $\boldsymbol{AY}$. The constants a_{ij}, $j = 1, 2, \ldots, n$ are the elements of the i^{th} row of the matrix $\boldsymbol{A}$. Moreover, if any two rows of the matrix $\boldsymbol{A}$ are multiplied together, the result is zero (try it!). Thus, L_1, $L_2, \ldots, L_n$ are independent linear functions of $Y_1, \ldots, Y_n$.

c. $\sum_{i=1}^{n}(Y_i - \overline{Y})^2 = \sum_{i=1}^{n} Y_i^2 - n\overline{Y}^2 = n\overline{Y}^2 + \sum_{i=1}^{n-1} U_i - n\overline{Y}^2 = \sum_{i=1}^{n-1} U_i$. Since U_i is independent of $\sqrt{n}\overline{Y}$ for $i = 1, 2, \ldots, n - 1$, $\sum_{i=1}^{n}(Y_i - \overline{Y})^2$ and $\overline{Y}$ are independent.

d. Define

$$W = \frac{\sum_{i=1}^{n}(Y_i - \mu)^2}{\sigma^2} = \frac{\sum_{i=1}^{n}(Y_i - \overline{Y} + \overline{Y} - \mu)^2}{\sigma^2} = \frac{\sum_{i=1}^{n}(Y_i - \overline{Y})^2}{\sigma^2} + \frac{n(\overline{Y} - \mu)^2}{\sigma^2} = X_1 + X_2 .$$

Now, W is chi–square with n degrees of freedom, and X_2 is chi–square with 1 degree of freedom since $X_2 = \left(\dfrac{\overline{Y} - \mu}{\sigma / \sqrt{n}}\right)^2 = Z^2$. Since X_1 and X_2 are independent (from part **c**), we can use moment generating functions to show that

$$(1 - 2t)^{-n/2} = m_W(t) = m_{X_1}(t) m_{X_2}(t) = m_{X_1}(t)(1 - 2t)^{-1/2} .$$

Thus, $m_{X_1}(t) = (1 - 2t)^{-(n-1)/2}$ and this is seen to be the mgf for the chi–square distribution with n – 1 degrees of freedom, proving the result.

13.94 **a.** From Section 13.3, SSE can be written as $\text{SSE} = \sum_{i=1}^{k}(n_i - 1)S_i^2$. From Ex. 13.93, each $\overline{Y}_i$ is independent of $S_i^2 = \sum_{j=1}^{n_i}(Y_{ij} - \overline{Y}_i)^2$. Therefore, since the k samples are independent, $\overline{Y}_1, \ldots, \overline{Y}_k$ are independent of SSE.

b. Note that SST = $\sum_{i=1}^{k} n_i(\overline{Y}_i - \overline{Y})^2$, and $\overline{Y}$ can be written as

$$\overline{Y} = \frac{\sum_{i=1}^{k} n_i \overline{Y}_i}{n}.$$

Since SST can be expressed as a function of only $\overline{Y}_1, \ldots, \overline{Y}_k$, by part (a) above we have that SST and SSE are independent. The distribution of $F = \dfrac{\text{MST}}{\text{MSE}}$ was derived in Ex. 13.6.

Chapter 14: Analysis of Categorical Data

14.1 **a.** H_0: $p_1 = .41$, $p_2 = .10$, $p_3 = .04$, $p_4 = .45$ vs. H_a: not H_0. The observed and expected counts are:

	A	B	AB	O
observed	89	18	12	81
expected	200(.41) = 82	200(.10) = 20	200(.04) = 8	200(.45) = 90

The chi–square statistic is $X^2 = \frac{(89-82)^2}{82} + \frac{(18-20)^2}{20} + \frac{(12-8)^2}{8} + \frac{(81-90)^2}{90} = 3.696$ with 4 –1 = 3 degrees of freedom. Since $\chi^2_{.05} = 7.81473$, we fail to reject H_0; there is not enough evidence to conclude the proportions differ.

b. Using the Applet, p–value = $P(\chi^2 > 3.696) = .29622$.

14.2 **a.** H_0: $p_1 = .60$, $p_2 = .05$, $p_3 = .35$ vs. H_a: not H_0. The observed and expected counts are:

	admitted unconditionally	admitted conditionally	refused
observed	329	43	128
expected	500(.60) = 300	500(.05) = 25	500(.35) = 175

The chi–square test statistic is $X^2 = \frac{(329-300)^2}{300} + \frac{(43-25)^2}{25} + \frac{(128-175)^2}{175} = 28.386$ with 3 – 1 = 2 degrees of freedom. Since $\chi^2_{.05} = 7.37776$, we can reject H_0 and conclude that the current admission rates differ from the previous records.

b. Using the Applet, p–value = $P(\chi^2 > 28.386) = .00010$.

14.3 The null hypothesis is H_0: $p_1 = p_2 = p_3 = p_4 = \frac{1}{4}$ vs. H_a: not H_0. The observed and expected counts are:

lane	1	2	3	4
observed	294	276	238	192
expected	250	250	250	250

The chi–square statistic is $X^2 = \frac{(294-250)^2 + (276-250)^2 + (238-250)^2 + (192-250)^2}{250} = 24.48$ with 4 –1 = 3 degrees of freedom. Since $\chi^2_{.05} = 7.81473$, we reject H_0 and conclude that the lanes are not preferred equally. From Table 6, p–value < .005.

Note that `R` can be used by:

```
> lanes <- c(294,276,238,192)
> chisq.test(lanes,p = c(.25,.25,.25,.25))    # p is not necessary here

        Chi-squared test for given probabilities

data:  lanes
X-squared = 24.48, df = 3, p-value = 1.983e-05
```

14.4 The null hypothesis is H_0: $p_1 = p_2 = \ldots = p_7 = \frac{1}{7}$ vs. H_a: not H_0. The observed and expected counts are:

	SU	M	T	W	R	F	SA
observed	24	36	27	26	32	26	29
expected	28.571	28.571	28.571	28.571	28.571	28.571	28.571

The chi–square statistic is $X^2 = \frac{(24-28.571)^2+(36-28.571)^2+\ldots+(29-28.571)^2}{28.571}$ = 24.48 with 7 –1 = 6 degrees of freedom. Since $\chi^2_{.05}$ = 12.5916, we can reject the null hypothesis and conclude that there is evidence of a difference in percentages of heart attacks for the days of the week

14.5 **a.** Let p = proportion of heart attacks on Mondays. Then, H_0: $p = \frac{1}{7}$ vs. H_a: $p > \frac{1}{7}$. Then, $\hat{p}$ = 36/200 = .18 and from Section 8.3, the test statistic is

$$z = \frac{.18-1/7}{\sqrt{\frac{(1/7)(6/7)}{200}}} = 1.50.$$

Since $z_{.05}$ = 1.645, we fail to reject H_0.

b. The test was suggested by the data, and this is known as "data snooping" or "data dredging." We should always apply the scientific method: first form a hypothesis and then collect data to test the hypothesis.

c. Monday has often been referred to as the most stressful workday of the week: it is the day that is farthest from the weekend, and this realization gets to some people.

14.6 **a.** $E(n_i - n_j) = E(n_i) - E(n_j) = np_i - np_j$.

b. Define the sample proportions $\hat{p}_i = n_i / n$ and $\hat{p}_j = n_j / n$. Then, $\hat{p}_i - \hat{p}_j$ is unbiased for $p_i - p_j$ from part **a** above.

c. $V(n_i - n_j) = V(n_i) + V(n_j) - 2\text{Cov}(n_i, n_j) = np_i(1-p_i) + np_j(1-p_j) + 2np_i p_j$.

d. $V(\hat{p}_i - \hat{p}_j) = \frac{1}{n^2}V(n_i - n_j) = \frac{1}{n}\left(p_i(1-p_i) + p_j(1-p_j) + 2p_i p_j\right)$.

e. A consistent estimator is one that is unbiased and whose variance tends to 0 as the sample size increases. Thus, $\hat{p}_i - \hat{p}_j$ is a consistent estimator.

f. Given the information in the problem and for large n, the quantity

$$Z_n = \frac{\hat{p}_i - \hat{p}_j - (p_i - p_j)}{\sigma_{\hat{p}_i - \hat{p}_j}}$$

is approx. normally distributed, where $\sigma_{\hat{p}_i - \hat{p}_j} = \sqrt{\frac{1}{n}\left(p_i(1-p_i) + p_j(1-p_j) + 2p_i p_j\right)}$.

Now, since $\hat{p}_i$ and $\hat{p}_j$ are consistent estimators,

$$W_n = \frac{\sigma_{\hat{p}_i - \hat{p}_j}}{\hat{\sigma}_{\hat{p}_i - \hat{p}_j}} = \frac{\sqrt{\frac{1}{n}\left(p_i(1-p_i) + p_j(1-p_j) + 2p_i p_j\right)}}{\sqrt{\frac{1}{n}\left(\hat{p}_i(1-\hat{p}_i) + \hat{p}_j(1-\hat{p}_j) + 2\hat{p}_i \hat{p}_j\right)}}$$

tends to 1 (see Chapter 9). Therefore, the quantity

$$Z_n W_n = \frac{\hat{p}_i - \hat{p}_j - (p_i - p_j)}{\sigma_{\hat{p}_i - \hat{p}_j}} \left(\frac{\sigma_{\hat{p}_i - \hat{p}_j}}{\hat{\sigma}_{\hat{p}_i - \hat{p}_j}} \right) = \frac{\hat{p}_i - \hat{p}_j - (p_i - p_j)}{\sqrt{\frac{1}{n}\left(\hat{p}_i(1-\hat{p}_i) + \hat{p}_j(1-\hat{p}_j) + 2\hat{p}_i\hat{p}_j\right)}}$$

has a limiting standard normal distribution by Slutsky's Theorem. The expression for the confidence interval follows directly from the above.

14.7 From Ex. 14.3, $\hat{p}_1 = .294$ and $\hat{p}_4 = .192$. A 95% (large sample) CI for $p_1 - p_4$ is

$$.294 - .192 \pm 1.96\sqrt{\frac{.294(.706) + .192(.808) + 2(.294)(.192)}{1000}} = .102 \pm .043 \text{ or } (.059, .145).$$

There is evidence that a greater proportion use the "slow" lane since the CI does not contain 0.

14.8 The hypotheses are H_0: ratio is 9:3:3:1 vs. H_a: not H_0. The observed and expected counts are:

category	1 (RY)	2 (WY)	3 (RG)	4 (WG)
observed	56	19	17	8
expected	56.25	18.75	18.75	6.25

The chi–square statistic is $X^2 = \frac{(56-56.25)^2}{56.25} + \frac{(19-18.75)^2}{18.75} + \frac{(17-18.75)^2}{18.75} + \frac{(8-6.25)^2}{6.25} = .658$ with 3 degrees of freedom. Since $\chi^2_{.05} = 7.81473$, we fail to reject H_0: there is not enough evidence to conclude the ratio is not 9:3:3:1.

14.9 **a.** From Ex. 14.8, $\hat{p}_1 = .56$ and $\hat{p}_3 = .17$. A 95% (large sample) CI for $p_1 - p_3$ is

$$.56 - .17 \pm 1.96\sqrt{\frac{.56(.44) + .17(.83) + 2(.56)(.17)}{100}} = .39 \pm .149 \text{ or } (.241, .539).$$

b. There are three intervals to construct: $p_1 - p_2$, $p_1 - p_3$, and $p_1 - p_4$. So that the simultaneous confidence coefficient is at least 95%, each interval should have confidence coefficient 1 – (.05/3) = .98333. Thus, we require the critical value $z_{.00833} = 2.39$. The three intervals are

$$.56 - .19 \pm 2.39\sqrt{\frac{.56(.44) + .19(.81) + 2(.56)(.19)}{100}} = .37 \pm .187$$

$$.56 - .17 \pm 2.39\sqrt{\frac{.56(.44) + .17(.83) + 2(.56)(.17)}{100}} = .39 \pm .182$$

$$.56 - .08 \pm 2.39\sqrt{\frac{.56(.44) + .08(.92) + 2(.56)(.08)}{100}} = .48 \pm .153.$$

14.10 The hypotheses of interest are H_0: $p_1 = .5$, $p_2 = .2$, $p_3 = .2$, $p_4 = .1$ vs. H_a: not H_0. The observed and expected counts are:

defect	1	2	3	4
observed	48	18	21	13
expected	50	20	20	10

It is found that $X^2 = 1.23$ with 3 degrees of freedom. Since $\chi^2_{.05} = 7.81473$, we fail to reject H_0; there is not enough evidence to conclude the proportions differ.

14.11 This is similar to Example 14.2. The hypotheses are H_0: Y is Poisson(λ) vs. H_a: not H_0. Using $\bar{y}$ to estimate λ, calculate $\bar{y} = \frac{1}{400}\Sigma_i y_i f_i = 2.44$. The expected cell counts are estimated as $\hat{E}(n_i) = n\hat{p}_i = 400\frac{(2.44)^{y_i}\exp(-2.44)}{y_i!}$. However, after $Y = 7$, the expected cell count drops below 5. So, the final group will be compiled as $\{Y \geq 7\}$. The observed and (estimated) expected cell counts are below:

# of colonies	n_i	$\hat{p}_i$	$\hat{E}(n_i)$
0	56	.087	34.86
1	104	.2127	85.07
2	80	.2595	103.73
3	62	.2110	84.41
4	42	.1287	51.49
5	27	.0628	25.13
6	9	.0255	10.22
7 or more	20		400 – 394.96 = 5.04

The chi–square statistic is $X^2 = \frac{(56-34.86)^2}{34.86} + \ldots + \frac{(20-5.04)^2}{5.04} = 69.42$ with 8 – 2 = 6 degrees of freedom. Since $\chi^2_{.05} = 12.59$, we can reject H_0 and conclude that the observations do not follow a Poisson distribution.

14.12 This is similar to Ex. 14.11. First, $\bar{y} = \frac{1}{414}\Sigma_i y_i f_i = 0.48309$. The observed and (estimated) expected cell counts are below; here, we collapsed cells into $\{Y \geq 3\}$:

# of accidents	n_i	$\hat{p}_i$	$\hat{E}(n_i)$
0	296	.6169	255.38
1	74	.298	123.38
2	26	.072	29.80
3	18	.0131	5.44

Then, $X^2 = \frac{(296-255.38)^2}{255.38} + \ldots + \frac{(18-5.44)^2}{5.44} = 55.71$ with 4 – 2 = 2 degrees of freedom. Since $\chi^2_{.05} = 5.99$, we can reject the claim that this is a sample from a Poisson distribution.

14.13 The contingency table with observed and expected counts is below.

	All facts known	Some facts withheld	Not sure	Total
Democrat	42 (53.48)	309 (284.378)	31 (44.142)	382
Republican	64 (49.84)	246 (265.022)	46 (41.138)	356
Other	20 (22.68)	115 (120.60)	27 (18.72)	162
Total	126	670	104	900

a. The chi–square statistic is $X^2 = \frac{(42-53.48)^2}{53.48} + \frac{(309-284.378)^2}{284.378} + \ldots + \frac{(27-18.72)^2}{18.72} = 18.711$ with degrees of freedom (3–1)(3–1) = 4. Since $\chi^2_{.05} = 9.48773$, we can reject H_0 and conclude that there is a dependence between part affiliation and opinion about a possible cover up.

b. From Table 6, p–value < .005.

c. Using the Applet, p–value = $P(\chi^2 > 18.711) = .00090$.

d. The p–value is approximate since the distribution of the test statistic is only approximately distributed as chi–square.

14.14 R will be used to answer this problem:

```
> p14.14 <- matrix(c(24,35,5,11,10,8),byrow=T,nrow=2)
> chisq.test(p14.14)

        Pearson's Chi-squared test

data:  p14.14
X-squared = 7.267, df = 2, p-value = 0.02642
```

a. In the above, $X^2 = 7.267$ with a p–value = .02642. Thus with α = .05, we can conclude that there is evidence of a dependence between attachment patterns and hours spent in child care.

b. See part **a** above.

14.15 **a.**

$$X^2 = \sum_{j=1}^{c}\sum_{i=1}^{r}\frac{\left[n_{ij} - E(\hat{n}_{ij})\right]^2}{E(\hat{n}_{ij})} = \sum_{j=1}^{c}\sum_{i=1}^{r}\frac{\left[n_{ij} - \frac{r_i c_j}{n}\right]^2}{\frac{r_i c_j}{n}} = n\sum_{j=1}^{c}\sum_{i=1}^{r}\frac{n_{ij}^2 - \frac{2n_{ij}r_i c_j}{n} + \left(\frac{r_i c_j}{n}\right)^2}{r_i c_j}$$

$$= n\left[\sum_{j=1}^{c}\sum_{i=1}^{r}\frac{n_{ij}^2}{r_i c_j} - 2\sum_{j=1}^{c}\sum_{i=1}^{r}\frac{n_{ij}}{n} + \sum_{j=1}^{c}\sum_{i=1}^{r}\frac{r_i c_j}{n^2}\right]$$

$$= n\left[\sum_{j=1}^{c}\sum_{i=1}^{r}\frac{n_{ij}^2}{r_i c_j} - 2 + \frac{\left(\sum_{i=1}^{r} r_i\right)\left(\sum_{j=1}^{c} c_j\right)}{n^2}\right] = n\left[\sum_{j=1}^{c}\sum_{i=1}^{r}\frac{n_{ij}^2}{r_i c_j} - 2 + \frac{n \cdot n}{n^2}\right]$$

$$= n\left[\sum_{j=1}^{c}\sum_{i=1}^{r}\frac{n_{ij}^2}{r_i c_j} - 1\right].$$

b. When every entry is multiplied by the same constant k, then

$$X^2 = kn\left[\sum_{j=1}^{c}\sum_{i=1}^{r}\frac{(kn_{ij})^2}{kr_i kc_j} - 1\right] = kn\left[\sum_{j=1}^{c}\sum_{i=1}^{r}\frac{n_{ij}^2}{r_i c_j} - 1\right].$$

Thus, X^2 will be increased by a factor of k.

14.16 The contingency table with observed and expected counts is below.

Church attendance	Bush	Democrat	Total
More than …	89	53	142
	(73.636)	(68.364)	
Once / week	87	68	155
	(80.378)	(74.622)	
Once / month	93	85	178
	(92.306)	(85.695)	
Once / year	114	134	248
	(128.604)	(119.400)	
Seldom / never	22	36	58
	(30.077)	(27.923)	
Total	405	376	781

The chi–square statistic is $X^2 = \frac{(89-73.636)^2}{73.636} + \ldots + \frac{(36-27.923)^2}{27.923} = 15.7525$ with $(5-1)(2-1) =$ 4 degrees of freedom. Since $\chi^2_{.05} = 9.48773$, we can conclude that there is evidence of a dependence between frequency of church attendance and choice of presidential candidate.

b. Let p = proportion of individuals who report attending church <u>at least once a week</u>. To estimate this parameter, we use $\hat{p} = \frac{89+53+87+68}{781} = .3803$. A 95% CI for p is

$$.3803 \pm 1.96\sqrt{\tfrac{.3803(.6197)}{781}} = .3803 \pm .0340.$$

14.17 R will be used to solve this problem:

<u>Part a</u>:

```
> p14.17a <- matrix(c(4,0,0,15,12,3,2,7,7,2,3,5),byrow=T,nrow=4)
> chisq.test(p14.17a)

        Pearson's Chi-squared test

data:  p14.17a
X-squared = 19.0434, df = 6, p-value = 0.004091

Warning message:
Chi-squared approximation may be incorrect in: chisq.test(p14.17a)
```

<u>Part b</u>:

```
> p14.17b <- matrix(c(19,6,2,19,41,27,3,7,31,0,3,3),byrow=T,nrow=4)
> chisq.test(p14.17b)
```

```
        Pearson's Chi-squared test

data:  p14.17b
X-squared = 60.139, df = 6, p-value = 4.218e-11

Warning message:
Chi-squared approximation may be incorrect in: chisq.test(p14.17b)
```

a. Using the first output, $X^2 = 19.0434$ with a p–value of .004091. Thus we can conclude at $\alpha = .01$ that the variables are dependent.

b. Using the second output, $X^2 = 60.139$ with a p–value of approximately 0. Thus we can conclude at $\alpha = .01$ that the variables are dependent.

c. Some of the expected cell counts are less than 5, so the chi–square approximation may be invalid (note the warning message in both outputs).

14.18 The contingency table with observed and expected counts is below.

	16–34	35–54	55+	Total
Low violence	8 (13.16)	12 (13.67)	21 (14.17)	41
High violence	18 (12.84)	15 (13.33)	7 (13.83)	40
Total	26	27	28	81

The chi–square statistic is $X^2 = \frac{(8-13.16)^2}{13.16} + \ldots + \frac{(7-13.83)^2}{13.83} = 11.18$ with 2 degrees of freedom. Since $\chi^2_{.05} = 5.99$, we can conclude that there is evidence that the two classifications are dependent.

14.19 The contingency table with the observed and expected counts is below.

	No	Yes	Total
Negative	166 (151.689)	1 (15.311)	167
Positive	260 (274.311)	42 (27.689)	302
Total	426	43	469

a. Here, $X^2 = \frac{(166-151.689)^2}{151.689} + \ldots + \frac{(42-26.689)^2}{26.689} = 22.8705$ with 1 degree of freedom. Since $\chi^2_{.05} = 3.84$, H_0 is rejected and we can conclude that the complications are dependent on the outcome of the initial ECG.

b. From Table 6, p–value < .005.

14.20 We can rearrange the data into a 2×2 contingency table by just considering the type A and B defects:

	B	$\overline{B}$	Total
A	48 (45.54)	18 (20.46)	66
$\overline{A}$	21 (23.46)	13 (10.54)	34
Total	69	31	100

Then, $X^2 = 1.26$ with 1 degree of freedom. Since $\chi^2_{.05} = 3.84$, we fail to reject H_0: there is not enough evidence to prove dependence of the defects.

14.21 Note that all the three examples have $n = 50$. The tests proceed as in previous exercises. For all cases, the critical value is $\chi^2_{.05} = 3.84$

a.

20 (13.44)	4 (10.56)
8 (14.56)	18 (11.44)

$X^2 = 13.99$, reject H_0: species segregate

b.

4 (10.56)	20 (13.44)
18 (11.44)	18 (14.56)

$X^2 = 13.99$, reject H_0: species overly mixed

c.

20 (18.24)	4 (5.76)
18 (19.76)	8 (6.24)

$X^2 = 1.36$, fail to reject H_0

14.22 **a.** The contingency table with the observed and expected counts is:

	Treated	Untreated	Total
Improved	117 (95.5)	74 (95.5)	191
Not Improved	83 (104.5)	126 (104.5)	209
Total	200	200	400

$X^2 = \frac{(117-95.5)^2}{95.5} + \ldots + \frac{(126-104.5)^2}{104.5} = 18.53$ with 1 degree of freedom. Since $\chi^2_{.05} = 3.84$, we reject H_0; there is evidence that the serum is effective.

b. Let p_1 = probability that a treated patient improves and let p_2 = probability that an untreated patient improves. The hypotheses are H_0: $p_1 - p_2 = 0$ vs. H_a: $p_1 - p_2 \neq 0$. Using the procedure from Section 10.3 (derived in Ex. 10.27), we have $\hat{p}_1 = 117/200 = .585$, $\hat{p}_2 = 74/200 = .37$, and the "pooled" estimator $\hat{p} = \frac{117+74}{400} = .4775$, the test statistic is

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}\hat{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{.585 - .37}{\sqrt{.4775(.5225)\left(\frac{2}{200}\right)}} = 4.3.$$

Since the rejection region is $|z| > 1.96$, we soundly reject H_0. Note that $z^2 = X^2$.

c. From Table 6, p–value < .005.

14.23 To test H_0: $p_1 - p_2 = 0$ vs. H_a: $p_1 - p_2 \neq 0$, the test statistic is

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}\hat{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}},$$

from Section 10.3. This is equivalent to

$$Z^2 = \frac{(\hat{p}_1 - \hat{p}_2)^2}{\hat{p}\hat{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} = \frac{n_1 n_2 (\hat{p}_1 - \hat{p}_2)^2}{(n_1 + n_2)\hat{p}\hat{q}}.$$

However, note that

$$\hat{p} = \frac{Y_1 + Y_2}{n_1 + n_2} = \frac{n_1 \hat{p}_1 + n_2 \hat{p}_2}{n_1 + n_2}.$$

Now, consider the X^2 test from Ex. 14.22. The hypotheses were H_0: independence of classification vs. H_a: dependence of classification. If H_0 is true, then $p_1 = p_2$ (serum has no affect). Denote the contingency table as

	Treated	Untreated	Total
Improved	$n_{11} = n_1 \hat{p}_1$	$n_{12} = n_2 \hat{p}_2$	$n_{11} + n_{12}$
Not Improved	$n_{21} = n_1 \hat{q}_1$	$n_{22} = n_2 \hat{q}_2$	$n_{21} + n_{22}$
Total	$n_{11} + n_{21} = n_1$	$n_{12} + n_{22} = n_2$	$\boxed{n_1 + n_2 = n}$

The expected counts are found as follows. $\hat{E}(n_{11}) = \frac{(n_{11}+n_{12})(n_{11}+n_{21})}{n_1+n_2} = \frac{(y_1+y_2)(n_{11}+n_{21})}{n_1+n_2} = n_1 \hat{p}$.

So similarly, $\hat{E}(n_{21}) = n_1 \hat{q}$, $\hat{E}(n_{12}) = n_2 \hat{p}$, and $\hat{E}(n_{22}) = n_2 \hat{q}$. Then, the X^2 statistic can be expressed as

$$X^2 = \frac{n_1^2(\hat{p}_1 - \hat{p})^2}{n_1 \hat{p}} + \frac{n_1^2(\hat{q}_1 - \hat{q})^2}{n_1 \hat{q}} + \frac{n_2^2(\hat{p}_2 - \hat{p})^2}{n_2 \hat{p}} + \frac{n_2^2(\hat{q}_2 - \hat{q})^2}{n_2 \hat{q}}$$

$$= \frac{n_1(\hat{p}_1 - \hat{p})^2}{\hat{p}} + \frac{n_1[(1-\hat{p}_1) - (1-\hat{p})]^2}{\hat{q}} + \frac{n_2(\hat{p}_2 - \hat{p})^2}{\hat{p}} + \frac{n_2[(1-\hat{p}_2) - (1-\hat{p})]^2}{\hat{q}}$$

However, by combining terms, this is equal to $X^2 = \frac{n_1(\hat{p}_1 - \hat{p})^2}{\hat{p}\hat{q}} + \frac{n_2(\hat{p}_2 - \hat{p})^2}{\hat{p}\hat{q}}$. By substituting the expression for $\hat{p}$ above in the numerator, this simplifies to

$$X^2 = \frac{n_1}{\hat{p}\hat{q}}\left(\frac{n_1\hat{p}_1 + n_2\hat{p}_1 - n_1\hat{p}_1 - n_2\hat{p}_2}{n_1 + n_2}\right)^2 + \frac{n_2}{\hat{p}\hat{q}}\left(\frac{n_1\hat{p}_2 + n_2\hat{p}_2 - n_1\hat{p}_1 - n_2\hat{p}_2}{n_1 + n_2}\right)^2$$

$$= \frac{n_1 n_2(\hat{p}_1 - \hat{p}_2)^2}{\hat{p}\hat{q}(n_1 + n_2)} = Z^2$$ from above. Thus, the tests are equivalent.

14.24 **a.** R output follows.

```
> p14.24 <- matrix(c(40,56,68,84,160,144,132,116),byrow=T,nrow=2)
> chisq.test(p14.24)

        Pearson's Chi-squared test

data:  p14.24
X-squared = 24.3104, df = 3, p-value = 2.152e-05     <-- reject H0
```

b. Denote the samples as 1, 2, 3, and 4. Then, the sample proportions that provide parental support for the four groups are $\hat{p}_1 = 40/200 = .20$, $\hat{p}_2 = 56/200 = .28$, $\hat{p}_3 = 68/200 = .34$, $\hat{p}_4 = 84/800 = .42$.

i. A 95% CI for $p_1 - p_4$ is $.20 - .42 \pm 1.96\sqrt{\frac{.20(.80)}{200} + \frac{.42(.58)}{200}} = -.22 \pm .088$.

ii. With 6 confidence intervals, each interval should have confidence coefficient $1 - (.05/6) = .991667$. Thus, we require the critical value $z_{.004167} = 2.638$. The six intervals are:

$p_1 - p_2$:	$-.08 \pm .112$	
$p_1 - p_3$:	$-.14 \pm .116$	(*)
$p_1 - p_4$:	$-.22 \pm .119$	(*)
$p_2 - p_3$:	$-.06 \pm .122$	
$p_2 - p_4$:	$-.14 \pm .124$	(*)
$p_3 - p_4$:	$-.08 \pm .128$	

iii. By considering the intervals that do not contain 0, these are noted by (*).

14.25 **a.** Three populations (income categories) are under investigation. In each population, members are classified as one out of the four education levels, thus creating the multinomial.

b. $X^2 = 19.1723$ with 6 degrees of freedom, and p-value = 0.003882 so reject H_0.

c. The sample proportions are:
- at least an undergraduate degree and marginally rich: 55/100 = .55
- at least an undergraduate degree and super rich: 66/100.66

The 95% CI is

$$.55 - .66 \pm 1.96\sqrt{\frac{.55(.45)}{100} + \frac{.66(.34)}{100}} = -.11 \pm .135.$$

14.26 **a.** Constructing the data using a contingency table, we have

Machine Number	Defectives	Nondefectives
1	16	384
2	24	376
3	9	391

In the chi–square test, $X^2 = 7.19$ with 2 degrees of freedom. Since $\chi^2_{.05} = 5.99$, we can reject the claim that the machines produce the same proportion of defectives.

b. The hypothesis of interest is H_0: $p_1 = p_2 = p_3 = p$ against an alternative that at least one equality is not correct. The likelihood function is

$$L(\boldsymbol{p}) = \prod_{i=1}^{3} \binom{400}{n_i} p_i^{n_i} (1 - p_i)^{400 - n_i}.$$

In Ω, the MLE of p_i is $\hat{p}_i = n_i / 400$, i = 1, 2, 3. In Ω_0, the MLE of p is $\hat{p} = \Sigma n_i / 1200$. Then,

$$\lambda = \frac{\left(\frac{\Sigma n_i}{1200}\right)^{\Sigma n_i}\left(1-\frac{\Sigma n_i}{1200}\right)^{1200-\Sigma y_i}}{\prod_{i=1}^{3}\left(\frac{n_i}{400}\right)^{y_i}\left(1-\frac{n_i}{400}\right)^{400-n_i}}.$$

Using the large sample properties, –2lnλ = –2(–3.689) = 7.378 with 2 degrees of freedom. Again, since $\chi^2_{.05}$ = 5.99, we can reject the claim that the machines produce the same proportion of defectives.

14.27 This exercise is similar to the others. Here, X^2 = 38.429 with 6 degrees of freedom. Since $\chi^2_{.05}$ = 12.59, we can conclude that age and probability of finding nodules are dependent.

14.28 **a.** The chi–square statistic is X^2 = 10.2716 with 1 degree of freedom. Since $\chi^2_{.05}$ = 3.84, we can conclude that the proportions in the two plants are different.

b. The 95% lower bound is

$$.73 - .51 - 1.645\sqrt{\tfrac{.73(.27)}{100} + \tfrac{.51(.49)}{100}} = .22 - .11 = .11.$$

Since the lower bound is greater than 0, this gives evidence that the proportion at the plant with active worker participation is greater.

c. No. The chi–square test in (a) only detects a difference in proportions (equivalent to a two–tailed alternative).

14.29 The contingency table with observed and expected counts is below.

	City A	City B	Nonurban 1	Nonurban 2	Total
w/ lung disease	34	42	21	18	115
	(28.75)	(28.75)	(28.75)	(28.75)	
w/o lung disease	366	358	379	382	1485
	(371.25)	(371.25)	(371.25)	(371.25)	
Total	400	400	400	400	1600

a. Using the above, it is found that X^2 = 14.19 with 3 degrees of freedom and since $\chi^2_{.05}$ = 7.81, we can conclude that there is a difference in the proportions of lung disease for the four locations.

b. It is known that cigarette smoking contributes to lung disease. If more smokers live in urban areas (which is possibly true), this could confound our results. Thus, smokers should probably be excluded from the study.

14.30 The CI is $.085-.105 \pm 1.96\sqrt{\frac{.085(.915)}{400}+\frac{.105(.895)}{400}} = -.02 \pm .041$.

14.31 The contingency table with observed and expected counts is below.

	RI	CO	CA	FL	Total
Participate	46	63	108	121	338
	(63.62)	(78.63)	(97.88)	(97.88)	
Don't participate	149	178	192	179	698
	(131.38)	(162.37)	(202.12)	(202.12)	
Total	195	241	300	300	1036

Here, $X^2 = 21.51$ with 3 degrees of freedom. Since $\chi^2_{.01} = 11.3449$, we can conclude that there is a difference in participation rates for the states.

14.32 See Section 5.9 of the text.

14.33 This is similar to the previous exercises. Here, $X^2 = 6.18$ with 2 degrees of freedom. From Table 6, we find that $.025 < p$–value $< .05$, so there is sufficient evidence that the attitudes are not independent of status.

14.34 R will be used here.

```
> p14.34a <- matrix(c(43,48,9,44,53,3),byrow=T,nrow=2)
> chisq.test(p14.34a)

        Pearson's Chi-squared test

data:  p14.34a
X-squared = 3.259, df = 2, p-value = 0.1960
>

> p14.34b <- matrix(c(4,42,41,13,3,48,35,14),byrow=T,nrow=2)
> chisq.test(p14.34b)

        Pearson's Chi-squared test

data:  p14.34b
X-squared = 1.0536, df = 3, p-value = 0.7883

Warning message:
Chi-squared approximation may be incorrect in: chisq.test(p14.34b)
```

a. For those drivers who rate themselves, the p–value for the test is .1960, so there is not enough evidence to conclude a dependence on gender and driver ratings.

b. For those drivers who rate others, the p–value for the test is .7883, so there is not enough evidence to conclude a dependence on gender and driver ratings.

c. Note in part **b**, the software is warning that two cells have expected counts that are less than 5, so the chi–square approximation may not be valid.

14.35 R:

```
> p14.35 <- matrix(c(49,43,34,31,57,62),byrow=T,nrow=2)
> p14.35
     [,1] [,2] [,3]
[1,]   49   43   34
[2,]   31   57   62
> chisq.test(p14.35)

        Pearson's Chi-squared test

data:  p14.35
X-squared = 12.1818, df = 2, p-value = 0.002263
```

In the above, the test statistic is significant at the .05 significance level, so we can conclude that the susceptibility to colds is affected by the number of relationships that people have.

14.36 R:

```
> p14.36 <- matrix(c(13,14,7,4,12,9,14,3),byrow=T,nrow=2)
> chisq.test(p14.36)

        Pearson's Chi-squared test

data:  p14.36
X-squared = 3.6031, df = 3, p-value = 0.3076

Warning message:
Chi-squared approximation may be incorrect in: chisq.test(p14.36)
```

a. From the above, we fail to reject the hypothesis that position played and knee injury type are independent.
b. From the above, p–value = .3076.
c. From the above, p–value = .3076.

14.37 The hypotheses are H_0: Y is binomial(4, p) vs. H_a: Y isn't binomial(4, p). The probability mass function is

$$p(y) = P(Y = y) = \binom{4}{y} p^y (1-p)^{4-y}, \; y = 0, 1, 2, 3, 4.$$

Similar to Example 14.2, we can estimate p by using the MLE (see Chapter 10; think of this as an experiment with 400 trials):

$$\hat{p} = \frac{\text{number of successes}}{\text{number of trials}} = \frac{0(11)+1(17)+2(42)+3(21)+4(9)}{400} = .5$$

So, the expected counts are $\hat{E}(n_i) = 100\,\hat{p}(i) = \binom{4}{i}(.5)^i(.5)^{4-i} = \binom{4}{i}(.5)^4$, $i = 0, \ldots, 4$. The observed and expected cell counts are below.

	0	1	2	3	4
n_i	11	17	42	21	9
$\hat{E}(n_i)$	6.25	25	37.5	21	6.25

Thus, $X^2 = 8.56$ with 5 – 1 – 1 = 3 degrees of freedom and the critical value is $\chi^2_{.05} = 7.81$. Thus, we can reject H_0 and conclude the data does not follow as binomial.

14.38 **a.** The likelihood function is

$$L(\theta) = (-1)^n \left[\ln(1-\theta)\right]^{-n} \frac{\theta^{\Sigma y_i}}{\Pi y_i}.$$

So, $\ln L(\theta) = k - n\ln[\ln(1-\theta)] + (\ln\theta)\sum_{i=1}^{n} y_i$ where k is a quantity that does not depend on θ. By taking a derivative and setting this expression equal to 0, this yields

$$\left(\frac{1}{1-\theta}\right)\frac{n}{\ln(1-\theta)} + \frac{1}{\theta}\sum_{i=1}^{n} y_i = 0,$$

or equivalently

$$\overline{Y} = \frac{\hat{\theta}}{-(1-\hat{\theta})\ln(1-\hat{\theta})}.$$

b. The hypotheses are H_0: data follow as logarithmic series vs. H_a: not H_0. From the table, $\overline{y} = \frac{1(359)+2(146)+3(57)+\ldots+7(29)}{675} = 2.105$. Thus, to estimate θ, we must solve the nonlinear equation $2.105 = \dfrac{\hat{\theta}}{-(1-\hat{\theta})\ln(1-\hat{\theta})}$, or equivalently we must find the root of

$$2.105(1-\hat{\theta})\ln(1-\hat{\theta}) + \hat{\theta} = 0.$$

By getting some help from R,

```
> uniroot(function(x) x + 2.101*(1-x)*log(1-x),c(.0001,.9999))
$root
[1] 0.7375882
```

Thus, we will use $\hat{\theta}$ = .7376. The probabilities are estimated as

$\hat{p}(1) = -\frac{.7376}{\ln(1-.7376)} = .5513$, $\hat{p}(2) = -\frac{(.7376)^2}{2\ln(1-.7376)} = .2033$, $\hat{p}(3) = .1000$,
$\hat{p}(4) = .0553$, $\hat{p}(5) = .0326$, $\hat{p}(6) = .0201$, $\hat{p}(7, 8, \ldots) = .0374$ (by subtraction)

The expected counts are obtained by multiplying these estimated probabilities by the total sample size of 675. The expected counts are

	1	2	3	4	5	6	7+
$\hat{E}(n_i)$	372.1275	137.2275	67.5000	37.3275	22.005	13.5675	25.245

Here, X^2 = 5.1708 with 7 – 1 – 1 = 5 degrees of freedom. Since $\chi^2_{.05}$ = 11.07, we fail to reject H_0.

14.39 Consider row i as a single cell with r_i observations falling in the cell. Then, $r_1, r_2, \ldots r_r$ follow a multinomial distribution so that the likelihood function is

$$L(\boldsymbol{p}) = \binom{n}{r_1\ r_2\ \cdots\ r_r} p_1^{r_1} p_2^{r_2} \ldots p_r^{r_r}\,.$$

so that

$$\ln L(\boldsymbol{p}) = k + \sum_{j=1}^{r} r_j \ln p_j\,,$$

where k does not involve any parameters and this is subject to $\sum_{j=1}^{r} p_j = 1$. Because of this restriction, we can substitute $p_r = 1 - \sum_{j=1}^{r-1} p_j$ and $r_r = n - \sum_{j=1}^{r-1} r_j$. Thus,

$$\ln L(\boldsymbol{p}) = k + \sum_{j=1}^{r-1} r_j \ln p_j + \left(n - \sum_{j=1}^{r-1} r_j\right)\ln\left(1 - \sum_{j=1}^{r-1} p_j\right).$$

Thus, the $n - 1$ equations to solve are

$$\frac{\partial \ln L}{\partial p_i} = \frac{r_i}{p_i} - \frac{n - \sum_{j=1}^{r-1} r_j}{\left(1 - \sum_{j=1}^{r-1} p_j\right)} = 0\,,$$

or equivalently

$$r_i\left(1 - \sum_{j=1}^{r-1} p_j\right) = p_i\left(n - \sum_{j=1}^{r-1} r_j\right),\ i = 1, 2, \ldots, r-1. \qquad (*)$$

In order to solve these simultaneously, add them together to obtain

$$\sum_{i=1}^{r-1} r_i\left(1 - \sum_{j=1}^{r-1} p_j\right) = \sum_{i=1}^{r-1} p_i\left(n - \sum_{j=1}^{r-1} r_j\right)$$

Thus, $\sum_{i=1}^{r-1} r_i = n\sum_{i=1}^{r-1} p_i$ and so $\sum_{i=1}^{r-1} \hat{p}_i = \frac{1}{n}\sum_{i=1}^{r-1} r_i$. Substituting this into (*) above yields the desired result.

14.40 **a.** The model specifies a trinomial distribution with $p_1 = p^2$, $p_2 = 2p(1 - p)$, $p_3 = (1 - p)^2$. Hence, the likelihood function is

$$L(p) = \frac{n!}{n_1!n_2!n_3!}\, p^{2n_1}[2p(1-p)]^{n_2}(1-p)^{2n_3}\,.$$

The student should verify that the MLE for p is $\hat{p} = \frac{2n_1+n_2}{2n}$. Using the given data, $\hat{p} = .5$ and the (estimated) expected cell counts are $\hat{E}(n_1) = 100(.5)^2 = 25$, $\hat{E}(n_2) = 50$, and $\hat{E}(n_3) = 25$. Using these, we find that $X^2 = 4$ with $3 - 1 - 1 = 1$ degree of freedom. Thus, since $\chi^2_{.05} = 3.84$ we reject H_0: there is evidence that the model is incorrect.

b. If the model specifies $p = .5$, it is not necessary to find the MLE as above. Thus, X^2 will have $3 - 1 = 2$ degrees of freedom. The computed test statistic has the same value as in part **a**, but since $\chi^2_{.05} = 5.99$, H_0 is not rejected in this case.

14.41 The problem describes a multinomial experiment with $k = 4$ cells. Under H_0, the four cell probabilities are $p_1 = p/2$, $p_2 = p^2/2 + pq$, $p_3 = q/2$, and $p_4 = q^2/2$, but $p = 1 - q$. To obtain an estimate of p, the likelihood function is

$$L = C(p/2)^{n_1}(p^2/2 + pq)^{n_2}(q/2)^{n_3}(q^2/2)^{n_4},$$

where C is the multinomial coefficient. By substituting $q = 1 - p$, this simplifies to

$$L = Cp^{n_1+n_2}(2-p)^{n_2}(1-p)^{n_3+2n_4}.$$

By taking logarithms, a first derivative, and setting the expression equal to 0, we obtain

$$(n_1 + 2n_2 + n_3 + 2n_4)p^2 - (3n_1 + 4n_2 + 2n_3 + 4n_4)p + 2(n_1 + n_2) = 0$$

(after some algebra). So, the MLE for p is the root of this quadratic equation. Using the supplied data and the quadratic formula, the valid solution is $\hat{p} = \frac{6960-\sqrt{1,941,760}}{6080} = .9155$.

Now, the estimated cell probabilities and estimated expected cell counts can be found by:

$\hat{p}_i$	$\hat{E}(n_i)$	n_i
$\hat{p}/2 = .45775$	915.50	880
$\hat{p}^2/2 + \hat{p}\hat{q} = .49643$	992.86	1032
$\hat{q}/2 = .04225$	84.50	80
$\hat{q}^2/2 = .00357$	7.14	8

Then, $X^2 = 3.26$ with $4 - 1 - 1 = 2$ degrees of freedom. Since $\chi^2_{.05} = 5.99$, the hypothesized model cannot be rejected.

14.42 Recall that from the description of the problem, it is required that $\sum_{i=1}^k p_i = \sum_{i=1}^k p_i^* = 1$. The likelihood function is given by (multiplication of two multinomial mass functions)

$$L = C\prod_{j=1}^k p_j^{n_j}\left(p_j^*\right)^{m_j},$$

where C are the multinomial coefficients. Now under H_0, this simplifies to

$$L_0 = C\prod_{j=1}^k p_j^{n_j+m_j}.$$

This is a special case of Ex. 14.39, so the MLEs are $\hat{p}_i = \frac{n_i+m_i}{n+m}$ and the estimated expected counts are $\hat{E}(n_i) = n\hat{p}_i = n\left(\frac{n_i+m_i}{n+m}\right)$ and $\hat{E}(m_i) = m\hat{p}_i = m\left(\frac{n_i+m_i}{n+m}\right)$ for $i = 1, \ldots, k$. The chi–square test statistic is given by

$$X^2 = \sum_{j=1}^k \frac{\left[n_i - n\left(\frac{n_i+m_i}{n+m}\right)\right]^2}{n\left(\frac{n_i+m_i}{n+m}\right)} + \sum_{j=1}^k \frac{\left[m_i - m\left(\frac{n_i+m_i}{n+m}\right)\right]^2}{m\left(\frac{n_i+m_i}{n+m}\right)}$$

which has a chi–square distribution with $2k - 2 - (k - 1) = k - 1$ degrees of freedom. Two degrees of freedom are lost due the two conditions first mentioned in the solution of this problem, and $k - 1$ degrees of freedom are lost in the estimation of cell probabilities. Hence, a rejection region will be based on $k - 1$ degrees of freedom in the chi–square distribution.

14.43 In this exercise there are 4 binomial experiments, one at each of the four dosage levels. So, with $i = 1, 2, 3, 4$, and p_i represents the binomial (success probability) parameter for dosage i, we have that $p_i = 1 + \beta i$. Thus, in order to estimate β, we form the likelihood function (product of four binomial mass functions):

$$L(\beta) = \prod_{i=1}^{4} \binom{1000}{n_i} (1 + i\beta)^{n_i} (-i\beta)^{1000-n_i} = K \prod_{i=1}^{4} \binom{1000}{n_i} (1 + i\beta)^{n_i} \beta^{1000-n_i},$$

where K is a constant that does not involve β. Then,

$$\frac{dL(\beta)}{d\beta} = \sum_{i=1}^{4} \frac{in_i}{1+i\beta} + \frac{1}{\beta} \sum_{i=1}^{4} (1000 - n_i).$$

By equating this to 0, we obtain a nonlinear function of β that must be solved numerically (to find the root). Below is the R code that does the job; note that in the association of β with probability and the dose levels, β must be contained in (–.25, 0):

```
> mle <- function(x)
+ {
+ ni <- c(820,650,310,50)
+ i <- 1:4
+ temp <- sum(1000-ni)
+ return(sum(i*ni/(1+i*x))+temp/x)
+ }
>
> uniroot(mle, c(-.2499,-.0001))  <– guessed range for the parameter
$root
[1] -0.2320990
```

Thus, we take $\hat{\beta} = -.232$ and so:

$\hat{p}_1 = 1 - .232 = .768$
$\hat{p}_2 = 1 + 2(-.232) = .536,$
$\hat{p}_3 = 1 + 3(-.232) = .304$
$\hat{p}_4 = 1 + 4(-.232) = .072.$

The observed and (estimated) expected cell counts are

Dosage	1	2	3	4
Survived	820	650	320	50
	(768)	(536)	(304)	(72)
Died	180	350	690	950
	(232)	(464)	(696)	(928)

The chi–square test statistic is $X^2 = 74.8$ with 8 – 4 – 1 = 3 degrees of freedom (see note below). Since $\chi^2_{.05} = 7.81$, we can soundly reject the claim that $p = 1 + \beta D$.

Note: there are 8 cells, but 5 restrictions:
- $p_i + q_i = 1$ for i = 1, 2, 3, 4
- estimation of β.

Chapter 15: Nonparametric Statistics

15.1 Let Y have a binomial distribution with $n = 25$ and $p = .5$. For the two–tailed sign test, the test rejects for extreme values (either too large or too small) of the test statistic whose null distribution is the same as Y. So, Table 1 in Appendix III can be used to define rejection regions that correspond to various significant levels. Thus:

Rejection region	α
$Y \le 6$ or $Y \ge 19$	$P(Y \le 6) + P(Y \ge 19) = .014$
$Y \le 7$ or $Y \ge 18$	$P(Y \le 7) + P(Y \ge 18) = .044$
$Y \le 8$ or $Y \ge 17$	$P(Y \le 8) + P(Y \ge 17) = .108$

15.2 Let p = P(blood levels are elevated after training). We will test H_0: $p = .5$ vs H_a: $p > .5$.

a. Since $m = 15$, so p–value = $P(M \ge 15) = \binom{17}{15}.5^{17} + \binom{17}{16}.5^{17} + \binom{17}{16}.5^{17} = 0.0012$.

b. Reject H_0.

c. $P(M \ge 15) = P(M > 14.5) \approx P(Z > 2.91) = .0018$, which is very close to part **a**.

15.3 Let $p = P$(recovery rate for A exceeds B). We will test H_0: $p = .5$ vs H_a: $p \ne .5$. The data are:

Hospital	A	B	Sign(A – B)
1	75.0	85.4	–
2	69.8	83.1	–
3	85.7	80.2	+
4	74.0	74.5	–
5	69.0	70.0	–
6	83.3	81.5	+
7	68.9	75.4	–
8	77.8	79.2	–
9	72.2	85.4	–
10	77.4	80.4	–

a. From the above, $m = 2$ so the p–value is given by $2P(M \le 2) = .110$. Thus, in order to reject H_0, it would have been necessary that the significance level $\alpha \ge .110$. Since this is fairly large, H_0 would probably not be rejected.

b. The t–test has a normality assumption that may not be appropriate for these data. Also, since the sample size is relatively small, a large–sample test couldn't be used either.

15.4 **a.** Let $p = P$(school A exceeds school B in test score). For H_0: $p = .5$ vs H_a: $p \ne .5$, the test statistic is M = # of times school A exceeds school B in test score. From the table, we find $m = 7$. So, the p–value = $2P(M \ge 7) = 2P(M \le 3) = 2(.172) = .344$. With $\alpha = .05$, we fail to reject H_0.

b. For the one–tailed test, H_0: $p = .5$ vs H_a: $p > .5$. Here, the p–value = $P(M \ge 7) = .173$ so we would still fail to reject H_0.

15.5 Let $p = P$(judge favors mixture B). For H_0: $p = .5$ vs H_a: $p \neq .5$, the test statistic is M = # of judges favoring mixture B. Since the observed value is $m = 2$, p–value = $2P(M \leq 2)$ = 2(.055) = .11. Thus, H_0 is not rejected at the $\alpha = .05$ level.

15.6 **a.** Let $p = P$(high elevation exceeds low elevation). For H_0: $p = .5$ vs H_a: $p > .5$, the test statistic is M = # of nights where high elevation exceeds low elevation. Since the observed value is $m = 9$, p–value = $P(M \geq 9) = .011$. Thus, the data favors H_a.

b. Extreme temperatures, such as the minimum temperatures in this example, often have skewed distributions, making the assumptions of the t–test invalid.

15.7 **a.** Let $p = P$(response for stimulus 1 is greater that for stimulus 2). The hypotheses are H_0: $p = .5$ vs H_a: $p > .5$, and the test statistic is M = # of times response for stimulus 1 exceeds stimulus 2. If it is required that $\alpha \leq .05$, note that

$$P(M \leq 1) + P(M \geq 8) = .04,$$

where M is binomial($n = 9$, $p = .5$) under H_0. Our rejection region is the set {0, 1, 8, 9}. From the table, $m = 2$ so we fail to reject H_0.

b. The proper test is the paired t–test. So, with H_0: $\mu_1 - \mu_2 = 0$ vs. H_a: $\mu_1 - \mu_2 \neq 0$, the summary statistics are $\overline{d} = -1.022$ and $s_D^2 = 3.467$, the computed test statistic is $|t| = \frac{|-1.022|}{\sqrt{\frac{3.467}{9}}} = 1.65$ with 8 degrees of freedom. Since $t_{.025} = 2.306$, we fail to reject H_0.

15.8 Let $p = P(B$ exceeds $A)$. For H_0: $p = .5$ vs H_a: $p \neq .5$, the test statistic is M = # of technicians for which B exceeds A with $n = 7$ (since one tied pair is deleted). The observed value of M is 1, so the p–value = $2P(M \leq 1) = .125$, so H_0 is not rejected.

15.9 **a.** Since two pairs are tied, $n = 10$. Let $p = P$(before exceeds after) so that H_0: $p = .5$ vs H_a: $p > .5$. From the table, $m = 9$ so the p–value is $P(M \geq 9) = .011$. Thus, H_0 is not rejected with $\alpha = .01$.

b. Since the observations are counts (and thus integers), the paired t–test would be inappropriate due to its normal assumption.

15.10 There are n ranks to be assigned. Thus, $T^+ + T^-$ = sum of all ranks = $\sum_{i=1}^{n} i = n(n+1)/2$ (see Appendix I).

15.11 From Ex. 15.10, $T^- = n(n+1)/2 - T^+$. If $T^+ > n(n+1)/4$, it must be so that $T^- < n(n+1)/4$. Therefore, since $T = \min(T^+, T^-)$, $T = T^-$.

15.12 **a.** Define d_i to be the difference between the math score and the art score for the i^{th} student, $i = 1, 2, \ldots, 15$. Then, $T^+ = 14$ and $T^- = 106$. So, $T = 14$ and from Table 9, since $14 < 16$, p–value < .01. Thus H_0 is rejected.
b. H_0: identical population distributions for math and art scores vs. H_a: population distributions differ by location.

15.13 Define d_i to be the difference between school A and school B. The differences, along with the ranks of $|d_i|$ are given below.

	1	2	3	4	5	6	7	8	9	10
d_i	28	5	–4	15	12	–2	7	9	–3	13
rank $\lvert d_i \rvert$	13	4	3	9	7	1	5	6	2	8

Then, $T^+ = 49$ and $T^- = 6$ so $T = 6$. Indexing $n = 10$ in Table 9, $.02 < T < .05$ so H_0 would be rejected if $\alpha = .05$. This is a different decision from Ex. 15.4

15.14 Using the data from Ex. 15.6, $T^- = 1$ and $T^+ = 54$, so $T = 1$. From Table 9, p–value < .005 for this one–tailed test and thus H_0 is rejected.

15.15 Here, R is used:

```
> x <- c(126,117,115,118,118,128,125,120)
> y <- c(130,118,125,120,121,125,130,120)
> wilcox.test(x,y,paired=T,alt="less",correct=F)

        Wilcoxon signed rank test

data:  x and y
V = 3.5, p-value = 0.0377
alternative hypothesis: true mu is less than 0
```

The test statistic is $T = 3.5$ so H_0 is rejected with $\alpha = .05$.

15.16 **a.** The sign test statistic is $m = 8$. Thus, p–value = $2P(M \geq 8) = .226$ (computed using a binomial with $n = 11$ and $p = .5$). H_0 should not be rejected.

b. For the Wilcoxon signed–rank test, $T^+ = 51.5$ and $T^- = 14.5$ with $n = 11$. With $\alpha = .05$, the rejection region is $\{T \leq 11\}$ so H_0 is not rejected.

15.17 From the sample, $T^+ = 44$ and $T^- = 11$ with $n = 10$ (two ties). With $T = 11$, we reject H_0 with $\alpha = .05$ using Table 9.

15.18 Using the data from Ex. 12.16:

d_i	3	6.1	2	4	2.5	8.9	.8	4.2	9.8	3.3	2.3	3.7	2.5	–1.8	7.5
$\lvert d_i \rvert$	3	6.1	2	4	2.5	8.9	.8	4.2	9.8	3.3	2.3	3.7	2.5	1.8	7.5
rank	7	12	3	10	5.5	14	1	11	15	8	4	9	5.5	2	13

Thus, $T^+ = 118$ and $T^- = 2$ with $n = 15$. From Table 9, since $T^- < 16$, p–value < .005 (a one–tailed test) so H_0 is rejected.

15.19 Recall for a continuous random variable Y, the median ξ is a value such that $P(Y > \xi) = P(Y < \xi) = .5$. It is desired to test H_0: $\xi = \xi_0$ vs. H_a: $\xi \neq \xi_0$.

a. Define $D_i = Y_i - \xi_0$ and let M = # of negative differences. Very large or very small values of M (compared against a binomial distribution with $p = .5$) lead to a rejection.
b. As in part a, define $D_i = Y_i - \xi_0$ and rank the D_i according to their absolute values according to the Wilcoxon signed–rank test.

15.20 Using the results in Ex. 15.19, we have H_0: $\xi = 15{,}000$ vs. H_a: $\xi > 15{,}000$ The differences $d_i = y_i - 15000$ are:

d_i	–200	1900	3000	4100	–1800	3500	5000	4200	100	1500
$\lvert d_i \rvert$	200	1900	3000	4100	1800	3500	5000	4200	100	1500
rank	2	5	6	8	4	7	10	9	1	3

a. With the sign test, $m = 2$, p–value = $P(M \le 2) = .055$ ($n = 10$) so H_0 is rejected.
b. $T^+ = 49$ and $T^- = 6$ so $T = 6$. From Table 9, $.01 < p$–value $< .025$ so H_0 is rejected.

15.21 **a.** $U = 4(7) + \frac{1}{2}(4)(5) - 34 = 4$. Thus, the p–value $= P(U \le 4) = .0364$
b. $U = 5(9) + \frac{1}{2}(5)(6) - 25 = 35$. Thus, the p–value $= P(U \ge 35) = P(U \le 10) = .0559$.
c. $U = 3(6) + \frac{1}{2}(3)(4) - 23 = 1$. Thus, p–value $= 2P(U \le 1) = 2(.0238) = .0476$

15.22 To test: H_0: the distributions of ampakine CX–516 are equal for the two groups
H_a: the distributions of ampakine CX–516 differ by a shift in location

The samples of ranks are:

Age group											
20s	20	11	7.5	14	7.5	16.5	2	18.5	3.5	7.5	$W_A = 108$
65–70	1	16.5	7.5	14	11	14	5	11	18.5	3.5	$W_B = 102$

Thus, $U = 100 + 10(11)/2 - 108 = 47$. By Table 8,
$$p\text{–value} = 2P(U \le 47) > 2P(U \le 39) = 2(.2179) = .4358.$$
Thus, there is not enough evidence to conclude that the population distributions of ampakine CX–516 are different for the two age groups.

15.23 The hypotheses to be tested are:
H_0: the population distributions for plastics 1 and 2 are equal
H_a: the populations distributions differ by location

The data (with ranks in parentheses) are:

Plastic 1	15.3 (2)	18.7 (6)	22.3 (10)	17.6 (4)	19.1 (7)	14.8 (1)
Plastic 2	21.2 (9)	22.4 (11)	18.3 (5)	19.3 (8)	17.1 (3)	27.7 (12)

By Table 8 with $n_1 = n_2 = 6$, $P(U \le 7) = .0465$ so $\alpha = 2(.0465) = .093$. The two possible values for U are $U_A = 36 + \frac{6(7)}{2} - W_A = 27$ and $U_B = 36 + \frac{6(7)}{2} - W_B = 9$. So, $U = 9$ and thus H_0 is not rejected.

15.24 **a.** Here, $U_A = 81 + \frac{9(10)}{2} - W_A = 126 - 94 = 32$ and $U_B = 81 + \frac{9(10)}{2} - W_B = 126 - 77 = 49$. Thus, $U = 32$ and by Table 8, p–value = $2P(U \le 32) = 2(.2447) = .4894$.

b. By conducting the two sample t–test, we have H_0: $\mu_1 - \mu_2 = 0$ vs. H_a: $\mu_1 = \mu_2 \ne 0$. The summary statistics are $\bar{y}_1 = 8.267$, $\bar{y}_2 = 8.133$, and $s_p^2 = .8675$. The computed test stat. is $|t| = \frac{|.1334|}{\sqrt{.8675\left(\frac{2}{9}\right)}} = .30$ with 16 degrees of freedom. By Table 5, p–value > 2(.1) = .20 so H_0 is not rejected.

c. In part **a**, we are testing for a shift in distribution. In part **b**, we are testing for unequal means. However, since in the t–test it is assumed that both samples were drawn from normal populations with common variance, under H_0 the two distributions are also equal.

15.25 With $n_1 = n_2 = 15$, it is found that $W_A = 276$ and $W_B = 189$. Note that although the actual failure times are not given, they are not necessary:

$$W_A = [1 + 5 + 7 + 8 + 13 + 15 + 20 + 21 + 23 + 24 + 25 + 27 + 28 + 29 + 30] = 276.$$

Thus, $U = 354 - 276 = 69$ and since $E(U) = \frac{n_1 n_2}{2} = 112.5$ and $V(U) = 581.25$,

$$z = \frac{69 - 112.5}{\sqrt{581.25}} = -1.80.$$

Since $-1.80 < -z_{.05} = -1.645$, we can conclude that the experimental batteries have a longer life.

15.26 R:

```
> DDT <- c(16,5,21,19,10,5,8,2,7,2,4,9)
> Diaz <- c(7.8,1.6,1.3)
> wilcox.test(Diaz,DDT,correct=F)

        Wilcoxon rank sum test

data:  Diaz and DDT
W = 6, p-value = 0.08271
alternative hypothesis: true mu is not equal to 0
```

With $\alpha = .10$, we can reject H_0 and conclude a difference between the populations.

15.27 Calculate $U_A = 4(6) + \frac{4(5)}{2} - W_A = 34 - 34 = 0$ and $U_B = 4(6) + \frac{6(7)}{2} - W_B = 45 - 21 = 24$. Thus, we use $U = 0$ and from Table 8, p–value = $2P(U \le 0) = 2(.0048) = .0096$. So, we would reject H_0 for $\alpha \approx .10$.

15.28 Similar to previous exercises. With $n_1 = n_2 = 12$, the two possible values for U are

$$U_A = 144 + \frac{12(13)}{2} - 89.5 = 132.5 \text{ and } U_B = 144 + \frac{12(13)}{2} - 210.5 = 11.5,$$

but since it is required to detect a shift of the "B" observations to the right of the "A" observations, we let $U = U_A = 132.5$. Here, we can use the large–sample approximation. The test statistic is $z = \frac{132.5 - 72}{\sqrt{300}} = 3.49$, and since $3.49 > z_{.05} = 1.645$, we can reject H_0 and conclude that rats in population "B" tend to survive longer than population A.

15.29 H_0: the 4 distributions of mean leaf length are identical, vs. H_a: at least two are different.
R:
```
> len <-
c(5.7,6.3,6.1,6.0,5.8,6.2,6.2,5.3,5.7,6.0,5.2,5.5,5.4,5.0,6,5.6,4,5.2,
3.7,3.2,3.9,4,3.5,3.6)
> site <- factor(c(rep(1,6),rep(2,6),rep(3,6),rep(4,6)))
> kruskal.test(len~site)

        Kruskal-Wallis rank sum test

data:  len by site
Kruskal-Wallis chi-squared = 16.974, df = 3, p-value = 0.0007155
```

We reject H_0 and conclude that there is a difference in at least two of the four sites.

15.30 **a.** This is a completely randomized design.

b. R:
```
> prop<-c(.33,.29,.21,.32,.23,.28,.41,.34,.39,.27,.21,.30,.26,.33,.31)
> campaign <- factor(c(rep(1,5),rep(2,5),rep(3,5)))
> kruskal.test(prop,campaign)

        Kruskal-Wallis rank sum test

data:  prop and campaign
Kruskal-Wallis chi-squared = 2.5491, df = 2, p-value = 0.2796
```

From the above, we cannot reject H_0.

c. R:
```
> wilcox.test(prop[6:10],prop[11:15], alt="greater")

        Wilcoxon rank sum test

data:  prop[6:10] and prop[11:15]
W = 19, p-value = 0.1111
alternative hypothesis: true mu is greater than 0
```

From the above, we fail to reject H_0: we cannot conclude that campaign 2 is more successful than campaign 3.

15.31 **a.** The summary statistics are: TSS = 14,288.933, SST = 2586.1333, SSE = 11,702.8. To test H_0: $\mu_A = \mu_B = \mu_C$, the test statistic is $F = \frac{2586.1333/2}{11,702.8/12} = 1.33$ with 2 numerator and 12 denominator degrees of freedom. Since $F_{.05} = 3.89$, we fail to reject H_0. We assumed that the three random samples were independently drawn from separate normal populations with common variance. Life–length data is typically right skewed.

b. To test H_0: the population distributions are identical for the three brands, the test statistic is $H = \frac{122}{15(16)}\left(\frac{36^2}{5} + \frac{35^2}{5} + \frac{49^2}{5}\right) - 3(16) = 1.22$ with 2 degrees of freedom. Since $\chi^2_{.05}$ = 5.99, we fail to reject H_0.

15.32 **a.** Using R:

```
> time<–c(20,6.5,21,16.5,12,18.5,9,14.5,16.5,4.5,2.5,14.5,12,18.5,9,
1,9,4.5, 6.5,2.5,12)
> strain<-factor(c(rep("Victoria",7),rep("Texas",7),rep("Russian",7)))
>
> kruskal.test(time~strain)

        Kruskal-Wallis rank sum test

data:  time by strain
Kruskal-Wallis chi-squared = 6.7197, df = 2, p-value = 0.03474
```

By the above, p–value = .03474 so there is evidence that the distributions of recovery times are not equal.

b. R: comparing the Victoria A and Russian strains:

```
> wilcox.test(time[1:7],time[15:21],correct=F)

        Wilcoxon rank sum test

data:  time[1:7] and time[15:21]
W = 43, p-value = 0.01733
alternative hypothesis: true mu is not equal to 0
```

With p–value = .01733, there is sufficient evidence that the distribution of recovery times with the two strains are different.

15.33 R:

```
> weight <- c(22,24,16,18,19,15,21,26,16,25,17,14,28,21,19,24,23,17,
18,13,20,21)
> temp <- factor(c(rep(38,5),rep(42,6),rep(46,6),rep(50,5)))
>
> kruskal.test(weight~temp)

        Kruskal-Wallis rank sum test

data:  weight by temp
Kruskal-Wallis chi-squared = 2.0404, df = 3, p-value = 0.5641
```

With a p–value = .5641, we fail to reject the hypothesis that the distributions of weights are equal for the four temperatures.

15.34 The rank sums are: $R_A = 141$, $R_B = 248$, and $R_C = 76$. To test H_0: the distributions of percentages of plants with weevil damage are identical for the three chemicals, the test statistic is $H = \frac{12}{30(31)}\left(\frac{141^2}{10} + \frac{248^2}{10} + \frac{76^2}{10}\right) - 3(31) = 19.47$. Since $\chi^2_{.005} = 10.5966$, the p–value is less than .005 and thus we conclude that the population distributions are not equal.

15.35 By expanding H,

$$H = \frac{12}{n(n+1)}\sum_{i=1}^{k} n_i\left(\overline{R}_i^2 - 2\overline{R}_i\frac{n+1}{2} + \frac{(n+1)^2}{4}\right)$$

$$= \frac{12}{n(n+1)}\sum_{i=1}^{k} n_i\left(\frac{R_i^2}{n_i^2} - (n+1)\frac{R_i}{n_i} + \frac{(n+1)^2}{4}\right)$$

$$= \frac{12}{n(n+1)}\sum_{i=1}^{k}\frac{R_i^2}{n_i} + \frac{12}{n}\sum_{i=1}^{k} R_i + \frac{3(n+1)}{n}\sum_{i=1}^{k} n_i$$

$$= \frac{12}{n(n+1)}\sum_{i=1}^{k}\frac{R_i^2}{n_i} + \frac{12}{n}\left(\frac{n(n+1)}{2}\right) + \frac{3(n+1)}{n}\cdot n$$

$$= \frac{12}{n(n+1)}\sum_{i=1}^{k}\frac{R_i^2}{n_i} - 3(n+1)\,.$$

15.36 There are 15 possible pairings of ranks: The statistic H is

$$H = \frac{12}{6(7)}\sum R_i^2\,/\,2 - 3(7) = \frac{1}{7}\left(\sum R_i^2 - 147\right).$$

The possible pairings are below, along with the value of H for each.

pairings			H
(1, 2)	(3, 4)	(5, 6)	32/7
(1, 2)	(3, 5)	(4, 6)	26/7
(1, 2)	(3, 6)	(5, 6)	24/7
(1, 3)	(2, 4)	(5, 6)	26/7
(1, 3)	(2, 5)	(4, 6)	18/7
(1, 3)	(2, 6)	(4, 5)	14/7
(1, 4)	(2, 3)	(5, 6)	24/7
(1, 4)	(2, 5)	(3, 6)	8/7
(1, 4)	(2, 6)	(3, 5)	6/7
(1, 5)	(2, 3)	(4, 6)	14/7
(1, 5)	(2, 4)	(3, 6)	6/7
(1, 5)	(2, 6)	(3, 4)	2/7
(1, 6)	(2, 3)	(4, 5)	8/7
(1, 6)	(2, 4)	(3, 5)	2/7
(1, 6)	(2, 5)	(3, 4)	0

Thus, the null distribution of H is (each of the above values are equally likely):

h	0	2/7	6/7	8/7	2	18/7	24/7	26/7	32/7
$p(h)$	1/15	2/15	2/15	2/15	2/15	1/15	2/15	2/15	1/15

15.37 R:

```
> score <- c(4.8,8.1,5.0,7.9,3.9,2.2,9.2,2.6,9.4,7.4,6.8,6.6,3.6,5.3,
2.1,6.2,9.6,6.5,8.5,2.0)
> anti <- factor(c(rep("I",5),rep("II",5),rep("III",5),rep("IV",5)))
> child <- factor(c(1:5, 1:5, 1:5, 1:5))
> friedman.test(score ~ anti | child)

        Friedman rank sum test

data:  score and anti and child
Friedman chi-squared = 1.56, df = 3, p-value = 0.6685
```

a. From the above, we do not have sufficient evidence to conclude the existence of a difference in the tastes of the antibiotics.

b. Fail to reject H_0.

c. Two reasons: more children would be required and the potential for significant child to child variability in the responses regarding the tastes.

15.38 R:

```
> cadmium <- c(162.1,199.8,220,194.4,204.3,218.9,153.7,199.6,210.7,
179,203.7,236.1,200.4,278.2,294.8,341.1,330.2,344.2)
> harvest <- c(rep(1,6),rep(2,6),rep(3,6))
> rate <- c(1:6,1:6,1:6)
> friedman.test(cadmium ~ rate | harvest)

        Friedman rank sum test

data:  cadmium and rate and harvest
Friedman chi-squared = 11.5714, df = 5, p-value = 0.04116
```

With $\alpha = .01$ we fail to reject H_0: we cannot conclude that the cadmium concentrations are different for the six rates of sludge application.

15.39 R:

```
> corrosion <- c(4.6,7.2,3.4,6.2,8.4,5.6,3.7,6.1,4.9,5.2,4.2,6.4,3.5,
5.3,6.8,4.8,3.7,6.2,4.1,5.0,4.9,7.0,3.4,5.9,7.8,5.7,4.1,6.4,4.2,5.1)
> sealant <- factor(c(rep("I",10),rep("II",10),rep("III",10)))
> ingot <- factor(c(1:10,1:10,1:10))
> friedman.test(corrosion~sealant|ingot)

        Friedman rank sum test

data:  corrosion and sealant and ingot
Friedman chi-squared = 6.6842, df = 2, p-value = 0.03536
```

With $\alpha = .05$, we can conclude that there is a difference in the abilities of the sealers to prevent corrosion.

15.40 A summary of the ranked data is

Ear	A	B	C
1	2	3	1
2	2	3	1
3	1	3	2
4	3	2	1
5	2	1	3
6	1	3	2
7	2.5	2.5	1
8	2	3	1
9	2	3	1
10	2	3	1

Thus, $R_A = 19.5$, $R_B = 26.5$, and $R_C = 14$.
To test: H_0: distributions of aflatoxin levels are equal
H_a: at least two distributions differ in location

$F_r = \frac{12}{10(3)(4)}[(19.5)^2 + (26.5)^2 + (14)^2] - 3(10)(4) = 7.85$ with 2 degrees of freedom. From Table 6, .01 < p–value < .025 so we can reject H_0.

15.41 **a.** To carry out the Friedman test, we need the rank sums, R_i, for each model. These can be found by adding the ranks given for each model. For model A, $R_1 = 8(15) = 120$. For model B, $R_2 = 4 + 2(6) + 7 + 8 + 9 + 2(14) = 68$, etc. The R_i values are:

120, 68, 37, 61, 31, 87, 100, 34, 32, 62, 85, 75, 30, 71, 67

Thus, $\sum R_i^2 = 71{,}948$ and then $F_r = \frac{12}{8(15)(16)}[71{,}948 - 3(8)(16)] = 65.675$ with 14 degrees of freedom. From Table 6, we find that p–value < .005 so we soundly reject the hypothesis that the 15 distributions are equal.

b. The highest (best) rank given to model H is lower than the lowest (worst) rank given to model M. Thus, the value of the test statistic is $m = 0$. Thus, using a binomial distribution with $n = 8$ and $p = .5$, p–value = $2P(M = 0) = 1/128$.

c. For the sign test, we must know whether each judge (exclusively) preferred model H or model M. This is not given in the problem.

15.42 H_0: the probability distributions of skin irritation scores are the same for the 3 chemicals vs. H_a: at least two of the distributions differ in location.
From the table of ranks, $R_1 = 15$, $R_2 = 19$, and $R_3 = 14$. The test statistic is

$$F_r = \frac{12}{8(3)(4)}[(15)^2 + (19)^2 + (14)^2] - 3(8)(4) = 1.75$$

with 2 degrees of freedom. Since $\chi^2_{.01} = 9.21034$, we fail to reject H_0: there is not enough evidence to conclude that the chemicals cause different degrees of irritation.

15.43 If $k = 2$ and $b = n$, then $F_r = \frac{2}{n}\left(R_1^2 + R_2^2\right) - 9n$. For $R_1 = 2n - M$ and $R_2 = n + M$, then

$$\begin{aligned}
F_r &= \frac{2}{n}\left[(2n - M)^2 + (n + M)^2\right] - 9n \\
&= \frac{2}{n}\left[(4n^2 - 4nM + M^2) + (n^2 + 2nM + M^2) - 4.5n^2\right] \\
&= \frac{2}{n}(-.5n^2 - 2nM + 2M^2) \\
&= \frac{4}{n}(M^2 - nM - \tfrac{1}{4}n^2) \\
&= \frac{4}{n}(M - \tfrac{1}{2}n)^2
\end{aligned}$$

The Z statistic from Section 15.3 is $Z = \dfrac{M - \frac{1}{2}n}{\frac{1}{2}\sqrt{n}} = \dfrac{2}{\sqrt{n}}(M - \tfrac{1}{2}n)$. So, $Z^2 = F_r$.

15.44 Using the hints given in the problem,

$$F_r = \tfrac{12b}{k(k+1)}\sum\left(\overline{R}_i^2 - 2\overline{R}_i\overline{R} + \overline{R}^2\right) = \tfrac{12b}{k(k+1)}\sum\left(R_i^2/b^2 - (k+1)R_i/b + (k+1)^2/4\right)$$

$$= \tfrac{12b}{k(k+1)}\sum R_i^2/b^2 - \tfrac{12}{k}\tfrac{bk(k+1)}{2} + \tfrac{12b(k+1)k}{4k} = \tfrac{12}{bk(k+1)}\sum R_i^2 - 3b(k+1).$$

15.45 This is similar to Ex. 15.36. We need only work about the 3! = 6 possible rank pairing. They are listed below, with the R_i values and F_r. When $b = 2$ and $k = 3$, $F_r = \frac{1}{2}\Sigma R_i^2 - 24$.

Block 1	Block 2	R_i
1	1	2
2	2	4
3	3	6

$F_r = 4$

Block 1	Block 2	R_i
1	1	2
2	3	5
3	2	5

$F_r = 3$

Block 1	Block 2	R_i
1	2	3
2	1	3
3	3	6

$F_r = 3$

Block 1	Block 2	R_i
1	2	3
2	3	5
3	1	4

$F_r = 1$

Block 1	Block 2	R_i
1	3	4
2	1	3
3	2	5

$F_r = 1$

Block 1	Block 2	R_i
1	3	4
2	2	4
3	1	4

$F_r = 0$

Thus, with each value being equally likely, the null distribution is given by

$$P(F_r = 0) = P(F_r = 4) = 1/6 \text{ and } P(F_r = 1) = P(F_r = 3) = 1/3.$$

15.46 Using Table 10, indexing row (5, 5):

a. $P(R = 2) = P(R \le 2) = .008$ (minimum value is 2).

b. $P(R \le 3) = .040$.

c. $P(R \le 4) = .167$.

15.47 Here, $n_1 = 5$ (blacks hired), $n_2 = 8$ (whites hired), and $R = 6$. From Table 10,

$$p\text{–value} = 2P(R \le 6) = 2(.347) = .694.$$

So, there is no evidence of nonrandom racial selection.

15.48 The hypotheses are H_0: no contagion (randomly diseased)
H_a: contagion (not randomly diseased)

Since contagion would be indicated by a grouping of diseased trees, a small numer of runs tends to support the alternative hypothesis. The computed test statistic is $R = 5$, so with $n_1 = n_2 = 5$, p–value = .357 from Table 10. Thus, we cannot conclude there is evidence of contagion.

15.49 **a.** To find $P(R \le 11)$ with $n_1 = 11$ and $n_2 = 23$, we can rely on the normal approximation. Since $E(R) = \frac{2(11)(23)}{11+23} + 1 = 15.88$ and $V(R) = 6.2607$, we have (in the second step the continuity correction is applied)

$$P(R \le 11) = P(R < 11.5) \approx P(Z < \tfrac{11.5-15.88}{\sqrt{6.2607}}) = P(Z < -1.75) = .0401.$$

b. From the sequence, the observed value of $R = 11$. Since an unusually large or small number of runs would imply a non–randomness of defectives, we employ a two–tailed test. Thus, since the p–value $= 2P(R \le 11) \approx 2(.0401) = .0802$, significance evidence for non–randomness does not exist here.

15.50 **a.** The measurements are classified as A if they lie above the mean and B if they fall below. The sequence of runs is given by

$$\underline{A\ A\ A\ A\ A}\ \underline{B\ B\ B\ B\ B\ B}\ \underline{A}\ \underline{B}\ \underline{A}\ \underline{B}\ \underline{A}$$

Thus, $R = 7$ with $n_1 = n_2 = 8$. Now, non–random fluctuation would be implied by a small number of runs, so by Table 10, p–value $= P(R \le 7) = .217$ so non–random fluctuation cannot be concluded.

b. By dividing the data into equal parts, $\bar{y}_1 = 68.05$ (first row) and $\bar{y}_2 = 67.29$ (second row) with $s_p^2 = 7.066$. For the two–sample t–test, $|t| = \frac{|68.05-67.27|}{\sqrt{7.066\left(\frac{2}{8}\right)}} = .57$ with 14 degrees of freedom. Since $t_{.05} = 1.761$, H_0 cannot be rejected.

15.51 From Ex. 15.18, let A represent school A and let B represent school B. The sequence of runs is given by

$$\underline{A}\ \underline{B}\ \underline{A}\ \underline{B}\ \underline{A}\ \underline{B\ B\ B}\ \underline{A}\ \underline{B\ B}\ \underline{A\ A}\ \underline{B}\ \underline{A}\ \underline{B}\ \underline{A\ A}$$

Notice that the 9th and 10th letters and the 13th and 14th letters in the sequence represent the two pairs of tied observations. If the tied observations were reversed in the sequence of runs, the value of R would remain the same: $R = 13$. Hence the order of the tied observations is irrelevant.

The alternative hypothesis asserts that the two distributions are not identical. Therein, a small number of runs would be expected since most of the observations from school A would fall below those from school B. So, a one–tailed test is employed (lower tail) so the p–value = $P(R \leq 13) = .956$. Thus, we fail to reject the null hypothesis (similar with Ex. 15.18).

15.52 Refer to Ex. 15.25. In this exercise, $n_1 = 15$ and $n_2 = 16$. If the experimental batteries have a greater mean life, we would expect that most of the observations from plant B to be smaller than those from plant A. Consequently, the number of runs would be small. To use the large sample test, note that $E(R) = 16$ and $V(R) = 7.24137$. Thus, since $R = 15$, the approximate p–value is given by

$$P(R \leq 15) = P(R < 15.5) \approx P(Z < -.1858) = .4263.$$

Of course, the hypotheses H_0: the two distributions are equal, would not be rejected.

15.53 R:

```
> grader <- c(9,6,7,7,5,8,2,6,1,10,9,3)
> moisture <- c(.22,.16,.17,.14,.12,.19,.10,.12,.05,.20,.16,.09)
> cor(grader,moisture,method="spearman")
[1] 0.911818
```

Thus, $r_S = .911818$. To test for association with $\alpha = .05$, index .025 in Table 11 so the rejection region is $|r_S| > .591$. Thus, we can safely conclude that the two variables are correlated.

15.54 R:

```
> days <- c(30,47,26,94,67,83,36,77,43,109,56,70)
> rating <- c(4.3,3.6,4.5,2.8,3.3,2.7,4.2,3.9,3.6,2.2,3.1,2.9)
> cor.test(days,rating,method="spearman")

        Spearman's rank correlation rho

data:  days and rating
S = 537.44, p-value = 0.0001651
alternative hypothesis: true rho is not equal to 0
sample estimates:
       rho
-0.8791607
```

From the above, $r_S = -.8791607$ and the p–value for the test H_0: there is no association is given by p–value = .0001651. Thus, H_0 is rejected.

15.55 R:

```
> rank <- c(8,5,10,3,6,1,4,7,9,2)
> score <- c(74,81,66,83,66,94,96,70,61,86)
> cor.test(rank,score,alt = "less",method="spearman")
```

```
        Spearman's rank correlation rho

data:  rank and score
S = 304.4231, p-value = 0.001043
alternative hypothesis: true rho is less than 0
sample estimates:
       rho
-0.8449887
```

a. From the above, $r_S = -.8449887$.

b. With the p–value = .001043, we can conclude that there exists a negative association between the interview rank and test score. Note that we only showed that the correlation is negative and not that the association has some specified level.

15.56 R:

```
> rating <- c(12,7,5,19,17,12,9,18,3,8,15,4)
> distance <- c(75,165,300,15,180,240,120,60,230,200,130,130)
> cor.test(rating,distance,alt = "less",method="spearman")

        Spearman's rank correlation rho

data:  rating and distance
S = 455.593, p-value = 0.02107
alternative hypothesis: true rho is less than 0
sample estimates:
       rho
-0.5929825
```

a. From the above, $r_S = -.5929825$.

b. With the p–value = .02107, we can conclude that there exists a negative association between rating and distance.

15.57 The ranks for the two variables of interest x_i and y_i corresponding the math and art, respectively) are shown in the table below.

Student	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$R(x_i)$	1	3	2	4	5	7.5	7.5	9	10.5	12	13.5	6	13.5	15	10.5
$R(y_i)$	5	11.5	1	2	3.5	8.5	3.5	13	6	15	11.5	7	10	14	8.5

Then, $r_S = \dfrac{15(1148.5) - 120(120)}{\sqrt{[15(1238.5) - 120^2]^2}} = .6768$ (the formula simplifies as shown since the samples of ranks are identical for both math and art). From Table 11 and with α = .10, the rejection region is $|r_S| > .441$ and thus we can conclude that there is a correlation between math and art scores.

15.58 R:

```
> bending <- c(419,407,363,360,257,622,424,359,346,556,474,441)
> twisting <- c(227,231,200,211,182,304,384,194,158,225,305,235)
> cor.test(bending,twisting,method="spearman",alt="greater")
```

```
        Spearman's rank correlation rho

data:  bending and twisting
S = 54, p-value = 0.001097
alternative hypothesis: true rho is greater than 0
sample estimates:
      rho
0.8111888
```

a. From the above, r_S = .8111888.

b. With a p–value = .001097, we can conclude that there is existence of a population association between bending and twisting stiffness.

15.59 The data are ranked below; since there are no ties in either sample, the alternate formula for r_S will be used.

$R(x_i)$	2	3	1	4	6	8	5	10	7	9
$R(y_i)$	2	3	1	4	6	8	5	10	7	9
d_i	0	0	0	0	0	0	0	0	0	0

Thus, $r_S = 1 - \frac{6[(0)^2+(0)^2+\ldots+(0)^2}{10(99)} = 1 - 0 = 1.$

From Table 11, note that 1 > .794 so the p–value < .005 and we soundly conclude that there is a positive correlation between the two variables.

15.60 It is found that r_S = .9394 with n = 10. From Table 11, the p–value < 2(.005) = .01 so we can conclude that correlation is present.

15.61 **a.** Since all five judges rated the three products, this is a randomized block design.

b. Since the measurements are ordinal values and thus integers, the normal theory would not apply.

c. Given the response to part b, we can employ the Friedman test. In R, this is (using the numbers 1–5 to denote the judges):

```
> rating <- c(16,16,14,15,13,9,7,8,16,11,7,8,4,9,2)
> brand <- factor(c(rep("HC",5),rep("S",5),rep("EB",5)))
> judge <- c(1:5,1:5,1:5)
> friedman.test(rating ~ brand | judge)

        Friedman rank sum test

data:  rating and brand and judge
Friedman chi-squared = 6.4, df = 2, p-value = 0.04076
```

With the (approximate) p–value = .04076, we can conclude that the distributions for rating the egg substitutes are not the same.

15.62 Let $p = P$(gourmet A's rating exceeds gourmet B's rating for a given meal). The hypothesis of interest is H_0: $p = .5$ vs H_a: $p \neq .5$. With M = # of meals for which A is superior, we find that

$$P(M \le 4) + P(M \ge 13) = 2P(M \le 4) = .04904.$$

using a binomial calculation with $n = 17$ (3 were ties) and $p = .5$. From the table, $m = 8$ so we fail to reject H_0.

15.63 Using the Wilcoxon signed–rank test,

```
> A <- c(6,4,7,8,2,7,9,7,2,4,6,8,4,3,6,9,9,4,4,5)
> B <- c(8,5,4,7,3,4,9,8,5,3,9,5,2,3,8,10,8,6,3,5)
> wilcox.test(A,B,paired=T)

        Wilcoxon signed rank test

data:  A and B
V = 73.5, p-value = 0.9043
alternative hypothesis: true mu is not equal to 0
```

With the p–value = .9043, the hypothesis of equal distributions is not rejected (as in Ex. 15.63).

15.64 For the Mann–Whitney U test, $W_A = 126$ and $W_B = 45$. So, with $n_1 = n_2 = 9$, $U_A = 0$ and $U_B = 81$. From Table 8, the lower tail of the two–tailed rejection region is $\{U \le 18\}$ with $\alpha = 2(.0252) = .0504$. With $U = 0$, we soundly reject the null hypothesis and conclude that the deaf children do differ in eye movement rate.

15.65 With $n_1 = n_2 = 8$, $U_A = 46.5$ and $U_B = 17.5$. From Table 8, the hypothesis of no difference will be rejected if $U \le 13$ with $\alpha = 2(.0249) = .0498$. Since our $U = 17.5$, we fail to reject H_0 (same as in Ex. 13.1).

15.66 **a.** The measurements are ordered below according to magnitude as mentioned in the exercise (from the "outside in"):

Instrument	*A*	*B*	*A*	*B*	*B*	*B*	*A*	*A*	*A*
Response	1060.21	1060.24	1060.27	1060.28	1060.30	1060.32	1060.34	1060.36	1060.40
Rank	1	3	5	7	9	8	6	4	2

To test H_0: $\sigma_A^2 = \sigma_B^2$ vs. H_a: $\sigma_A^2 > \sigma_B^2$, we use the Mann–Whitney U statistic. If H_a is true, then the measurements for A should be assigned lower ranks. For the significance level, we will use $\alpha = P(U \le 3) = .056$. From the above table, the values are $U_1 = 17$ and $U_2 = 3$. So, we reject H_0.

b. For the two samples, $s_A^2 = .00575$ and $s_B^2 = .00117$. Thus, $F = .00575/.00117 = 4.914$ with 4 numerator and 3 denominator degrees of freedom. From R:

```
> 1 - pf(4.914,4,3)
[1] 0.1108906
```

Since the p–value = .1108906, H_0 would not be rejected.

15.67 First, obviously $P(U \le 2) = P(U = 0) + P(U = 1) + P(U = 2)$. Denoting the five observations from samples 1 and 2 as A and B respectively (and $n_1 = n_2 = 5$), the only sample point associated with $U = 0$ is

$$B\ B\ B\ B\ B\ A\ A\ A\ A\ A$$

because there are no A's preceding any of the B's. The only sample point associated with $U = 1$ is

$$B\ B\ B\ B\ A\ B\ A\ A\ A\ A$$

since only one A observation precedes a B observation. Finally, there are two sample points associated with $U = 2$:

$$B\ B\ B\ A\ B\ B\ A\ A\ A\ A \qquad B\ B\ B\ B\ A\ A\ B\ A\ A\ A$$

Now, under the null hypothesis all of the $\binom{10}{5} = 252$ orderings are equally likely. Thus,

$$P(U \le 2) = 4/252 = 1/63 = .0159.$$

15.68 Let Y = # of positive differences and let T = the rank sum of the positive differences. Then, we must find $P(T \le 2) = P(T = 0) + P(T = 1) + P(T = 2)$. Now, consider the three pairs of observations and the ranked differences according to magnitude. Let d_1, d_2, and d_3 denote the ranked differences. The possible outcomes are:

d_1	d_2	d_3	Y	T
+	+	+	3	6
–	+	+	2	5
+	–	+	2	4
+	+	–	2	3
–	–	+	1	3
–	+	–	1	2
+	–	–	1	1
–	–	–	0	0

Now, under H_0 Y is binomial with $n = 3$ and $p = P(A \text{ exceeds } B) = .5$. Thus,
$P(T = 0) = P(T = 0, Y = 0) = P(Y = 0)P(T = 0 \mid Y = 0) = .125(1) = .125$.

Similarly, $P(T = 1) = P(T = 1, Y = 1) = P(Y = 1)P(T = 1 \mid Y = 1) = ..375(1/3) = .125$, since conditionally when $Y = 1$, there are three possible values for T (1, 2, or 3).

Finally, $P(T = 2) = P(T = 2, Y = 1) = P(Y = 1)P(T = 2 \mid Y = 1) = ..375(1/3) = .125$, using similar logic as in the above.

Thus, $P(T \le 2) = .125 + .125 + .125 = .375$.

15.69 **a.** A composite ranking of the data is:

Line 1	Line 2	Line 3
19	14	2
16	10	15
12	5	4
20	13	11
3	9	1
18	17	8
21	7	6
$R_1 = 109$	$R_2 = 75$	$R_3 = 47$

Thus,

$$H = \tfrac{12}{21(22)}\left[\tfrac{109^2}{7} + \tfrac{75^2}{7} + \tfrac{47}{7}\right] = 3(22) = 7.154$$

with 2 degrees of freedom. Since $\chi^2_{.05} = 5.99147$, we can reject the claim that the population distributions are equal.

15.70 **a.** R:

```
> rating <- c(20,19,20,18,17,17,11,13,15,14,16,16,15,13,18,11,8,
12,10,14,9,10)
> supervisor <- factor(c(rep("I",5),rep("II",6),rep("III",5),
rep("IV",6)))
> kruskal.test(rating~supervisor)

        Kruskal-Wallis rank sum test

data:  rating by supervisor
Kruskal-Wallis chi-squared = 14.6847, df = 3, p-value = 0.002107
```

With a p–value = .002107, we can conclude that one or more of the supervisors tend to receive higher ratings

b. To conduct a Mann–Whitney U test for only supervisors I and III,

```
> wilcox.test(rating[12:16],rating[1:5], correct=F)

        Wilcoxon rank sum test

data:  rating[12:16] and rating[1:5]
W = 1.5, p-value = 0.02078
alternative hypothesis: true mu is not equal to 0
```

Thus, with a p–value = .02078, we can conclude that the distributions of ratings for supervisors I and III differ by location.

15.71 Using Friedman's test (people are blocks), $R_1 = 19$, $R_2 = 21.5$, $R_3 = 27.5$ and $R_4 = 32$. To test H_0: the distributions for the items are equal vs.
H_a: at least two of the distributions are different

the test statistic is $F_r = \frac{12}{10(4)(5)}\left[19^2 + (21.5)^2 + (27.5)^2 + 32^2\right] - 3(10)(5) = 6.21$.

With 3 degrees of freedom, $\chi^2_{.05} = 7.81473$ and so H_0 is not rejected.

15.72 In R:

```
> perform <- c(20,25,30,37,24,16,22,25,40,26,20,18,24,27,39,41,21,25)
> group <- factor(c(1:6,1:6,1:6))
> method <- factor(c(rep("lect",6),rep("demonst",6),rep("machine",6)))
> friedman.test(perform ~ method | group)

        Friedman rank sum test

data:  perform and method and group
Friedman chi-squared = 4.2609, df = 2, p-value = 0.1188
```

With a p–value = .1188, it is unwise to reject the claim of equal teach method effectiveness, so fail to reject H_0.

15.73 Following the methods given in Section 15.9, we must obtain the probability of observing exactly Y_1 runs of S and Y_2 runs of F, where $Y_1 + Y_2 = R$. The joint probability mass functions for Y_1 and Y_2 is given by

$$p(y_1, y_2) = \frac{\binom{7}{y_1-1}\binom{7}{y_2-1}}{\binom{16}{8}}.$$

(1) For the event $R = 2$, this will only occur if $Y_1 = 1$ and $Y_2 = 1$, with either the S elements or the F elements beginning the sequence. Thus,
$$P(R = 2) = 2p(1, 1) = \tfrac{2}{12{,}870}.$$
(2) For $R = 3$, this will occur if $Y_1 = 1$ and $Y_2 = 2$ or $Y_1 = 2$ and $Y_2 = 1$. So,
$$P(R = 3) = p(1, 2) + p(2, 1) = \tfrac{14}{12{,}870}.$$
(3) Similarly, $P(R = 4) = 2p(2, 2) = \frac{98}{12{,}870}$.

(4) Likewise, $P(R = 5) = p(3, 2) + p(2, 3) = \frac{294}{12{,}870}$.

(5) In the same manor, $P(R = 6) = 2p(3, 3) = \frac{882}{12{,}870}$.

Thus, $P(R \le 6) = \frac{2+14+98+294+882}{12{,}870} = .100$, agreeing with the entry found in Table 10.

15.74 From Ex. 15.67, it is not difficult to see that the following pairs of events are equivalent:

$$\{W = 15\} \equiv \{U = 0\},\ \{W = 16\} \equiv \{U = 2\},\text{ and } \{W = 17\} \equiv \{U = 3\}.$$

Therefore, $P(W \le 17) = \mathrm{P}(U \le 3) = .0159$.

15.75 Assume there are n_1 "A" observations and n_2 "B" observations, The Mann–Whitney U statistic is defined as

$$U = \sum_{i=1}^{n_2} U_i \ ,$$

where U_i is the number of A observations preceding the i^{th} B. With $B_{(i)}$ to be the i^{th} B observation in the combined sample after it is ranked from smallest to largest, and write $R[B_{(i)}]$ to be the rank of the i^{th} ordered B in the total ranking of the combined sample. Then, U_i is the number of A observations the precede $B_{(i)}$. Now, we know there are $(i - 1)$ B's that precede $B_{(i)}$, and that there are $R[B_{(i)}] - 1$ A's and B's preceding $B_{(i)}$. Then,

$$U = \sum_{i=1}^{n_2} U_i = \sum_{i=1}^{n_2} [R(B_{(i)}) - i] = \sum_{i=1}^{n_2} R(B_{(i)}) - \sum_{i=1}^{n_2} i = W_B - n_2(n_2+1)/2$$

Now, let $N = n_1 + n_2$. Since $W_A + W_B = N(N + 1)/2$, so $W_B = N(N + 1)/2 - W_A$. Plugging this expression in to the one for U yields

$$U = N(N+1)/2 - n_2(n_2+1)/2 - W_A = \tfrac{N^2+N+n_2^2+n_2}{2} - W_A$$
$$= \tfrac{n_1^2+2n_1n_2+n_2^2+n_1+n_2-n_2^2-n_2}{2} - W_A = n_1n_2 + \tfrac{n_1(n_1+1)}{2} - W_A \, .$$

Thus, the two tests are equivalent.

15.76 Using the notation introduced in Ex. 15.65, note that

$$W_A = \sum_{i=1}^{n_1} R(A_i) = \sum_{i=1}^{N} X_i \ ,$$

where

$$X_i = \begin{cases} R(z_i) & \text{if } z_i \text{ is from sample } A \\ 0 & \text{if } z_i \text{ is from sample } B \end{cases}$$

If H_0 is true,

$$E(X_i) = R(z_i)P[X_i = R(z_i)] + 0\cdot P(X_i = 0) = R(z_i)\tfrac{n_1}{N}$$
$$E(X_i^2) = [R(z_i)]^2 \tfrac{n_1}{N}$$
$$V(X_i) = [R(z_i)]^2 \tfrac{n_1}{N} - \left(R(z_i)\tfrac{n_1}{N}\right)^2 = [R(z_i)]^2\left(\tfrac{n_1(N-n_1)}{N^2}\right).$$
$$E(X_i, X_j) = R(z_i)R(z_i)P[X_i = R(z_i), X_j = R(z_j)] = R(z_i)R(z_i)\left(\tfrac{n_1}{N}\right)\left(\tfrac{n_1-1}{N-1}\right).$$

From the above, it can be found that $\text{Cov}(X_i, X_j) = R(z_i)R(z_i)\left[\tfrac{-n_1(N-n_1)}{N^2(N-1)}\right]$.

Therefore,

$$E(W_A) = \sum_{i=1}^{N} E(X_i) = \tfrac{n_1}{N}\sum_{i=1}^{N} R(z_i) = \tfrac{n_1}{N}\left(\tfrac{N(N+1)}{2}\right) = \tfrac{n_1(N+1)}{2}$$

and

$$\begin{aligned}
V(W_A) &= \sum_{i=1}^{N} V(X_i) + \sum_{i \neq j} \text{Cov}(X_i, X_j) \\
&= \tfrac{n_1(N-n_1)}{N^2} \sum_{i=1}^{N} [R(z_i)]^2 - \tfrac{n_1(N-n_1)}{N^2(N-1)} \left[\sum_{i=1}^{N} \sum_{j=1}^{N} R(z_i)R(z_j) - \sum_{i=1}^{N} [R(z_i)]^2 \right] \\
&= \tfrac{n_1(N-n_1)}{N^2} \left[\tfrac{N(N+1)N2N+1)}{6} \right] - \tfrac{n_1(N-n_1)}{N^2(N-1)} \left\{ \left[\sum_{i=1}^{N} R(z_i) \right]^2 - \sum_{i=1}^{N} [R(z_i)]^2 \right\} \\
&= \tfrac{2n_1(N-n_1)(N+1)(2N+1)}{12N} - \tfrac{n_1(N-n_1)}{N^2(N-1)} \left[\tfrac{N^2(N+1)^2}{4} - \tfrac{N(N+1)(2N+1)}{6} \right] \\
&= \tfrac{n_1 n_2 (n_1+n_2+1)}{12} \left[\tfrac{4N+2}{N} - \tfrac{(3N+2)(N-1)}{n(N-1)} \right] = \tfrac{n_1 n_2 (n_1+n_2+1)}{12} \,.
\end{aligned}$$

From Ex. 15.75 it was shown that $U = n_1 n_2 + \frac{n_1(n_1+1)}{2} - W_A$. Thus,

$$\begin{aligned}
E(U) &= n_1 n_2 + \tfrac{n_1(n_1+1)}{2} - E(W_A) = \tfrac{n_1 n_2}{2} \\
V(U) &= V(W_A) = \tfrac{n_1 n_2 (n_1+n_2+1)}{12} \,.
\end{aligned}$$

15.77 Recall that in order to obtain T, the Wilcoxon signed–rank statistic, the differences d_i are calculated and ranked according to absolute magnitude. Then, using the same notation as in Ex. 15.76,

$$T^+ = \sum_{i=1}^{N} X_i$$

where

$$X_i = \begin{cases} R(D_i) & \text{if } D_i \text{ is positive} \\ 0 & \text{if } D_i \text{ is negative} \end{cases}$$

When H_0 is true, $p = P(D_i > 0) = \frac{1}{2}$. Thus,

$$\begin{aligned}
E(X_i) &= R(D_i)P[X_i = R(D_i)] = \tfrac{1}{2} R(D_i) \\
E(X_i^2) &= [R(D_i)]^2 P[X_i = R(D_i)] = \tfrac{1}{2}[R(D_i)]^2 \\
V(X_i) &= \tfrac{1}{2}[R(D_i)]^2 = [\tfrac{1}{2} R(D_i)]^2 = \tfrac{1}{4}[R(D_i)]^2 \\
E(X_i, X_j) &= R(D_i)R(D_j)P[X_i = R(D_i), X_j = R(D_j)] = \tfrac{1}{4} R(D_i)R(D_j) \,.
\end{aligned}$$

Then, $\text{Cov}(X_i, X_j) = 0$ so

$$E(T^+) = \sum_{i=1}^{n} E(X_i) = \tfrac{1}{2} \sum_{i=1}^{n} R(D_i) = \tfrac{1}{2} \left(\tfrac{n(n+1)}{2} \right) = \tfrac{n(n+1)}{4}$$

$$V(T^+) = \sum_{i=1}^{n} V(X_i) = \tfrac{1}{4} \sum_{i=1}^{n} [R(D_i)]^2 = \tfrac{1}{4} \left(\tfrac{n(n+1)(2n+1)}{6} \right) = \tfrac{n(n+1)(2n+1)}{24} \,.$$

Since $T^- = \frac{n(n+1)}{2} - T^+$ (see Ex. 15.10),

$$\begin{aligned}
E(T^-) &= E(T^+) = E(T) \\
V(T^-) &= V(T^+) = V(T) \,.
\end{aligned}$$

15.78 Since we use X_i to denote the rank of the i^{th} "X" sample value and Y_i to denote the rank of the i^{th} "Y" sample value,

$$\sum_{i=1}^{n} X_i = \sum_{i=1}^{n} Y_i = \tfrac{n(n+1)}{2} \text{ and } \sum_{i=1}^{n} X_i^2 = \sum_{i=1}^{n} Y_i^2 = \tfrac{n(n+1)(2n+1)}{6}.$$

Then, define $d_i = X_i - Y_i$ so that

$$\sum_{i=1}^{n} d_i^2 = \sum_{i=1}^{n}\left(X_i^2 - 2X_iY_i + Y_i^2\right) = \tfrac{n(n+1)(2n+1)}{6} - 2\sum_{i=1}^{n} X_iY_i + \tfrac{n(n+1)(2n+1)}{6}$$

and thus

$$\sum_{i=1}^{n} X_iY_i = \tfrac{n(n+1)(2n+1)}{6} - \tfrac{1}{2}\sum_{i=1}^{n} d_i^2 .$$

Now, we have

$$\begin{aligned}
r_S &= \frac{n\sum_{i=1}^{n} X_iY_i - \left(\sum_{i=1}^{n} X_i\right)\left(\sum_{i=1}^{n} Y_i\right)}{\sqrt{\left[n\sum_{i=1}^{n} X_i^2 - \left(\sum_{i=1}^{n} X_i\right)^2\right]}\sqrt{\left[n\sum_{i=1}^{n} Y_i^2 - \left(\sum_{i=1}^{n} Y_i\right)^2\right]}} \\
&= \frac{\frac{n^2(n+1)(2n+1)}{6} - \frac{n}{2}\sum_{i=1}^{n} d_i^2 - \frac{n^2(n+1)^2}{4}}{\frac{n^2(n+1)(2n+1)}{6} - \frac{n^2(n+1)^2}{4}} \\
&= \frac{\frac{n^2(n+1)(n-1)}{12} - \frac{n}{2}\sum_{i=1}^{n} d_i^2}{\frac{n^2(n+1)(n-1)}{12}} \\
&= 1 - \frac{\frac{n}{2}\sum_{i=1}^{n} d_i^2}{\frac{n^2(n^2-1)}{12}} \\
&= 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2-1)} .
\end{aligned}$$

Chapter 16: Introduction to Bayesian Methods of Inference

16.1 Refer to Table 16.1.

a. $\beta(10,30)$

b. $n = 25$

c. $\beta(10,30)$**,** $n = 25$

d. Yes

e. Posterior for the $\beta(1,3)$ prior.

16.2 **a.-d.** Refer to Section 16.2

16.3 **a.-e.** Applet exercise, so answers vary.

16.4 **a.-d.** Applex exercise, so answers vary.

16.5 It should take more trials with a beta(10, 30) prior.

16.6 Here, $L(y \mid p) = p(y \mid p) = \binom{n}{y} p^y (1-p)^{n-y}$, where $y = 0, 1, \ldots, n$ and $0 < p < 1$. So,

$$f(y, p) = \binom{n}{y} p^y (1-p)^{n-y} \times \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1}(1-p)^{\beta-1}$$

so that

$$m(y) = \int_0^1 \binom{n}{y} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{y+\alpha-1}(1-p)^{n-y+\beta-1} dp = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(y+\alpha)\Gamma(n-y+\beta)}{\Gamma(n+\alpha+\beta)}.$$

The posterior density of p is then

$$g^*(p \mid y) = \frac{\Gamma(n+\alpha+\beta)}{\Gamma(y+\alpha)\Gamma(n-y+\beta)} p^{y+\alpha-1}(1-p)^{n-y+\beta-1},\ 0 < p < 1.$$

This is the identical beta density as in Example 16.1 (recall that the sum of n i.i.d. Bernoulli random variables is binomial with n trials and success probability p).

16.7 **a.** The Bayes estimator is the mean of the posterior distribution, so with a beta posterior with $\alpha = y + 1$ and $\beta = n - y + 3$ in the prior, the posterior mean is

$$\hat{p}_B = \frac{Y+1}{n+4} = \frac{Y}{n+4} + \frac{1}{n+4}.$$

b. $E(\hat{p}_B) = \frac{E(Y)+1}{n+4} = \frac{np+1}{n+4} \neq p$, $V(\hat{p}) = \frac{V(Y)}{(n+4)^2} = \frac{np(1-p)}{(n+4)^2}$

16.8 **a.** From Ex. 16.6, the Bayes estimator for p is $\hat{p}_B = E(p \mid Y) = \frac{Y+1}{n+2}$.

b. This is the uniform distribution in the interval (0, 1).

c. We know that $\hat{p} = Y/n$ is an unbiased estimator for p. However, for the Bayes estimator,

$$E(\hat{p}_B) = \frac{E(Y)+1}{n+2} = \frac{np+1}{n+2} \text{ and } V(\hat{p}_B) = \frac{V(Y)}{(n+2)^2} = \frac{np(1-p)}{(n+2)^2}.$$

Thus, $MSE(\hat{p}_B) = V(\hat{p}_B) + [B(\hat{p}_B)]^2 = \frac{np(1-p)}{(n+2)^2} + \left(\frac{np+1}{n+2} - p\right)^2 = \frac{np(1-p)+(1-2p)^2}{(n+2)^2}$.

d. For the unbiased estimator $\hat{p}$, MSE($\hat{p}$) = V($\hat{p}$) = $p(1-p)/n$. So, holding n fixed, we must determine the values of p such that

$$\frac{np(1-p)+(1-2p)^2}{(n+2)^2} < \frac{p(1-p)}{n}.$$

The range of values of p where this is satisfied is solved in Ex. 8.17(c).

16.9 **a.** Here, $L(y \mid p) = p(y \mid p) = (1-p)^{y-1} p$, where $y = 1, 2, \ldots$ and $0 < p < 1$. So,

$$f(y,p) = (1-p)^{y-1} p \times \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1}(1-p)^{\beta-1}$$

so that

$$m(y) = \int_0^1 \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha}(1-p)^{\beta+y-2} dp = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(\alpha+1)\Gamma(y+\beta-1)}{\Gamma(y+\alpha+\beta)}.$$

The posterior density of p is then

$$g^*(p \mid y) = \frac{\Gamma(\alpha+\beta+y)}{\Gamma(\alpha+1)\Gamma(\beta+y-1)} p^{\alpha}(1-p)^{\beta+y-2}, \ 0 < p < 1.$$

This is a beta density with shape parameters $\alpha^* = \alpha + 1$ and $\beta^* = \beta + y - 1$.

b. The Bayes estimators are

(1) $\hat{p}_B = E(p \mid Y) = \frac{\alpha+1}{\alpha+\beta+Y}$,

(2) $$[p(1-p)]_B = E(p \mid Y) - E(p^2 \mid Y) = \frac{\alpha+1}{\alpha+\beta+Y} - \frac{(\alpha+2)(\alpha+1)}{(\alpha+\beta+Y+1)(\alpha+\beta+Y)}$$
$$= \frac{(\alpha+1)(\beta+Y-1)}{(\alpha+\beta+Y+1)(\alpha+\beta+Y)},$$

where the second expectation was solved using the result from Ex. 4.200. (Alternately, the answer could be found by solving $E[p(1-p) \mid Y] = \int_0^1 p(1-p) g^*(p \mid Y) dp$.

16.10 **a.** The joint density of the random sample and θ is given by the product of the marginal densities multiplied by the gamma prior:

$$
\begin{aligned}
f(y_1,\ldots,y_n,\theta) &= \left[\prod_{i=1}^{n} \theta \exp(-\theta y_i)\right] \frac{1}{\Gamma(\alpha)\beta^{\alpha}} \theta^{\alpha-1} \exp(-\theta/\beta) \\
&= \frac{\theta^{n+\alpha-1}}{\Gamma(\alpha)\beta^{\alpha}} \exp\left(-\theta \sum_{i=1}^{n} y_i - \theta/\beta\right) = \frac{\theta^{n+\alpha-1}}{\Gamma(\alpha)\beta^{\alpha}} \exp\left(-\theta \Big/ \frac{\beta}{\beta\sum_{i=1}^{n} y_i + 1}\right)
\end{aligned}
$$

b. $m(y_1,\ldots,y_n) = \dfrac{1}{\Gamma(\alpha)\beta^{\alpha}} \displaystyle\int_0^{\infty} \theta^{n+\alpha-1} \exp\left(-\theta \Big/ \frac{\beta}{\beta\sum_{i=1}^{n} y_i + 1}\right) d\theta$, but this integral resembles that of a gamma density with shape parameter $n + \alpha$ and scale parameter $\dfrac{\beta}{\beta\sum_{i=1}^{n} y_i + 1}$.

Thus, the solution is $m(y_1,\ldots,y_n) = \dfrac{1}{\Gamma(\alpha)\beta^{\alpha}} \Gamma(n+\beta) \left(\dfrac{\beta}{\beta\sum_{i=1}^{n} y_i + 1}\right)^{n+\alpha}$.

c. The solution follows from parts (a) and (b) above.

d. Using the result in Ex. 4.111,

$$
\begin{aligned}
\hat{\mu}_B = E(\mu \mid \boldsymbol{Y}) = E(1/\theta \mid \boldsymbol{Y}) &= \frac{1}{\beta^*(\alpha^* - 1)} = \left[\frac{\beta}{\beta\sum_{i=1}^{n} Y_i + 1}(n+\alpha-1)\right]^{-1} \\
&= \frac{\beta\sum_{i=1}^{n} Y_i + 1}{\beta(n+\alpha-1)} = \frac{\sum_{i=1}^{n} Y_i}{n+\alpha-1} + \frac{1}{\beta(n+\alpha-1)}
\end{aligned}
$$

e. The prior mean for $1/\theta$ is $E(1/\theta) = \dfrac{1}{\beta(\alpha-1)}$ (again by Ex. 4.111). Thus, $\hat{\mu}_B$ can be written as

$$
\hat{\mu}_B = \overline{Y}\left(\frac{n}{n+\alpha-1}\right) + \frac{1}{\beta(\alpha-1)}\left(\frac{\alpha-1}{n+\alpha-1}\right),
$$

which is a weighted average of the MLE and the prior mean.

f. We know that $\overline{Y}$ is unbiased; thus $E(\overline{Y}) = \mu = 1/\theta$. Therefore,

$$
E(\hat{\mu}_B) = E(\overline{Y})\left(\frac{n}{n+\alpha-1}\right) + \frac{1}{\beta(\alpha-1)}\left(\frac{\alpha-1}{n+\alpha-1}\right) = \frac{1}{\theta}\left(\frac{n}{n+\alpha-1}\right) + \frac{1}{\beta(\alpha-1)}\left(\frac{\alpha-1}{n+\alpha-1}\right).
$$

Therefore, $\hat{\mu}_B$ is biased. However, it is asymptotically unbiased since

$$
E(\hat{\mu}_B) - 1/\theta \to 0.
$$

Also,

$$V(\hat{\mu}_B) = V(\bar{Y})\left(\frac{n}{n+\alpha-1}\right)^2 = \frac{1}{\theta^2 n}\left(\frac{n}{n+\alpha-1}\right)^2 = \frac{1}{\theta^2}\frac{n}{(n+\alpha-1)^2} \to 0.$$

So, $\hat{\mu}_B \xrightarrow{p} 1/\theta$ and thus it is consistent.

16.11 **a.** The joint density of U and λ is

$$\begin{aligned} f(u,\lambda) = p(u \mid \lambda)g(\lambda) &= \frac{(n\lambda)^u \exp(-n\lambda)}{u!} \times \frac{1}{\Gamma(\alpha)\beta^\alpha}\lambda^{\alpha-1}\exp(-\lambda/\beta) \\ &= \frac{n^u}{u!\Gamma(\alpha)\beta^\alpha}\lambda^{u+\alpha-1}\exp(-n\lambda-\lambda/\beta) \\ &= \frac{n^u}{u!\Gamma(\alpha)\beta^\alpha}\lambda^{u+\alpha-1}\exp\left[-\lambda\Big/\left(\frac{\beta}{n\beta+1}\right)\right] \end{aligned}$$

b. $m(u) = \dfrac{n^u}{u!\Gamma(\alpha)\beta^\alpha}\displaystyle\int_0^\infty \lambda^{u+\alpha-1}\exp\left[-\lambda\Big/\left(\frac{\beta}{n\beta+1}\right)\right]d\lambda$, but this integral resembles that of a gamma density with shape parameter $u + \alpha$ and scale parameter $\dfrac{\beta}{n\beta+1}$. Thus, the solution is $m(u) = \dfrac{n^u}{u!\Gamma(\alpha)\beta^\alpha}\Gamma(u+\alpha)\left(\dfrac{\beta}{n\beta+1}\right)^{u+\alpha}$.

c. The result follows from parts (a) and (b) above.

d. $\hat{\lambda}_B = E(\lambda \mid U) = \alpha^*\beta^* = (U+\alpha)\left(\dfrac{\beta}{n\beta+1}\right)$.

e. The prior mean for λ is $E(\lambda) = \alpha\beta$. From the above,

$$\hat{\lambda}_B = \left(\sum_{i=1}^n Y_i + \alpha\right)\left(\frac{\beta}{n\beta+1}\right) = \bar{Y}\left(\frac{n\beta}{n\beta+1}\right) + \alpha\beta\left(\frac{1}{n\beta+1}\right),$$

which is a weighted average of the MLE and the prior mean.

f. We know that $\bar{Y}$ is unbiased; thus $E(\bar{Y}) = \lambda$ Therefore,

$$E(\hat{\lambda}_B) = E(\bar{Y})\left(\frac{n\beta}{n\beta+1}\right) + \alpha\beta\left(\frac{1}{n\beta+1}\right) = \lambda\left(\frac{n\beta}{n\beta+1}\right) + \alpha\beta\left(\frac{1}{n\beta+1}\right).$$

So, $\hat{\lambda}_B$ is biased but it is asymptotically unbiased since

$$E(\hat{\lambda}_B) - \lambda \to 0.$$

Also,

$$V(\hat{\lambda}_B) = V(\bar{Y})\left(\frac{n\beta}{n\beta+1}\right)^2 = \frac{\lambda}{n}\left(\frac{n\beta}{n\beta+1}\right)^2 = \lambda\frac{n\beta}{(n\beta+1)^2} \to 0.$$

So, $\hat{\lambda}_B \xrightarrow{p} \lambda$ and thus it is consistent.

16.12 First, it is given that $W = vU = v\sum_{i=1}^{n}(Y_i - \mu_0)^2$ is chi–square with n degrees of freedom. Then, the density function for U (conditioned on v) is given by

$$f_U(u \mid v) = |v| f_W(uv) = v\frac{1}{\Gamma(n/2)2^{n/2}}(uv)^{n/2-1}e^{-uv/2} = \frac{1}{\Gamma(n/2)2^{n/2}}u^{n/2-1}v^{n/2}e^{-uv/2}.$$

a. The joint density of U and v is then

$$\begin{aligned} f(u,v) = f_U(u \mid v)g(v) &= \frac{1}{\Gamma(n/2)2^{n/2}}u^{n/2-1}v^{n/2}\exp(-uv/2)\times\frac{1}{\Gamma(\alpha)\beta^{\alpha}}v^{\alpha-1}\exp(-v/\beta) \\ &= \frac{1}{\Gamma(n/2)\Gamma(\alpha)2^{n/2}\beta^{\alpha}}u^{n/2-1}v^{n/2+\alpha-1}\exp(-uv/2 - v/\beta) \\ &= \frac{1}{\Gamma(n/2)\Gamma(\alpha)2^{n/2}\beta^{\alpha}}u^{n/2-1}v^{n/2+\alpha-1}\exp\left[-v\Big/\left(\frac{2\beta}{u\beta+2}\right)\right]. \end{aligned}$$

b. $m(u) = \dfrac{1}{\Gamma(n/2)\Gamma(\alpha)2^{n/2}\beta^{\alpha}}u^{n/2-1}\displaystyle\int_0^{\infty} v^{n/2+\alpha-1}\exp\left[-v\Big/\left(\frac{2\beta}{u\beta+2}\right)\right]dv$, but this integral resembles that of a gamma density with shape parameter $n/2 + \alpha$ and scale parameter $\dfrac{2\beta}{u\beta+2}$. Thus, the solution is $m(u) = \dfrac{u^{n/2-1}}{\Gamma(n/2)\Gamma(\alpha)2^{n/2}\beta^{\alpha}}\Gamma(n/2+\alpha)\left(\dfrac{2\beta}{u\beta+2}\right)^{n/2+\alpha}$.

c. The result follows from parts (a) and (b) above.

d. Using the result in Ex. 4.111(e),

$$\hat{\sigma}_B^2 = E(\sigma^2 \mid U) = E(1/v \mid U) = \frac{1}{\beta^*(\alpha^*-1)} = \frac{1}{n/2+\alpha-1}\left(\frac{U\beta+2}{2\beta}\right) = \frac{U\beta+2}{\beta(n+2\alpha-2)}.$$

e. The prior mean for $\sigma^2 = 1/v = \dfrac{1}{\beta(\alpha-1)}$. From the above,

$$\hat{\sigma}_B^2 = \frac{U\beta+2}{\beta(n+2\alpha-2)} = \frac{U}{n}\left(\frac{n}{n+2\alpha-2}\right) + \frac{1}{\beta(\alpha-1)}\left(\frac{2(\alpha-1)}{n+2\alpha-2}\right).$$

16.13 **a.** (.099, .710)
b. Both probabilities are .025.

c. $P(.099 < p < .710) = .95$.
d.-g. Answers vary.
h. The credible intervals should decrease in width with larger sample sizes.

16.14 **a.-b.** Answers vary.

16.15 With $y = 4$, $n = 25$, and a beta(1, 3) prior, the posterior distribution for p is beta(5, 24). Using R, the lower and upper endpoints of the 95% credible interval are given by:
```
> qbeta(.025,5,24)
[1] 0.06064291
> qbeta(.975,5,24)
[1] 0.3266527
```

16.16 With $y = 4$, $n = 25$, and a beta(1, 1) prior, the posterior distribution for p is beta(5, 22). Using R, the lower and upper endpoints of the 95% credible interval are given by:
```
> qbeta(.025,5,22)
[1] 0.06554811
> qbeta(.975,5,22)
[1] 0.3486788
```

This is a wider interval than what was obtained in Ex. 16.15.

16.17 With $y = 6$ and a beta(10, 5) prior, the posterior distribution for p is beta(11, 10). Using R, the lower and upper endpoints of the 80% credible interval for p are given by:
```
> qbeta(.10,11,10)
[1] 0.3847514
> qbeta(.90,11,10)
[1] 0.6618291
```

16.18 With $n = 15$, $\sum_{i=1}^{n} y_i = 30.27$, and a gamma(2.3, 0.4) prior, the posterior distribution for θ is gamma(17.3, .030516). Using R, the lower and upper endpoints of the 80% credible interval for θ are given by
```
> qgamma(.10,shape=17.3,scale=.0305167)
[1] 0.3731982
> qgamma(.90,shape=17.3,scale=.0305167)
[1] 0.6957321
```

The 80% credible interval for θ is (.3732, .6957). To create a 80% credible interval for $1/\theta$, the end points of the previous interval can be inverted:

$$.3732 < \theta < .6957$$
$$1/(.3732) > 1/\theta > 1/(.6957)$$

Since $1/(.6957) = 1.4374$ and $1/(.3732) = 2.6795$, the 80% credible interval for $1/\theta$ is (1.4374, 2.6795).

16.19 With $n = 25$, $\sum_{i=1}^{n} y_i = 174$, and a gamma(2, 3) prior, the posterior distribution for λ is gamma(176, .0394739). Using R, the lower and upper endpoints of the 95% credible interval for λ are given by

```
> qgamma(.025,shape=176,scale=.0394739)
[1] 5.958895
> qgamma(.975,shape=176,scale=.0394739)
[1] 8.010663
```

16.20 With $n = 8$, $u = .8579$, and a gamma(5, 2) prior, the posterior distribution for v is gamma(9, 1.0764842). Using R, the lower and upper endpoints of the 90% credible interval for v are given by

```
> qgamma(.05,shape=9,scale=1.0764842)
[1] 5.054338
> qgamma(.95,shape=9,scale=1.0764842)
[1] 15.53867
```

The 90% credible interval for v is (5.054, 15.539). Similar to Ex. 16.18, the 90% credible interval for $\sigma^2 = 1/v$ is found by inverting the endpoints of the credible interval for v, given by (.0644, .1979).

16.21 From Ex. 6.15, the posterior distribution of p is beta(5, 24). Now, we can find $P^*(p \in \Omega_0) = P^*(p < .3)$ by (in R):

```
> pbeta(.3,5,24)
[1] 0.9525731
```

Therefore, $P^*(p \in \Omega_a) = P^*(p \geq .3) = 1 - .9525731 = .0474269$. Since the probability associated with H_0 is much larger, our decision is to not reject H_0.

16.22 From Ex. 6.16, the posterior distribution of p is beta(5, 22). We can find $P^*(p \in \Omega_0) = P^*(p < .3)$ by (in R):

```
> pbeta(.3,5,22)
[1] 0.9266975
```

Therefore, $P^*(p \in \Omega_a) = P^*(p \geq .3) = 1 - .9266975 = .0733025$. Since the probability associated with H_0 is much larger, our decision is to not reject H_0.

16.23 From Ex. 6.17, the posterior distribution of p is beta(11, 10). Thus, $P^*(p \in \Omega_0) = P^*(p < .4)$ is given by (in R):

```
> pbeta(.4,11,10)
[1] 0.1275212
```

Therefore, $P^*(p \in \Omega_a) = P^*(p \geq .4) = 1 - .1275212 = .8724788$. Since the probability associated with H_a is much larger, our decision is to reject H_0.

16.24 From Ex. 16.18, the posterior distribution for θ is gamma(17.3, .0305). To test

$$H_0: \theta > .5 \text{ vs. } H_a: \theta \leq .5,$$

we calculate $P^*(\theta \in \Omega_0) = P^*(\theta > .5)$ as:

```
> 1 - pgamma(.5,shape=17.3,scale=.0305)
[1] 0.5561767
```

Therefore, $P^*(\theta \in \Omega_a) = P^*(\theta \geq .5)$ = 1 – .5561767 = .4438233. The probability associated with H_0 is larger (but only marginally so), so our decision is to not reject H_0.

16.25 From Ex. 16.19, the posterior distribution for λ is gamma(176, .0395). Thus, $P^*(\lambda \in \Omega_0) = P^*(\lambda > 6)$ is found by

```
> 1 - pgamma(6,shape=176,scale=.0395)
[1] 0.9700498
```

Therefore, $P^*(\lambda \in \Omega_a) = P^*(\lambda \leq 6)$ = 1 – .9700498 = .0299502. Since the probability associated with H_0 is much larger, our decision is to not reject H_0.

16.26 From Ex. 16.20, the posterior distribution for v is gamma(9, 1.0765). To test:

$$H_0: v < 10 \text{ vs. } H_a: v \geq 10,$$

we calculate $P^*(v \in \Omega_0) = P^*(v < 10)$ as

```
> pgamma(10,9, 1.0765)
[1] 0.7464786
```

Therefore, $P^*(\lambda \in \Omega_a) = P^*(v \geq 10)$ = 1 – .7464786 = .2535214. Since the probability associated with H_0 is larger, our decision is to not reject H_0.