Chapter 2

Random Variables, Distributions, and Expectations

2.1 Discrete; continuous; continuous; discrete; discrete; continuous.

2.2 A table of sample space and assigned values of the random variable is shown next.

Sample Space	x
NNN	0
NNB	1
NBN	1
BNN	1
NBB	2
BNB	2
BBN	2
BBB	3

2.3 A table of sample space and assigned values of the random variable is shown next.

Sample Space	w
HHH	3
HHT	1
HTH	1
THH	1
HTT	-1
THT	-1
TTH	-1
TTT	-3

2.4 $S = \{HHH, THHH, HTHHH, TTHHH, TTTHHH, HTTHHH, THTHHH, HHTHHH\}$; The sample space is discrete.

2.5 (a) $c = 1/30$ since $1 = \sum_{x=0}^{3} c(x^2+4) = 30c$.

(b) $c = 1/10$ since

$$1 = \sum_{x=0}^{2} c\binom{2}{x}\binom{3}{3-x} = c\left[\binom{2}{0}\binom{3}{3} + \binom{2}{1}\binom{3}{2} + \binom{2}{2}\binom{3}{1}\right] = 10c.$$

2.6 (a) $P(X > 200) = \int_{200}^{\infty} \frac{20000}{(x+100)^3}\, dx = -\left.\frac{10000}{(x+100)^2}\right|_{200}^{\infty} = \frac{1}{9}$.

(b) $P(80 < X < 200) = \int_{80}^{120} \frac{20000}{(x+100)^3}\, dx = -\left.\frac{10000}{(x+100)^2}\right|_{80}^{120} = \frac{1000}{9801} = 0.1020$.

2.7 (a) $P(X < 1.2) = \int_0^1 x\, dx + \int_1^{1.2}(2-x)\, dx = \left.\frac{x^2}{2}\right|_0^1 + \left.\left(2x - \frac{x^2}{2}\right)\right|_1^{1.2} = 0.68$.

(b) $P(0.5 < X < 1) = \int_{0.5}^{1} x\, dx = \left.\frac{x^2}{2}\right|_{0.5}^{1} = 0.375$.

2.8 (a) $P(0 < X < 1) = \int_0^1 \frac{2(x+2)}{5}\, dx = \left.\frac{(x+2)^2}{5}\right|_0^1 = 1$.

(b) $P(1/4 < X < 1/2) = \int_{1/4}^{1/2} \frac{2(x+2)}{5}\, dx = \left.\frac{(x+2)^2}{5}\right|_{1/4}^{1/2} = 19/80$.

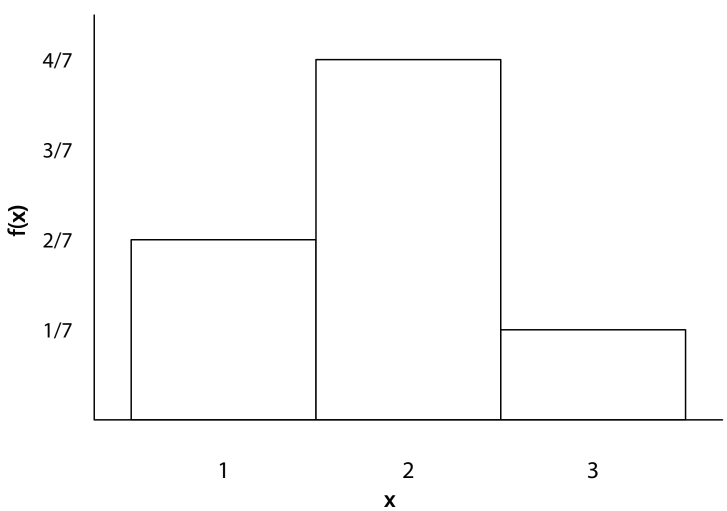
2.9 We can select x defective sets from 2, and $3-x$ good sets from 5 in $\binom{2}{x}\binom{5}{3-x}$ ways. A random selection of 3 from 7 sets can be made in $\binom{7}{3}$ ways. Therefore,

$$f(x) = \frac{\binom{2}{x}\binom{5}{3-x}}{\binom{7}{3}}, \qquad x = 0, 1, 2.$$

In tabular form

x	0	1	2
$f(x)$	2/7	4/7	1/7

The following is a probability histogram:



2.10 (a) $P(T = 5) = F(5) - F(4) = 3/4 - 1/2 = 1/4$.

(b) $P(T > 3) = 1 - F(3) = 1 - 1/2 = 1/2$.

(c) $P(1.4 < T < 6) = F(6) - F(1.4) = 3/4 - 1/4 = 1/2$.

(d) $P(T \le 5 | T \ge 2) = \frac{P(2 \le T \le 5)}{P(T \ge 2)} = \frac{3/4 - 1/4}{1 - 1/4} = \frac{2}{3}$.

2.11 The c.d.f. of X is

$$F(x) = \begin{cases} 0, & \text{for } x < 0, \\ 0.41, & \text{for } 0 \le x < 1, \\ 0.78, & \text{for } 1 \le x < 2, \\ 0.94, & \text{for } 2 \le x < 3, \\ 0.99, & \text{for } 3 \le x < 4, \\ 1, & \text{for } x \ge 4. \end{cases}$$

2.12 (a) $P(X < 0.2) = F(0.2) = 1 - e^{-1.6} = 0.7981$;

(b) $f(x) = F'(x) = 8e^{-8x}$. Therefore, $P(X < 0.2) = 8 \int_0^{0.2} e^{-8x}\, dx = -e^{-8x}\Big|_0^{0.2} = 0.7981$.

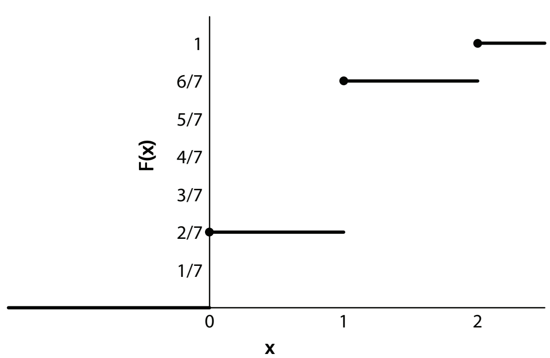
2.13 The c.d.f. of X is

$$F(x) = \begin{cases} 0, & \text{for } x < 0, \\ 2/7, & \text{for } 0 \le x < 1, \\ 6/7, & \text{for } 1 \le x < 2, \\ 1, & \text{for } x \ge 2. \end{cases}$$

(a) $P(X = 1) = P(X \le 1) - P(X \le 0) = 6/7 - 2/7 = 4/7$;

(b) $P(0 < X \le 2) = P(X \le 2) - P(X \le 0) = 1 - 2/7 = 5/7$.

2.14 A graph of the c.d.f. is shown next.



2.15 (a) $1 = k \int_0^1 \sqrt{x}\, dx = \frac{2k}{3} x^{3/2}\Big|_0^1 = \frac{2k}{3}$. Therefore, $k = \frac{3}{2}$.

(b) For $0 \leq x < 1$, $F(x) = \frac{3}{2} \int_0^x \sqrt{t}\, dt = t^{3/2} \Big|_0^x = x^{3/2}$.
Hence,

$$F(x) = \begin{cases} 0, & x < 0 \\ x^{3/2}, & 0 \leq x < 1 \\ 1, & x \geq 1 \end{cases}$$

$P(0.3 < X < 0.6) = F(0.6) - F(0.3) = (0.6)^{3/2} - (0.3)^{3/2} = 0.3004$.

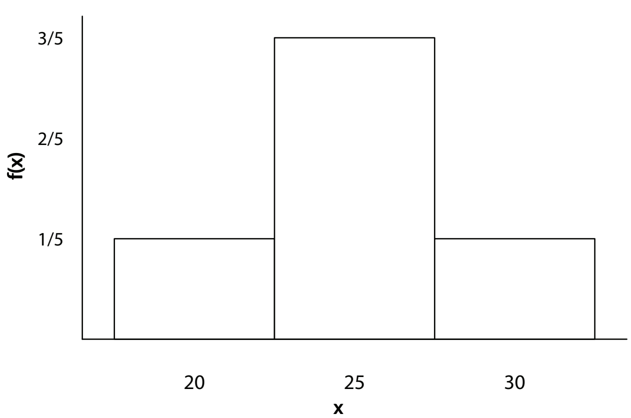
2.16 Denote by X the number of spades int he three draws. Let S and N stand for a spade and not a spade, respectively. Then
$P(X = 0) = P(NNN) = (39/52)(38/51)(37/50) = 703/1700$,
$P(X = 1) = P(SNN) + P(NSN) + P(NNS) = 3(13/52)(39/51)(38/50) = 741/1700$,
$P(X = 3) = P(SSS) = (13/52)(12/51)(11/50) = 11/850$, and
$P(X = 2) = 1 - 703/1700 - 741/1700 - 11/850 = 117/850$.
The probability mass function for X is then

x	0	1	2	3
$f(x)$	703/1700	741/1700	117/850	11/850

2.17 Let T be the total value of the three coins. Let D and N stand for a dime and nickel, respectively. Since we are selecting without replacement, the sample space containing elements for which $t = 20, 25,$ and 30 cents corresponding to the selecting of 2 nickels and 1 dime, 1 nickel and 2 dimes, and 3 dimes. Therefore, $P(T = 20) = \frac{\binom{2}{2}\binom{4}{1}}{\binom{6}{3}} = \frac{1}{5}$,
$P(T = 25) = \frac{\binom{2}{1}\binom{4}{2}}{\binom{6}{3}} = \frac{3}{5}$,
$P(T = 30) = \frac{\binom{4}{3}}{\binom{6}{3}} = \frac{1}{5}$,
and the probability distribution in tabular form is

t	20	25	30
$P(T = t)$	1/5	3/5	1/5

As a probability histogram



2.18 There are $\binom{10}{4}$ ways of selecting any 4 CDs from 10. We can select x jazz CDs from 5 and $4-x$ from the remaining CDs in $\binom{5}{x}\binom{5}{4-x}$ ways. Hence

$$f(x) = \frac{\binom{5}{x}\binom{5}{4-x}}{\binom{10}{4}}, \qquad x = 0, 1, 2, 3, 4.$$

2.19 (a) For $x \geq 0$, $F(x) = \int_0^x \frac{1}{2000} \exp(-t/2000)\, dt = -\exp(-t/2000)|_0^x$ $= 1 - \exp(-x/2000)$. So

$$F(x) = \begin{cases} 0, & x < 0, \\ 1 - \exp(-x/2000), & x \geq 0. \end{cases}$$

(b) $P(X > 1000) = 1 - F(1000) = 1 - [1 - \exp(-1000/2000)] = 0.6065$.

(c) $P(X < 2000) = F(2000) = 1 - \exp(-2000/2000) = 0.6321$.

2.20 (a) $f(x) \geq 0$ and $\int_{23.75}^{26.25} \frac{2}{5}\, dx = \frac{2}{5}t\Big|_{23.75}^{26.25} = \frac{2.5}{2.5} = 1$.

(b) $P(X < 24) = \int_{23.75}^{24} \frac{2}{5}\, dx = \frac{2}{5}(24 - 23.75) = 0.1$.

(c) $P(X > 26) = \int_{26}^{26.25} \frac{2}{5}\, dx = \frac{2}{5}(26.25 - 26) = 0.1$. It is not extremely rare.

2.21 (a) $f(x) \geq 0$ and $\int_1^\infty 3x^{-4}\, dx = -3\, \frac{x^{-3}}{3}\Big|_1^\infty = 1$. So, this is a valid density function.

(b) For $x \geq 1$, $F(x) = \int_1^x 3t^{-4}\, dt = 1 - x^{-3}$. So,

$$F(x) = \begin{cases} 0, & x < 1, \\ 1 - x^{-3}, & x \geq 1. \end{cases}$$

(c) $P(X > 4) = 1 - F(4) = 4^{-3} = 0.0156$.

2.22 (a) $1 = k \int_{-1}^1 (3 - x^2)\, dx = k \left(3x - \frac{x^3}{3}\right)\Big|_{-1}^1 = \frac{16}{3}k$. So, $k = \frac{3}{16}$.

(b) For $-1 \leq x < 1$, $F(x) = \frac{3}{16} \int_{-1}^x (3 - t^2)\, dt = \left(3t - \frac{1}{3}t^3\right)\Big|_{-1}^x = \frac{1}{2} + \frac{9}{16}x - \frac{x^3}{16}$. So, $P\left(X < \frac{1}{2}\right) = \frac{1}{2} - \left(\frac{9}{16}\right)\left(\frac{1}{2}\right) - \frac{1}{16}\left(\frac{1}{2}\right)^3 = \frac{99}{128}$.

(c) $P(|X| < 0.8) = P(X < -0.8) + P(X > 0.8) = F(-0.8) + 1 - F(0.8)$ $= 1 + \left(\frac{1}{2} - \frac{9}{16}0.8 + \frac{1}{16}0.8^3\right) - \left(\frac{1}{2} + \frac{9}{16}0.8 - \frac{1}{16}0.8^3\right) = 0.164$.

2.23 (a) For $y \geq 0$, $F(y) = \frac{1}{4} \int_0^y e^{-t/4}\, dy = 1 - e^{y/4}$. So, $P(Y > 6) = e^{-6/4} = 0.2231$. This probability certainly cannot be considered as "unlikely."

(b) $P(Y \leq 1) = 1 - e^{-1/4} = 0.2212$, which is not so small either.

2.24 (a) $f(y) \geq 0$ and $\int_0^1 5(1 - y)^4\, dy = -(1 - y)^5|_0^1 = 1$. So, this is a valid density function.

(b) $P(Y < 0.1) = -\,(1-y)^5|_0^{0.1} = 1-(1-0.1)^5 = 0.4095.$

(c) $P(Y > 0.5) = (1-0.5)^5 = 0.03125.$

2.25 (a) Using integral by parts and setting $1 = k\int_0^1 y^4(1-y)^3\,dy$, we obtain $k = 280$.

(b) For $0 \le y < 1$, $F(y) = 56y^5(1-y)^3 + 28y^6(1-y)^2 + 8y^7(1-y) + y^8$. So, $P(Y \le 0.5) = 0.3633$.

(c) Using the cdf in (b), $P(Y > 0.8) = 0.0563$.

2.26 (a) The event $Y = y$ means that among 5 selected, exactly y tubes meet the specification (M) and $5-y$ (M') does not. The probability for one combination of such a situation is $(0.99)^y(1-0.99)^{5-y}$ if we assume independence among the tubes. Since there are $\frac{5!}{y!(5-y)!}$ permutations of getting y Ms and $5-y$ M's, the probability of this event ($Y = y$) would be what it is specified in the problem.

(b) Three out of 5 is outside of specification means that $Y = 2$. $P(Y = 2) = 9.8\times 10^{-6}$ which is extremely small. So, the conjecture is false.

2.27 (a) $P(X > 8) = 1 - P(X \le 8) = \sum\limits_{x=0}^{8} e^{-6}\frac{6^x}{x!} = 1 - e^{-6}\left(\frac{6^0}{0!} + \frac{6^1}{1!} + \cdots + \frac{6^8}{8!}\right) = 0.1528.$

(b) $P(X = 2) = e^{-6}\frac{6^2}{2!} = 0.0446.$

2.28 For $0 < x < 1$, $F(x) = 2\int_0^x (1-t)\,dt = -\,(1-t)^2|_0^x = 1-(1-x)^2$.

(a) $P(X \le 1/3) = 1-(1-1/3)^2 = 5/9.$

(b) $P(X > 0.5) = (1-1/2)^2 = 1/4.$

(c) $P(X < 0.75 \mid X \ge 0.5) = \frac{P(0.5 \le X < 0.75)}{P(X \ge 0.5)} = \frac{(1-0.5)^2-(1-0.75)^2}{(1-0.5)^2} = \frac{3}{4}.$

2.29 (a) $\sum\limits_{x=0}^{3}\sum\limits_{y=0}^{3} f(x,y) = c\sum\limits_{x=0}^{3}\sum\limits_{y=0}^{3} xy = 36c = 1$. Hence $c = 1/36$.

(b) $\sum\limits_{x}\sum\limits_{y} f(x,y) = c\sum\limits_{x}\sum\limits_{y} |x-y| = 15c = 1$. Hence $c = 1/15$.

2.30 The joint probability distribution of (X,Y) is

	$f(x,y)$	x = 0	1	2	3
y	0	0	1/30	2/30	3/30
	1	1/30	2/30	3/30	4/30
	2	2/30	3/30	4/30	5/30

(a) $P(X \le 2, Y = 1) = f(0,1) + f(1,1) + f(2,1) = 1/30 + 2/30 + 3/30 = 1/5.$

(b) $P(X > 2, Y \le 1) = f(3,0) + f(3,1) = 3/30 + 4/30 = 7/30.$

(c) $P(X > Y) = f(1,0) + f(2,0) + f(3,0) + f(2,1) + f(3,1) + f(3,2)$
$= 1/30 + 2/30 + 3/30 + 3/30 + 4/30 + 5/30 = 3/5.$

(d) $P(X+Y=4) = f(2,2) + f(3,1) = 4/30 + 4/30 = 4/15$.

(e) The possible outcomes of X are 0, 1, 2, and 3, and the possible outcomes of Y are 0, 1, and 2. The marginal distribution of X can be calculated such as $f_X(0) = 1/30 + 2/30 = 1/10$. Finally, we have the distribution tables.

x	0	1	2	3
$f_X(x)$	1/10	1/5	3/10	4/10

y	0	1	2
$f_Y(y)$	1/5	1/3	7/15

2.31 (a) We can select x oranges from 3, y apples from 2, and $4-x-y$ bananas from 3 in $\binom{3}{x}\binom{2}{y}\binom{3}{4-x-y}$ ways. A random selection of 4 pieces of fruit can be made in $\binom{8}{4}$ ways. Therefore,

$$f(x,y) = \frac{\binom{3}{x}\binom{2}{y}\binom{3}{4-x-y}}{\binom{8}{4}}, \qquad x = 0,1,2,3; \quad y = 0,1,2; \quad 1 \le x+y \le 4.$$

Hence, we have the following joint probability table with the marginal distributions on the last row and last column.

$f(x,y)$		$x=0$	$x=1$	$x=2$	$x=3$	$f_Y(y)$
y	0	0	3/70	9/70	3/70	3/14
	1	2/70	18/70	18/70	2/70	8/14
	2	3/70	9/70	3/70	0	3/14
$f_X(x)$		1/14	6/14	6/14	1/14	

(b) $P[(X,Y) \in A] = P(X+Y \le 2) = f(1,0) + f(2,0) + f(0,1) + f(1,1) + f(0,2)$
$= 3/70 + 9/70 + 2/70 + 18/70 + 3/70 = 1/2$.

(c) $P(Y=0|X=2) = \dfrac{P(X=2, Y=0)}{P(X=2)} = \dfrac{9/70}{6/14} = \dfrac{3}{10}$.

(d) We know from (c) that $P(Y=0|X=2) = 3/10$, and we can calculate

$$P(Y=1|X=2) = \frac{18/70}{6/14} = \frac{3}{5}, \text{ and } P(Y=2|X=2) = \frac{3/70}{6/14} = \frac{1}{10}.$$

2.32 (a) $g(x) = \frac{2}{3}\int_0^1 (x+2y)\,dy = \frac{2}{3}(x+1)$, for $0 \le x \le 1$.

(b) $h(y) = \frac{2}{3}\int_0^1 (x+2y)\,dx = \frac{1}{3}(1+4y)$, for $0 \le y \le 1$.

(c) $P(X < 1/2) = \frac{2}{3}\int_0^{1/2} (x+1)\,dx = \frac{5}{12}$.

2.33 (a) $P(X+Y \le 1/2) = \int_0^{1/2}\int_0^{1/2-y} 24xy\,dx\,dy = 12\int_0^{1/2}\left(\frac{1}{2}-y\right)^2 y\,dy = \frac{1}{16}$.

(b) $g(x) = \int_0^{1-x} 24xy\,dy = 12x(1-x)^2$, for $0 \le x < 1$.

(c) $f(y|x) = \frac{24xy}{12x(1-x)^2} = \frac{2y}{(1-x)^2}$, for $0 \le y \le 1-x$.
Therefore, $P(Y < 1/8 \mid X = 3/4) = 32\int_0^{1/8} y\,dy = 1/4$.

2.34 Since $h(y) = e^{-y}\int_0^{\infty} e^{-x}\ dx = e^{-y}$, for $y > 0$, then $f(x|y) = f(x,y)/h(y) = e^{-x}$, for $x > 0$. So, $P(0 < X < 1 \mid Y = 2) = \int_0^1 e^{-x}\ dx = 0.6321$.

2.35 (a) $P(0 \le X \le 1/2,\ 1/4 \le Y \le 1/2) = \int_0^{1/2}\int_{1/4}^{1/2} 4xy\ dy\ dx = 3/8\int_0^{1/2} x\ dx = 3/64$.

(b) $P(X < Y) = \int_0^1\int_0^y 4xy\ dx\ dy = 2\int_0^1 y^3\ dy = 1/2$.

2.36 (a) $1 = k\int_{30}^{50}\int_{30}^{50}(x^2+y^2)\ dx\ dy = k(50-30)\left(\int_{30}^{50} x^2\ dx + \int_{30}^{50} y^2\ dy\right) = \frac{392k}{3}\cdot 10^4$.
So, $k = \frac{3}{392}\cdot 10^{-4}$.

(b) $P(30 \le X \le 40,\ 40 \le Y \le 50) = \frac{3}{392}\cdot 10^{-4}\int_{30}^{40}\int_{40}^{50}(x^2+y^2)\ dy\ dx$
$= \frac{3}{392}\cdot 10^{-3}(\int_{30}^{40} x^2\ dx + \int_{40}^{50} y^2\ dy) = \frac{3}{392}\cdot 10^{-3}\left(\frac{40^3-30^3}{3} + \frac{50^3-40^3}{3}\right) = \frac{49}{196}$.

(c) $P(30 \le X \le 40,\ 30 \le Y \le 40) = \frac{3}{392}\cdot 10^{-4}\int_{30}^{40}\int_{30}^{40}(x^2+y^2)\ dx\ dy$
$= 2\frac{3}{392}\cdot 10^{-4}(40-30)\int_{30}^{40} x^2\ dx = \frac{3}{196}\cdot 10^{-3}\frac{40^3-30^3}{3} = \frac{37}{196}$.

2.37 $P(X+Y > 1/2) = 1 - P(X+Y < 1/2) = 1 - \int_0^{1/4}\int_x^{1/2-x}\frac{1}{y}\ dy\ dx$
$= 1 - \int_0^{1/4}\left[\ln\left(\frac{1}{2}-x\right) - \ln x\right]\ dx = 1 + \left[\left(\frac{1}{2}-x\right)\ln\left(\frac{1}{2}-x\right) - x\ln x\right]\Big|_0^{1/4}$
$= 1 + \frac{1}{4}\ln\left(\frac{1}{4}\right) = 0.6534$.

The upper limit of y-integral comes from $x + y < 1/2$ and $0 < x < y$.

2.38 (a) $g(x) = 2\int_x^1\ dy = 2(1-x)$ for $0 < x < 1$;
$h(y) = 2\int_0^y\ dx = 2y$, for $0 < y < 1$.
Since $f(x,y) \ne g(x)h(y)$, X and Y are not independent.

(b) $f(x|y) = f(x,y)/h(y) = 1/y$, for $0 < x < y$.
Therefore, $P(1/4 < X < 1/2 \mid Y = 3/4) = \frac{4}{3}\int_{1/4}^{1/2} dx = \frac{1}{3}$.

2.39 (a)

x	1	2	3
$g(x)$	0.10	0.35	0.55

(b)

y	1	3	5
$h(y)$	0.20	0.50	0.30

(c) $P(Y = 3 \mid X = 2) = \frac{0.1}{0.05+0.10+0.20} = 0.2857$.

2.40

	$f(x,y)$	x = 2	4	$h(y)$
y	1	0.10	0.15	0.25
	3	0.20	0.30	0.50
	5	0.10	0.15	0.25
	$g(x)$	0.40	0.60	

(a)

x	2	4
$g(x)$	0.40	0.60

(b)

y	1	3	5
$h(y)$	0.25	0.50	0.25

2.41 $g(x) = \frac{1}{8}\int_2^4 (6-x-y)\,dy = \frac{3-x}{4}$, for $0 < x < 2$.
So, $f(y|x) = \frac{f(x,y)}{g(x)} = \frac{6-x-y}{2(3-x)}$, for $2 < y < 4$,
and $P(1 < Y < 3 \mid X = 1) = \frac{1}{4}\int_2^3 (5-y)\,dy = \frac{5}{8}$.

2.42 (a) $P(H) = 0.4$, $P(T) = 0.6$, and $S = \{HH, HT, TH, TT\}$. Let (W, Z) represent a typical outcome of the experiment. The particular outcome $(1, 0)$ indicating a total of 1 head and no heads on the first toss corresponds to the event TH. Therefore, $f(1, 0) = P(W = 1, Z = 0) = P(TH) = P(T)P(H) = (0.6)(0.4) = 0.24$. Similar calculations for the outcomes $(0, 0)$, $(1, 1)$, and $(2, 1)$ lead to the following joint probability distribution:

		w		
	$f(w,z)$	0	1	2
z	0	0.36	0.24	
	1		0.24	0.16

(b) Summing the columns, the marginal distribution of W is

w	0	1	2
$g(w)$	0.36	0.48	0.16

(c) Summing the rows, the marginal distribution of Z is

z	0	1
$h(z)$	0.60	0.40

(d) $P(W \geq 1) = f(1, 0) + f(1, 1) + f(2, 1) = 0.24 + 0.24 + 0.16 = 0.64$.

2.43 X and Y are independent since $f(x, y) = g(x)h(y)$ for all (x, y).

2.44 Since $f(1, 1) \neq g(1)h(1)$, the variables are not independent.

2.45 (a) $1 = k\int_0^2\int_0^1\int_0^1 xy^2z\,dx\,dy\,dz = 2k\int_0^1\int_0^1 y^2z\,dy\,dz = \frac{2k}{3}\int_0^1 z\,dz = \frac{k}{3}$. So, $k = 3$.

(b) $P\left(X < \frac{1}{4}, Y > \frac{1}{2}, 1 < Z < 2\right) = 3\int_1^2\int_{1/2}^1\int_0^{1/4} xy^2z\,dx\,dy\,dz = \frac{9}{2}\int_0^{1/4}\int_{1/2}^1 y^2z\,dy\,dz$
$= \frac{21}{16}\int_1^2 z\,dz = \frac{21}{512}$.

2.46 (a) $h(y) = 6\int_0^{1-y} x\,dx = 3(1-y)^2$, for $0 < y < 1$. Since $f(x|y) = \frac{f(x,y)}{h(y)} = \frac{2x}{(1-y)^2}$, for $0 < x < 1-y$, involves the variable y, X and Y are not independent.

(b) $P(X > 0.3 \mid Y = 0.5) = 8\int_{0.3}^{0.5} x\,dx = 0.64$.

2.47 $g(x) = 4\int_0^1 xy\,dy = 2x$, for $0 < x < 1$; $h(y) = 4\int_0^1 xy\,dx = 2y$, for $0 < y < 1$. Since $f(x, y) = g(x)h(y)$ for all (x, y), X and Y are independent.

2.48 (a) $g(y,z) = \frac{4}{9}\int_0^1 xyz^2\,dx = \frac{2}{9}yz^2$, for $0 < y < 1$ and $0 < z < 3$.

(b) $h(y) = \frac{2}{9}\int_0^3 yz^2\,dz = 2y$, for $0 < y < 1$.

(c) $P\left(\frac{1}{4} < X < \frac{1}{2}, Y > \frac{1}{3}, Z < 2\right) = \frac{4}{9}\int_1^2\int_{1/3}^1\int_{1/4}^{1/2} xyz^2\,dx\,dy\,dz = \frac{7}{162}$.

(d) Since $f(x|y,z) = \frac{f(x,y,z)}{g(y,z)} = 2x$, for $0 < x < 1$, $P\left(0 < X < \frac{1}{2} \mid Y = \frac{1}{4}, Z = 2\right) = 2\int_0^{1/2} x\,dx = \frac{1}{4}$.

2.49 $g(x) = k\int_{30}^{50}(x^2+y^2)\,dy = k\left(x^2y + \frac{y^3}{3}\right)\Big|_{30}^{50} = k\left(20x^2 + \frac{98{,}000}{3}\right)$, and
$h(y) = k\left(20y^2 + \frac{98{,}000}{3}\right)$.
Since $f(x,y) \neq g(x)h(y)$, X and Y are not independent.

2.50 $E(X) = \sum_{x=0}^{3} x\ f(x) = (0)(27/64) + (1)(27/64) + (2)(9/64) + (3)(1/64) = 3/4$.

2.51 $\mu = E(X) = (0)(0.41) + (1)(0.37) + (2)(0.16) + (3)(0.05) + (4)(0.01) = 0.88$.

2.52 Assigning wrights of $3w$ and w for a head and tail, respectively. We obtain $P(H) = 3/4$ and $P(T) = 1/4$. The sample space for the experiment is $S = \{HH, HT, TH, TT\}$. Now if X represents the number of tails that occur in two tosses of the coin, we have

$$P(X = 0) = P(HH) = (3/4)(3/4) = 9/16,$$
$$P(X = 1) = P(HT) + P(TH) = (2)(3/4)(1/4) = 3/8,$$
$$P(X = 2) = P(TT) = (1/4)(1/4) = 1/16.$$

The probability distribution for X is then

x	0	1	2
$f(x)$	9/16	3/8	1/16

from which we get $\mu = E(X) = (0)(9/16) + (1)(3/8) + (2)(1/16) = 1/2$.

2.53 $\mu = E(X) = (20)(1/5) + (25)(3/5) + (30)(1/5) = 25$ cents.

2.54 Let c = amount to play the game and Y = amount won.

y	$5-c$	$3-c$	$-c$
$f(y)$	2/13	2/13	9/13

$E(Y) = (5-c)(2/13) + (3-c)(2/13) + (-c)(9/13) = 0$. So, $13c = 16$ which implies $c = \$1.23$.

2.55 Expected gain $= E(X) = (4000)(0.3) + (-1000)(0.7) = \$500$.

2.56 Let X = profit. Then

$$\mu = E(X) = (250)(0.22) + (150)(0.36) + (0)(0.28) + (-150)(0.14) = \$88.$$

2.57 $E(X) = \frac{4}{\pi}\int_0^1 \frac{x}{1+x^2}\ dx = \frac{\ln 4}{\pi}$.

2.58 $\mu_X = \sum xg(x) = (1)(0.17) + (2)(0.5) + (3)(0.33) = 2.16$,
$\mu_Y = \sum yh(y) = (1)(0.23) + (2)(0.5) + (3)(0.27) = 2.04$.

2.59 $E(X) = \int_0^1 x^2\ dx + \int_1^2 x(2-x)\ dx = 1$. Therefore, the average number of hours per year is $(1)(100) = 100$ hours.

2.60 $E(X) = \int_0^1 2x(1-x)\ dx = 1/3$. So, $(1/3)(\$5,000) = \$1,667.67$.

2.61 $E(X) = \frac{1}{\pi a^2}\int_{-a}^{a}\int_{-\sqrt{a^2-y^2}}^{\sqrt{a^2-y^2}} x\ dx\ dy = \frac{1}{\pi a^2}\int_{-a}^{a}\left[\left(\frac{a^2-y^2}{2}\right) - \left(\frac{a^2-y^2}{2}\right)\right]\ dy = 0.$

2.62 $E(X) = \int_0^1 \frac{2x(x+2)}{5}\ dx = \frac{8}{15}$.

2.63 The probability density function is,

x	-3	6	9
$f(x)$	1/6	1/2	1/3
$g(x)$	25	169	361

$\mu_{g(X)} = E[(2X+1)^2] = (25)(1/6) + (169)(1/2) + (361)(1/3) = 209.$

2.64 $P(X_1 + X_2 \geq 1) = 1 - P(X_1 = 0, X_2 = 0)$
$= 1 - \frac{\binom{980}{2}\binom{20}{0}}{\binom{1000}{2}} = 1 - 0.9604 = 0.040.$

2.65 Let $Y = 1200X - 50X^2$ be the amount spent.

x	0	1	2	3
$f(x)$	1/10	3/10	2/5	1/5
$y = g(x)$	0	1150	2200	3150

$\mu_Y = E(1200X - 50X^2) = (0)(1/10) + (1150)(3/10) + (2200)(2/5) + (3150)(1/5)$
$= \$1,855.$

2.66 $E(Y) = E(X+4) = \int_0^\infty 32(x+4)\frac{1}{(x+4)^3}\ dx = 8$ days.

2.67 (a) $E[g(X,Y)] = E(XY^2) = \sum_x\sum_y xy^2 f(x,y)$
$= (2)(1)^2(0.10) + (2)(3)^2(0.20) + (2)(5)^2(0.10) + (4)(1)^2(0.15) + (4)(3)^2(0.30)$
$+ (4)(5)^2(0.15) = 35.2.$

(b) $\mu_X = E(X) = (2)(0.40) + (4)(0.60) = 3.20$,
$\mu_Y = E(Y) = (1)(0.25) + (3)(0.50) + (5)(0.25) = 3.00.$

2.68 (a) $E(X^2Y - 2XY) = \sum_{x=0}^{3}\sum_{y=0}^{2}(x^2y - 2xy)f(x,y) = (1-2)(18/70) + (4-4)(18/70) +$
$\cdots + (8-8)(3/70) = -3/7.$

(b)

x	0	1	2	3
$g(x)$	5/70	30/70	30/70	5/70

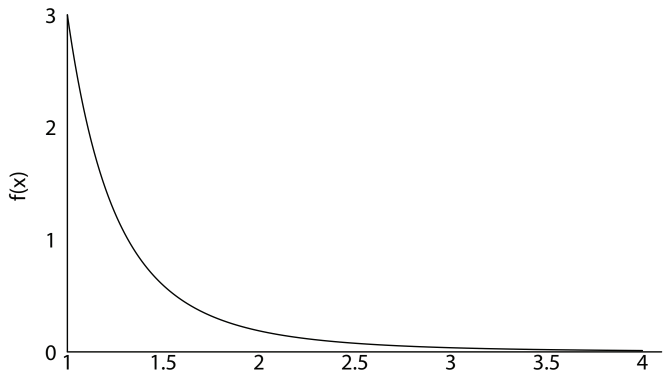
y	0	1	2
$h(y)$	15/70	40/70	15/70

$\mu_X = E(X) = (0)(5/70) + (1)(30/70) + (2)(30/70) + (3)(5/70) = 3/2,$
$\mu_Y = E(Y) = (0)(15/70) + (1)(40/70) + (2)(15/70) = 1.$
Hence $\mu_X - \mu_Y = 3/2 - 1 = 1/2$.

2.69 $E(X) = \frac{1}{2000}\int_0^\infty x\exp(-x/2000)\ dx = 2000\int_0^\infty y\exp(-y)\ dy = 2000.$

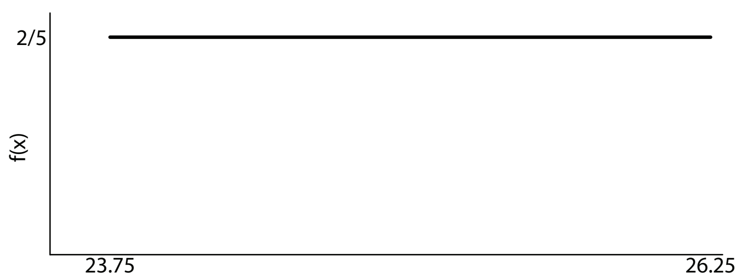
2.70 $E(Z) = E(\sqrt{X^2+Y^2}) = \int_0^1\int_0^1 4xy\sqrt{x^2+y^2}\ dx\ dy = \frac{4}{3}\int_0^1 [y(1+y^2)^{3/2} - y^4]\ dy$
$= 8(2^{3/2}-1)/15 = 0.9752.$

2.71 (a) The density function is shown next



(b) $\mu = E(X) = \int_1^\infty 3x^{-3}\ dx = \frac{3}{2}$.

2.72 (a) The density function is shown next.



(b) $E(X) = \frac{2}{5}\int_{23.75}^{26.25} x\ dx = \frac{1}{5}(26.25^2 - 23.75^2) = 25.$

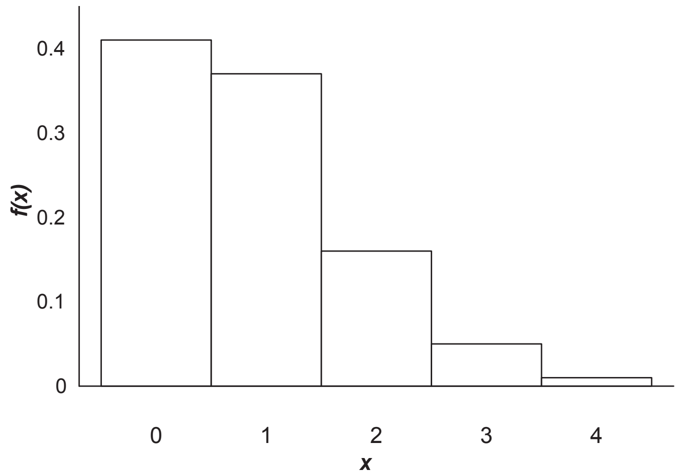
(c) The mean is exactly in the middle of the interval. This should not be surprised due to the symmetry of the density at 25.

2.73 (a) $\mu = E(Y) = 5\int_0^1 y(1-y)^4\ dy = -\int_0^1 y\ d(1-y)^5 = \int_0^1 (1-y)^5\ dy = \frac{1}{6}$.

(b) $P(Y > 1/6) = \int_{1/6}^1 5(1-y)^4\ dy = -\ (1-y)^5\big|_{1/6}^1 = (1-1/6)^5 = 0.4019.$

2.74 $E(Y) = \frac{1}{4}\int_0^\infty ye^{-y/4}\ dy = 4.$

2.75 (a) A histogram is shown next.



(b) $\mu = (0)(0.41) + (1)(0.37) + (2)(0.16) + (3)(0.05) + (4)(0.01) = 0.88$.

(c) $E(X^2) = (0)^2(0.41) + (1)^2(0.37) + (2)^2(0.16) + (3)^2(0.05) + (4)^2(0.01) = 1.62$.

2.76 $\mu = \$500$. So, $\sigma^2 = E[(X-\mu)^2] = \sum_x (x-\mu)^2 f(x) = (-1500)^2(0.7) + (3500)^2(0.3) = \$5,250,000$.

2.77 $\mu = (2)(0.01) + (3)(0.25) + (4)(0.4) + (5)(0.3) + (6)(0.04) = 4.11$,
$E(X^2) = (2)^2(0.01) + (3)^2(0.25) + (4)^2(0.4) + (5)^2(0.3) + (6)^2(0.04) = 17.63$.
So, $\sigma^2 = 17.63 - 4.11^2 = 0.74$.

2.78 $\mu = (0)(0.4) + (1)(0.3) + (2)(0.2) + (3)(0.1) = 1.0$,
and $E(X^2) = (0)^2(0.4) + (1)^2(0.3) + (2)^2(0.2) + (3)^2(0.1) = 2.0$.
So, $\sigma^2 = 2.0 - 1.0^2 = 1.0$.

2.79 It is know $\mu = 1/3$.
So, $E(X^2) = \int_0^1 2x^2(1-x)\ dx = 1/6$ and $\sigma^2 = 1/6 - (1/3)^2 = 1/18$. So, in the actual profit, the variance is $\frac{1}{18}(5000)^2$.

2.80 It is known $\mu = 8/15$.
Since $E(X^2) = \int_0^1 \frac{2}{5}x^2(x+2)\ dx = \frac{11}{30}$, then $\sigma^2 = 11/30 - (8/15)^2 = 37/450$.

2.81 It is known $\mu = 1$.
Since $E(X^2) = \int_0^1 x^2 \cdot x\ dx + \int_1^2 x^2(2-x)\ dx = 7/6$, then $\sigma^2 = 7/6 - (1)^2 = 1/6$.

2.82 $\mu_{g(X)} = E[g(X)] = \int_0^1 (3x^2+4)\left(\frac{2x+4}{5}\right)\ dx = \frac{1}{5}\int_0^1 (6x^3 + 12x^2 + 8x + 16)\ dx = 5.1$.
So, $\sigma^2 = E[g(X) - \mu]^2 = \int_0^1 (3x^2 + 4 - 5.1)^2\left(\frac{2x+4}{5}\right)\ dx$
$= \int_0^1 (9x^4 - 6.6x^2 + 1.21)\left(\frac{2x+4}{5}\right)\ dx = 0.83$.

2.83 $\mu_Y = E(3X-2) = \frac{1}{4}\int_0^\infty (3x-2)e^{-x/4}\ dx = 10$. So
$\sigma_Y^2 = E\{[(3X-2) - 10]^2\} = \frac{9}{4}\int_0^\infty (x-4)^2 e^{-x/4}\ dx = 144$.

2.84 From previous exercise, $k = \left(\frac{3}{392}\right) 10^{-4}$, and $g(x) = k\left(20x^2 + \frac{98000}{3}\right)$, with
$\mu_X = E(X) = \int_{30}^{50} xg(x)\,dx = k\int_{30}^{50}\left(20x^3 + \frac{98000}{3}x\right)\,dx = 40.8163$.
Similarly, $\mu_Y = 40.8163$. On the other hand,
$E(XY) = k\int_{30}^{50}\int_{30}^{50} xy(x^2+y^2)\,dy\,dx = 1665.3061$.
Hence, $\sigma_{XY} = E(XY) - \mu_X\mu_Y = 1665.3061 - (40.8163)^2 = -0.6642$.

2.85 $g(x) = \frac{2}{3}\int_0^1 (x+2y)\,dy = \frac{2}{3}(x+1)$, for $0 < x < 1$, so $\mu_X = \frac{2}{3}\int_0^1 x(x+1)\,dx = \frac{5}{9}$;
$h(y) = \frac{2}{3}\int_0^1 (x+2y)\,dx = \frac{2}{3}\left(\frac{1}{2}+2y\right)$, so $\mu_Y = \frac{2}{3}\int_0^1 y\left(\frac{1}{2}+2y\right)\,dy = \frac{11}{18}$; and
$E(XY) = \frac{2}{3}\int_0^1\int_0^1 xy(x+2y)\,dy\,dx = \frac{1}{3}$.
So, $\sigma_{XY} = E(XY) - \mu_X\mu_Y = \frac{1}{3} - \left(\frac{5}{9}\right)\left(\frac{11}{18}\right) = -0.0062$.

2.86 Since $\sigma_{XY} = Cov(a+bX, X) = b\sigma_X^2$ and $\sigma_Y^2 = b^2\sigma_X^2$, then
$\rho = \frac{\sigma_{XY}}{\sigma_X}\sigma_Y = \frac{b\sigma_X^2}{\sqrt{\sigma_X^2 b^2\sigma_X^2}} = \frac{b}{|b|} = \text{sign of } b$.
Hence $\rho = 1$ if $b > 0$ and $\rho = -1$ if $b < 0$.

2.87 $E(X) = (0)(0.41) + (1)(0.37) + (2)(0.16) + (3)(0.05) + (4)(0.01) = 0.88$
and $E(X^2) = (0)^2(0.41) + (1)^2(0.37) + (2)^2(0.16) + (3)^2(0.05) + (4)^2(0.01) = 1.62$.
So, $Var(X) = 1.62 - 0.88^2 = 0.8456$ and $\sigma = \sqrt{0.8456} = 0.9196$.

2.88 $E(X) = 2\int_0^1 x(1-x)\,dx = 2\left(\frac{x^2}{2} - \frac{x^3}{3}\right)\Big|_0^1 = \frac{1}{3}$ and
$E(X^2) = 2\int_0^1 x^2(1-x)\,dx = 2\left(\frac{x^3}{3} - \frac{x^4}{4}\right)\Big|_0^1 = \frac{1}{6}$. Hence,
$Var(X) = \frac{1}{6} - \left(\frac{1}{3}\right)^2 = \frac{1}{18}$, and $\sigma = \sqrt{1/18} = 0.2357$.

2.89 The joint and marginal probability mass functions are given in the following table.

		y = 0	1	2	$f_X(x)$
x	0	0	1/35	3/70	1/14
	1	3/70	9/35	9/70	3/7
	2	9/70	9/35	3/70	3/7
	3	3/70	1/35	0	1/14
	$f_X(x)$	3/14	4/7	3/14	

Hence, $E(X) = \frac{3}{2}$, $E(Y) = 1$, $E(XY) = \frac{9}{7}$, $Var(X) = \frac{15}{28}$, and $Var(Y) = \frac{3}{7}$. Finally, $\rho_{XY} = -\frac{1}{\sqrt{5}}$.

2.90 Since $f_X(x) = 2(1-x)$, for $0 < x < 1$, and $f_Y(y) = 2y$, for $0 < y < 1$, we obtain $E(X) = \frac{1}{3}$, $E(Y) = \frac{2}{3}$, $Var(X) = Var(Y) = \frac{1}{18}$, and $E(XY) = \frac{1}{4}$. Hence, $\rho_{XY} = \frac{1}{2}$.

2.91 Let X = number of cartons sold and Y = profit.
We can write $Y = 1.65X + (0.90)(5-X) - (1.20)(5) = 0.75X - 1.50$. Now
$E(X) = (0)(1/15)+(1)(2/15)+(2)(2/15)+(3)(3/15)+(4)(4/15)+(5)(3/15) = 46/15$,
and $E(Y) = (0.75)E(X) - 1.50 = (0.75)(46/15) - 1.50 = \$0.80$.

2.92 $\mu_X = E(X) = \frac{1}{4}\int_0^\infty xe^{-x/4}\,dx = 4.$
Therefore, $\mu_Y = E(3X-2) = 3E(X) - 2 = (3)(4) - 2 = 10.$
Since $E(X^2) = \frac{1}{4}\int_0^\infty x^2 e^{-x/4}\,dx = 32$, therefore, $\sigma_X^2 = E(X^2) - \mu_X^2 = 32 - 16 = 16.$
Hence $\sigma_Y^2 = 9\sigma_X^2 = (9)(16) = 144.$

2.93 The equations $E[(X-1)^2] = 10$ and $E[(X-2)^2] = 6$ may be written in the form:

$$E(X^2) - 2E(X) = 9, \qquad E(X^2) - 4E(X) = 2.$$

Solving these two equations simultaneously we obtain

$$E(X) = 7/2, \quad \text{and} \quad E(X^2) = 16.$$

Hence $\mu = 7/2$ and $\sigma^2 = 16 - (7/2)^2 = 15/4.$

2.94 Since $E(X) = \int_0^1 x^2\,dx + \int_1^2 x(2-x)\,dx = 1$, and
$E(X^2) = \int_0^1 x^3\,dx + \int_1^2 x^2(2-x)\,dx = 7/6$, then

$$E(Y) = 60E(X^2) + 39E(X) = (60)(7/6) + (39)(1) = 109 \text{ kilowatt hours.}$$

2.95 $E(2XY^2 - X^2Y) = 2E(XY^2) - E(X^2Y)$. Now,
$E(XY^2) = \sum\limits_{x=0}^{2}\sum\limits_{y=0}^{2} xy^2 f(x,y) = (1)(1)^2(3/14) = 3/14$, and
$E(X^2Y) = \sum\limits_{x=0}^{2}\sum\limits_{y=0}^{2} x^2 y f(x,y) = (1)^2(1)(3/14) = 3/14.$
Therefore, $E(2XY^2 - X^2Y) = (2)(3/14) - (3/14) = 3/14.$

2.96 $\sigma_Z^2 = \sigma_{-2X+4Y-3}^2 = 4\sigma_X^2 + 16\sigma_Y^2 = (4)(5) + (16)(3) = 68$ since $\rho_{XY} = 0$.

2.97 $\sigma_Z^2 = \sigma_{-2X+4Y-3}^2 = 4\sigma_X^2 + 16\sigma_Y^2 - 16\sigma_{XY} = (4)(5) + (16)(3) - (16)(1) = 52.$

2.98 $E(Z) = E(XY) = E(X)E(Y) = \int_0^1\int_2^\infty 16xy(y/x^3)\,dx\,dy = 8/3.$

2.99 For $0 < a < 1$, since $g(a) = \sum\limits_{x=0}^{\infty} a^x = \frac{1}{1-a}$, $g'(a) = \sum\limits_{x=1}^{\infty} xa^{x-1} = \frac{1}{(1-a)^2}$ and
$g''(a) = \sum\limits_{x=2}^{\infty} x(x-1)a^{x-2} = \frac{2}{(1-a)^3}$. These results will be used to evaluate the sums below.

(a) The marginal distributions $g(x)$ and $h(y)$ are given in Review Exercise 2.118.
$E(X) = (3/4)\sum\limits_{x=1}^{\infty} x(1/4)^x = (3/4)(1/4)\sum\limits_{x=1}^{\infty} x(1/4)^{x-1} = (3/16)[1/(1-1/4)^2]$
$= 1/3$, and $E(Y) = E(X) = 1/3$.
$E(X^2) - E(X) = E[X(X-1)] = (3/4)\sum\limits_{x=2}^{\infty} x(x-1)(1/4)^x$
$= (3/4)(1/4)^2\sum\limits_{x=2}^{\infty} x(x-1)(1/4)^{x-2} = (3/4^3)[2/(1-1/4)^3] = 2/9.$
So, $Var(X) = E(X^2) - [E(X)]^2 = [E(X^2) - E(X)] + E(X) - [E(X)]^2$
$2/9 + 1/3 - (1/3)^2 = 4/9$, and $Var(Y) = 4/9$.

(b) $E(Z) = E(X) + E(Y) = (1/3) + (1/3) = 2/3$, and
$Var(Z) = Var(X+Y) = Var(X) + Var(Y) = (4/9) + (4/9) = 8/9$, since X and Y are independent (from Exercise 3.79).

2.100 (a) $g(x) = \frac{3}{2}\int_0^1 (x^2 + y^2)\ dy = \frac{1}{2}(3x^2 + 1)$ for $0 < x < 1$ and
$h(y) = \frac{1}{2}(3y^2 + 1)$ for $0 < y < 1$.
Since $f(x, y) \neq g(x)h(y)$, X and Y are not independent.

(b) $E(X + Y) = E(X) + E(Y) = 2E(X) = \int_0^1 x(3x^2 + 1)\ dx = 3/4 + 1/2 = 5/4$.
$E(XY) = \frac{3}{2}\int_0^1\int_0^1 xy(x^2 + y^2)\ dx\ dy = \frac{3}{2}\int_0^1 y\left(\frac{1}{4} + \frac{y^2}{2}\right)\ dy$
$= \frac{3}{2}\left[\left(\frac{1}{4}\right)\left(\frac{1}{2}\right) + \left(\frac{1}{2}\right)\left(\frac{1}{4}\right)\right] = \frac{3}{8}$.

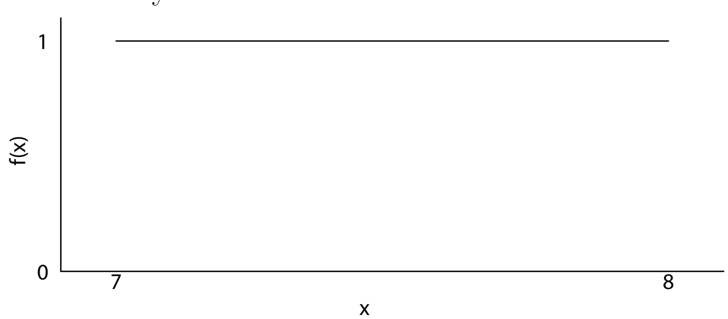
(c) $Var(X) = E(X^2) - [E(X)]^2 = \frac{1}{2}\int_0^1 x^2(3x^2 + 1)\ dx - \left(\frac{5}{8}\right)^2 = \frac{7}{15} - \frac{25}{64} = \frac{73}{960}$, and
$Var(Y) = \frac{73}{960}$. Also, $Cov(X, Y) = E(XY) - E(X)E(Y) = \frac{3}{8} - \left(\frac{5}{8}\right)^2 = -\frac{1}{64}$.

(d) $Var(X + Y) = Var(X) + Var(Y) + 2Cov(X, Y) = 2\frac{73}{960} - 2\frac{1}{64} = \frac{29}{240}$.

2.101 (a) $E(Y) = \int_0^\infty ye^{-y/4}\ dy = 4$.

(b) $E(Y^2) = \int_0^\infty y^2 e^{-y/4}\ dy = 32$ and $Var(Y) = 32 - 4^2 = 16$.

2.102 (a) The density function is shown next.



(b) $E(Y) = \int_7^8 y\ dy = \frac{1}{2}[8^2 - 7^2] = \frac{15}{2} = 7.5$,
$E(Y^2) = \int_7^8 y^2\ dy = \frac{1}{3}[8^3 - 7^3] = \frac{169}{3}$, and $Var(Y) = \frac{169}{3} - \left(\frac{15}{2}\right)^2 = \frac{1}{12}$.

2.103 $g(x) = 24\int_0^{1-x} xy\ dy = 12x(1 - x)^2$, for $0 < x < 1$.

(a) $P(X \geq 0.5) = 12\int_{0.5}^1 x(1 - x)^2\ dx = \int_{0.5}^1 (12x - 24x^2 + 12x^3)\ dx = \frac{5}{16} = 0.3125$.

(b) $h(y) = 24\int_0^{1-y} xy\ dx = 12y(1 - y)^2$, for $0 < y < 1$.

(c) $f(x|y) = \frac{f(x,y)}{h(y)} = \frac{24xy}{12y(1-y)^2} = \frac{2x}{(1-y)^2}$, for $0 < x < 1 - y$.
So, $P\left(X < \frac{1}{8} \mid Y = \frac{3}{4}\right) = \int_0^{1/8} \frac{2x}{1/16}\ dx = 32\int_0^{1/8} x\ dx = 0.25$.

2.104 (a)

x	1	3	5	7
$f(x)$	0.4	0.2	0.2	0.2

(b) $P(4 < X \leq 7) = P(X \leq 7) - P(X \leq 4) = F(7) - F(4) = 1 - 0.6 = 0.4$.

2.105 (a) $g(x) = \int_0^\infty ye^{-y(1+x)}\ dy = -\frac{1}{1+x} ye^{-y(1+x)}\Big|_0^\infty + \frac{1}{1+x}\int_0^\infty e^{-y(1+x)}\ dy$
$= -\frac{1}{(1+x)^2}\ e^{-y(1+x)}\Big|_0^\infty$
$= \frac{1}{(1+x)^2}$, for $x > 0$.
$h(y) = ye^{-y}\int_0^\infty e^{-yx}\ dx = -e^{-y}\ e^{-yx}|_0^\infty = e^{-y}$, for $y > 0$.

(b) $P(X \geq 2, Y \geq 2) = \int_2^\infty \int_2^\infty ye^{-y(1+x)}\ dx\ dy = -\int_2^\infty e^{-y}\ e^{-yx}|_2^\infty\ dy = \int_2^\infty e^{-3y} dy$
$= -\frac{1}{3}e^{-3y}\Big|_2^\infty = \frac{1}{3e^6}$.

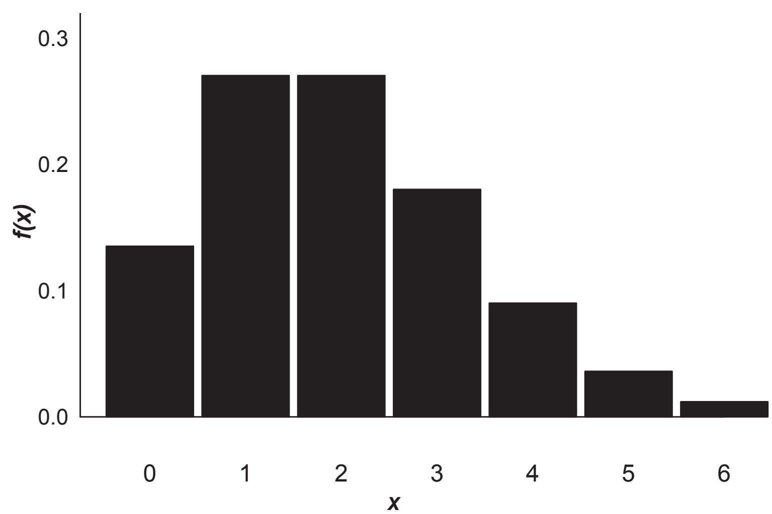
2.106 (a) $P\left(X \leq \frac{1}{2}, Y \leq \frac{1}{2}\right) = \frac{3}{2}\int_0^{1/2}\int_0^{1/2}(x^2+y^2)\ dydx = \frac{3}{2}\int_0^{1/2}\left(x^2 y + \frac{y^3}{3}\right)\Big|_0^{1/2}\ dx$
$= \frac{3}{4}\int_0^{1/2}\left(x^2 + \frac{1}{12}\right)\ dx = \frac{1}{16}$.

(b) $P\left(X \geq \frac{3}{4}\right) = \frac{3}{2}\int_{3/4}^1 \left(x^2 + \frac{1}{3}\right)\ dx = \frac{53}{128}$.

2.107 (a)

x	0	1	2	3	4	5	6
$f(x)$	0.1353	0.2707	0.2707	0.1804	0.0902	0.0361	0.0120

(b) A histogram is shown next.



(c)

x	0	1	2	3	4	5	6
$F(x)$	0.1353	0.4060	0.6767	0.8571	0.9473	0.9834	0.9954

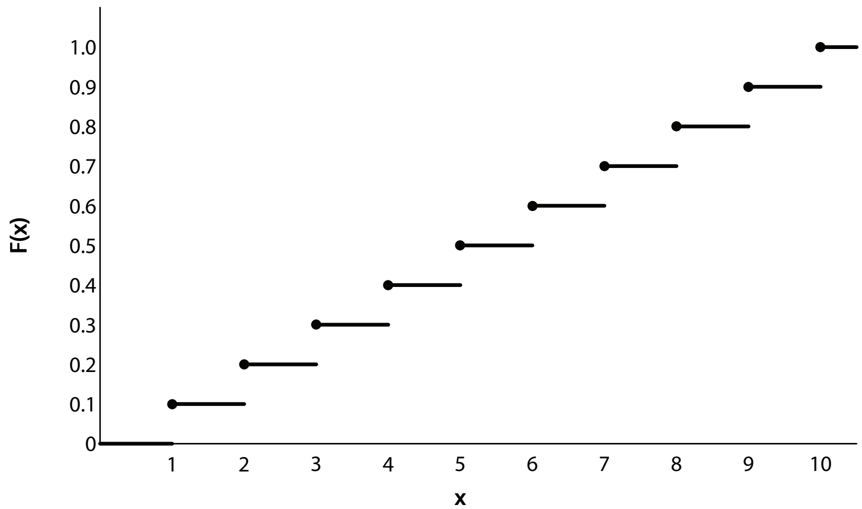
2.108 $f(x) = \binom{5}{x}(0.1)^x(1-0.1)^{5-x}$, for $x = 0, 1, 2, 3, 4, 5$.

2.109 (a) $f(x) = \frac{d}{dx}F(x) = \frac{1}{50}e^{-x/50}$, for $x > 0$.

(b) $P(X > 70) = 1 - P(X \leq 70) = 1 - F(70) = 1 - (1 - e^{-70/50}) = 0.2466$.

2.110 (a) $f(x) = \frac{1}{10}$, for $x = 1, 2, \ldots, 10$.

(b) A c.d.f. plot is shown next.



2.111 $P(X \geq 3) = \frac{1}{2}\int_3^\infty e^{-y/2} = e^{-3/2} = 0.2231.$

2.112 (a) $f(x) \geq 0$ and $\int_0^{10} \frac{1}{10}\, dx = 1$. This is a continuous uniform distribution.

(b) $P(X \leq 7) = \frac{1}{10}\int_0^7\, dx = 0.7.$

2.113 (a) $f(y) \geq 0$ and $\int_0^1 f(y)\, dy = 10\int_0^1 (1-y)^9\, dy = -\frac{10}{10}(1-y)^{10}\Big|_0^1 = 1.$

(b) $P(Y > 0.6) = \int_{0.6}^1 f(y)\, dy = -(1-y)^{10}\Big|_{0.6}^1 = (1-0.6)^{10} = 0.0001.$

2.114 (a) $P(Z > 20) = \frac{1}{10}\int_{20}^\infty e^{-z/10}\, dz = -e^{-z/10}\Big|_{20}^\infty = e^{-20/10} = 0.1353.$

(b) $P(Z \leq 10) = -e^{-z/10}\Big|_0^{10} = 1 - e^{-10/10} = 0.6321.$

2.115 (a) $g(x_1) = \int_{x_1}^1 2\, dx_2 = 2(1-x_1)$, for $0 < x_1 < 1$.

(b) $h(x_2) = \int_0^{x_2} 2\, dx_1 = 2x_2$, for $0 < x_2 < 1$.

(c) $P(X_1 < 0.2, X_2 > 0,5) = \int_{0.5}^1\int_0^{0.2} 2\, dx_1\, dx_2 = 2(1-0.5)(0.2-0) = 0.2.$

(d) $f_{X_1|X_2}(x_1|x_2) = \frac{f(x_1,x_2)}{h(x_2)} = \frac{2}{2x_2} = \frac{1}{x_2}$, for $0 < x_1 < x_2$.

2.116 (a) $f_{X_1}(x_1) = \int_0^{x_1} 6x_2\, dx_2 = 3x_1^2$, for $0 < x_1 < 1$. Apparently, $f_{X_1}(x_1) \geq 0$ and $\int_0^1 f_{X_1}(x_1)\, dx_1 = \int_0^1 3x_1^2\, dx_1 = 1$. So, $f_{X_1}(x_1)$ is a valid density function.

(b) $f_{X_2|X_1}(x_2|x_1) = \frac{f(x_1,x_2)}{f_{X_1}(x_1)} = \frac{6x_2}{3x_1^2} = 2\frac{x_2}{x_1^2}$, for $0 < x_2 < x_1$.
So, $P(X_2 < 0.5 \mid X_1 = 0.7) = \frac{2}{0.7^2}\int_0^{0.5} x_2\, dx_2 = \frac{25}{49}$.

2.117 (a) To calculate the marginal distribution of X, the sum is a geometric series: $g(x) = \frac{9}{(16)4^y}\sum_{x=0}^\infty \frac{1}{4^x} = \frac{9}{(16)4^y}\frac{1}{1-1/4} = \frac{3}{4}\cdot\frac{1}{4^x}$, for $x = 0,1,2,\ldots$; similarly, $h(y) = \frac{3}{4}\cdot\frac{1}{4^y}$, for $y = 0,1,2,\ldots$. Since $f(x,y) = g(x)h(y)$, X and Y are independent.

(b) $P(X+Y<4) = f(0,0)+f(0,1)+f(0,2)+f(0,3)+f(1,0)+f(1,1)+f(1,2)+f(2,0)+f(2,1)+f(3,0) = \frac{9}{16}\left(1+\frac{1}{4}+\frac{1}{4^2}+\frac{1}{4^3}+\frac{1}{4}+\frac{1}{4^2}+\frac{1}{4^3}+\frac{1}{4^2}+\frac{1}{4^3}+\frac{1}{4^3}\right) = \frac{9}{16}\left(1+\frac{2}{4}+\frac{3}{4^2}+\frac{4}{4^3}\right) = \frac{63}{64}$.

2.118 $P(\text{the system works}) = P(\text{all components work}) = (0.95)(0.99)(0.92) = 0.86526$.

2.119 $P(\text{the system does not fail}) = P(\text{at least one of the components works})$
$= 1 - P(\text{all components fail}) = 1-(1-0.95)(1-0.94)(1-0.90)(1-0.97) = 0.999991$.

2.120 Denote by X the number of components (out of 5) work.
Then, $P(\text{the system is operational}) = P(X \geq 3) = P(X=3)+P(X=4)+P(X=5) = \binom{5}{3}(0.92)^3(1-0.92)^2 + \binom{5}{4}(0.92)^4(1-0.92) + \binom{5}{5}(0.92)^5 = 0.9955$.

2.122 $E(Y-X) = \int_0^1 \int_0^y 2(y-x)\, dx\, dy = \int_0^1 y^2\, dy = \frac{1}{3}$. Therefore, the average amount of kerosene left in the tank at the end of each day is $(1/3)(1000) = 333$ liters.

2.123 (a) $E(X) = \int_0^\infty \frac{x}{5} e^{-x/5}\, dx = 5$.

(b) $E(X^2) = \int_0^\infty \frac{x^2}{5} e^{-x/5}\, dx = 50$, so $Var(X) = 50 - 5^2 = 25$, and $\sigma = 5$.

(c) $E[(X+5)^2] = E\{[(X-5)+10]^2\} = E[(X-5)^2] + 10^2 + 20E(X-5)$
$= Var(X) + 100 = 125$.

2.124 (a) $E(X) = \int_0^\infty \frac{x}{900} e^{-x/900}\, dx = 900$ hours.

(b) $E(X^2) = \int_0^\infty \frac{x^2}{900} e^{-x/900}\, dx = 1620000$ hours^2.

(c) $Var(X) = E(X^2) - [E(X)]^2 = 810000$ hours^2 and $\sigma = 900$ hours.

2.125 It is known $g(x) = \frac{2}{3}(x+1)$, for $0<x<1$, and $h(y) = \frac{1}{3}(1+4y)$, for $0<y<1$.

(a) $\mu_X = \int_0^1 \frac{2}{3}x(x+1)\, dx = \frac{5}{9}$ and $\mu_Y = \int_0^1 \frac{1}{3}y(1+4y)\, dy = \frac{11}{18}$.

(b) $E[(X+Y)/2] = \frac{1}{2}[E(X)+E(Y)] = \frac{7}{12}$.

2.126 $Cov(aX, bY) = E[(aX - a\mu_X)(bY - b\mu_Y)] = abE[(X-\mu_X)(Y-\mu_Y)] = abCov(X,Y)$.

2.127 Since $g(1)h(1) = (0.17)(0.23) \neq 0.10 = f(1,1)$, X and Y are not independent.

2.128 $E(X) = (-5000)(0.2) + (10000)(0.5) + (30000)(0.3) = \$13,000$.

2.129 (a) $f(x) = \binom{3}{x}(0.15)^x(0.85)^{3-x}$, for $x = 0,1,2,3$.

x	0	1	2	3
$f(x)$	0.614125	0.325125	0.057375	0.003375

(b) $E(X) = 0.45$.

(c) $E(X^2) = 0.585$, so $Var(X) = 0.585 - 0.45^2 = 0.3825$.

(d) $P(X \leq 2) = 1 - P(X=3) = 1 - 0.003375 = 0.996625$.

(e) 0.003375.

(f) Yes.

2.130 (a) $E(X) = \frac{3}{4\times 50^3}\int_{-50}^{50} x(50^2 - x^2)\ dx = 0.$

(b) $E(X^2) = \frac{3}{4\times 50^3}\int_{-50}^{50} x^2(50^2 - x^2)\ dx = 500.$

(c) $\sigma = \sqrt{E(X^2) - [E(X)]^2} = \sqrt{500 - 0} = 22.36.$

2.131 (a) The marginal density of X is

x_1	0	1	2	3	4
$f_{X_1}(x_1)$	0.13	0.21	0.31	0.23	0.12

(b) The marginal density of Y is

x_2	0	1	2	3	4
$f_{X_2}(x_2)$	0.10	0.30	0.39	0.15	0.06

(c) Given $X_2 = 3$, the conditional density function of X_1 is $f(x_1, 3)/0.15$. So

x_1	0	1	2	3	4
$f_{X_2}(x_2)$	$\frac{7}{15}$	$\frac{1}{5}$	$\frac{1}{15}$	$\frac{1}{5}$	$\frac{1}{15}$

(d) $E(X_1) = (0)(0.13) + (1)(0.21) + (2)(0.31) + (3)(0.23) + (4)(0.12) = 2.$

(e) $E(X_2) = (0)(0.10) + (1)(0.30) + (2)(0.39) + (3)(0.15) + (4)(0.06) = 1.77.$

(f) $E(X_1|X_2 = 3) = (0)\left(\frac{7}{15}\right) + (1)\left(\frac{1}{5}\right) + (2)\left(\frac{1}{15}\right) + (3)\left(\frac{1}{5}\right) + (4)\left(\frac{1}{15}\right) = \frac{18}{15} = \frac{6}{5} = 1.2.$

(g) $E(X_1^2) = (0)^2(0.13) + (1)^2(0.21) + (2)^2(0.31) + (3)^2(0.23) + (4)^2(0.12) = 5.44.$
So, $\sigma_{X_1} = \sqrt{E(X_1^2) - [E(X_1)]^2} = \sqrt{5.44 - 2^2} = \sqrt{1.44} = 1.2.$

2.132 (a) The marginal densities of X and Y are, respectively,

x	0	1	2
$g(x)$	0.2	0.32	0.48

y	0	1	2
$h(y)$	0.26	0.35	0.39

The conditional density of X given $Y = 2$ is

x	0	1	2
$f_{X\|Y=2}(x\|2)$	$\frac{4}{39}$	$\frac{5}{39}$	$\frac{30}{39}$

(b) $E(X) = (0)(0.2) + (1)(0.32) + (2)(0.48) = 1.28,$
$E(X^2) = (0)^2(0.2) + (1)^2(0.32) + (2)^2(0.48) = 2.24$, and
$Var(X) = 2.24 - 1.28^2 = 0.6016.$

(c) $E(X|Y = 2) = (1)\frac{5}{39} + (2)\frac{30}{39} = \frac{65}{39}$ and $E(X^2|Y = 2) = (1)^2\frac{5}{39} + (2)^2\frac{30}{39} = \frac{125}{39}$. So,
$Var(X) = \frac{125}{39} - \left(\frac{65}{39}\right)^2 = \frac{650}{1521} = \frac{50}{117}.$

2.133 Using the approximation formula, $Var(Y) \approx \sum_{i=1}^{k} \left[\frac{\partial h(x_1,x_2,\ldots,x_k)}{\partial x_i}\right]^2 \Big|_{x_i=\mu_i,\ 1\le i\le k} \sigma_i^2$, we have

$$Var(\hat{Y}) \approx \sum_{i=0}^{2} \left(\frac{\partial e^{b_0+b_1k_1+b_2k_2}}{\partial b_i}\right)^2 \Big|_{b_i=\beta_i,\ 0\le i\le 2} \sigma_{b_i}^2 = e^{2(\beta_0+k_1\beta_1+k_2\beta_2)}(\sigma_0^2 + k_1^2\sigma_1^2 + k_2^2\sigma_2^2).$$

2.134 (a) $E(Y) = 10\int_0^1 y(1-y)^9\ dy = -\ y(1-y)^{10}|_0^1 + \int_0^1 (1-y)^{10}\ dy = \frac{1}{11}$.

(b) $E(1-Y) = 1 - E(Y) = \frac{10}{11}$.

(c) $Var(Z) = Var(1-Y) = Var(Y) = E(Y^2) - [E(Y)]^2 = \frac{10}{11^2\times 12} = 0.006887$.

(b) For $n = 20$, $P(X < 10) = P(X \le 9) = 0.0171$.

3.6 For $n = 8$ and $p = 0.6$, we have

(a) $P(X = 3) = b(3; 8, 0.6) = P(X \le 3) - P(X \le 2) = 0.1737 - 0.0498 = 0.1239$.

(b) $P(X \ge 5) = 1 - P(X \le 4) = 1 - 0.4059 = 0.5941$.

3.7 For $n = 15$ and $p = 0.25$, we have

(a) $P(3 \le X \le 6) = P(X \le 6) - P(X \le 2) = 0.9434 - 0.2361 = 0.7073$.

(b) $P(X < 4) = P(X \le 3) = 0.4613$.

(c) $P(X > 5) = 1 - P(X \le 5) = 1 - 0.8516 = 0.1484$.

3.8 From Table A.1 with $n = 9$ and $p = 0.25$, we have $P(X < 4) = 0.8343$.

3.9 From Table A.1 with $n = 7$ and $p = 0.9$, we have
$P(X = 5) = P(X \le 5) - P(X \le 4) = 0.1497 - 0.0257 = 0.1240$.

3.10 (a) $n = 4$, $P(X = 4) = (0.9)^4 = 0.6561$.

(b) The bulls could win in 4, 5, 6 or 7 games. So the probability is

$$0.6561 + \binom{4}{3}(0.9)^4(0.1) + \binom{5}{3}(0.9)^4(0.1)^2 + \binom{6}{3}(0.9)^4(0.1)^3 = 0.9973.$$

(c) The probability that the Bulls win is always 0.9.

3.11 $p = 0.4$ and $n = 5$.

(a) $P(X = 0) = 0.0778$.

(b) $P(X < 2) = P(X \le 1) = 0.3370$.

(c) $P(X > 3) = 1 - P(X \le 3) = 1 - 0.9130 = 0.0870$.

3.12 Probability of 2 or more of 4 engines operating when $p = 0.6$ is

$$P(X \ge 2) = 1 - P(X \le 1) = 0.8208,$$

and the probability of 1 or more of 2 engines operating when $p = 0.6$ is

$$P(X \ge 1) = 1 - P(X = 0) = 0.8400.$$

The 2-engine plane has a slightly higher probability for a successful flight when $p = 0.6$.

3.13 Let X_1 = number of times encountered green light with $P(\text{Green}) = 0.35$,
X_2 = number of times encountered yellow light with $P(\text{Yellow}) = 0.05$, and
X_3 = number of times encountered red light with $P(\text{Red}) = 0.60$. Then

$$f(x_1, x_2, x_3) = \binom{n}{x_1, x_2, x_3}(0.35)^{x_1}(0.05)^{x_2}(0.60)^{x_3}.$$

3.14 (a) $\mu = np = (15)(0.25) = 3.75$.

(b) $\sigma^2 = npq = (15)(0.25)(0.75) = 2.8125$.

3.15 (a) $\binom{10}{2,5,3}(0.225)^2(0.544)^5(0.231)^3 = 0.0749$.

(b) $\binom{10}{10}(0.544)^{10}(0.456)^0 = 0.0023$.

(c) $\binom{10}{0}(0.225)^0(0.775)^{10} = 0.0782$.

3.16 $p = 0.40$ and $n = 6$, so $P(X = 4) = P(X \le 4) - P(X \le 3) = 0.9590 - 0.8208 = 0.1382$.

3.17 $n = 20$ and the probability of a defective is $p = 0.10$. So, $P(X \le 3) = 0.8670$.

3.18 $n = 8$ and $p = 0.60$;

(a) $P(X = 6) = \binom{8}{6}(0.6)^6(0.4)^2 = 0.2090$.

(b) $P(X = 6) = P(X \le 6) - P(X \le 5) = 0.8936 - 0.6846 = 0.2090$.

3.19 $n = 20$ and $p = 0.90$;

(a) $P(X = 18) = P(X \le 18) - P(X \le 17) = 0.6083 - 0.3231 = 0.2852$.

(b) $P(X \ge 15) = 1 - P(X \le 14) = 1 - 0.0113 = 0.9887$.

(c) $P(X \le 18) = 0.6083$.

3.20 $n = 20$;

(a) $p = 0.20$, $P(X \ge x) \le 0.5$ and $P(X < x) > 0.5$ yields $x = 5$.

(b) $p = 0.80$, $P(Y \ge y) \ge 0.8$ and $P(Y < y) < 0.2$ yields $y = 15$.

3.21 $P(X \ge 1) = 1 - P(X = 0) = 1 - h(0; 15, 3, 6) = 1 - \frac{\binom{6}{0}\binom{9}{3}}{\binom{15}{3}} = \frac{53}{65}$.

3.22 (a) Probability that all 4 fire $= h(4; 10, 4, 7) = \frac{1}{6}$.

(b) Probability that at most 2 will not fire $= \sum_{x=0}^{2} h(x; 10, 4, 3) = \frac{29}{30}$.

3.23 $P(X \le 2) = \sum_{x=0}^{2} h(x; 50, 5, 10) = 0.9517$.

3.24 $h(2; 9, 5, 4) = \frac{\binom{4}{2}\binom{5}{3}}{\binom{9}{5}} = \frac{10}{21}$.

3.25 Using the binomial approximation of the hypergeometric distribution with $p = 30/150 = 0.2$, the probability is $1 - \sum_{x=0}^{2} b(x; 10, 0.2) = 0.3222$.

3.26 (a) $P(X = 0) = h(0; 25, 3, 3) = \frac{77}{115}$.

(b) $P(X = 1) = h(1; 25, 3, 1) = \frac{3}{25}$.

3.27 (a) $P(X = 0) = b(0; 3, 3/25) = 0.6815$.

(b) $P(1 \leq X \leq 3) = \sum_{x=1}^{3} b(x; 3, 1/25) = 0.1153$.

3.28 The binomial approximation of the hypergeometric with $p = 1 - 4000/10000 = 0.6$ gives a probability of $\sum_{x=0}^{7} b(x; 15, 0.6) = 0.2131$.

3.29 $h(5; 25, 15, 10) = \frac{\binom{10}{5}\binom{15}{10}}{\binom{25}{15}} = 0.2315$.

3.30 Using the extension of the hypergeometric distribution the probability is

$$\frac{\binom{13}{5}\binom{13}{2}\binom{13}{3}\binom{13}{3}}{\binom{52}{13}} = 0.0129.$$

3.31 (a) $\frac{\binom{3}{0}\binom{17}{5}}{\binom{20}{5}} = 0.3991$.

(b) $\frac{\binom{3}{2}\binom{17}{3}}{\binom{20}{5}} = 0.1316$.

3.32 (a) $\frac{\binom{2}{1}\binom{13}{4}}{\binom{15}{5}} = 0.4762$.

(b) $\frac{\binom{2}{2}\binom{13}{3}}{\binom{15}{5}} = 0.0952$.

3.33 $N = 10000$, $n = 30$ and $k = 300$. Using binomial approximation to the hypergeometric distribution with $p = 300/10000 = 0.03$, the probability of $\{X \geq 1\}$ can be determined by

$$1 - b(0; 30, 0.03) = 1 - (0.97)^{30} = 0.599.$$

3.34 From the negative binomial distribution, we obtain

$$b^*(8; 2, 1/6) = \binom{7}{1}(1/6)^2(5/6)^6 = 0.0651.$$

3.35 The probability that all coins turn up the same is 1/4. Using the geometric distribution with $p = 3/4$ and $q = 1/4$, we have

$$P(X < 4) = \sum_{x=1}^{3} g(x; 3/4) = \sum_{x=1}^{3} (3/4)(1/4)^{x-1} = \frac{63}{64}.$$

3.36 (a) Using the geometric distribution, we have $g(5; 2/3) = (2/3)(1/3)^4 = 2/243$.

(b) Using the negative binomial distribution, we have

$$b^*(5; 3, 2/3) = \binom{4}{2}(2/3)^3(1/3)^2 = \frac{16}{81}.$$

3.37 (a) $P(X > 5) = \sum_{x=6}^{\infty} p(x; 5) = 1 - \sum_{x=0}^{5} p(x; 5) = 0.3840.$

(b) $P(X = 0) = p(0; 5) = 0.0067.$

3.38 (a) Using the Poisson distribution with $x = 5$ and $\mu = 3$, we find from Table A.2 that

$$p(5; 3) = \sum_{x=0}^{5} p(x; 3) - \sum_{x=0}^{4} p(x; 3) = 0.1008.$$

(b) $P(X < 3) = P(X \leq 2) = 0.4232.$

(c) $P(X \geq 2) = 1 - P(X \leq 1) = 0.8009.$

3.39 Using the geometric distribution

(a) $P(X = 3) = g(3; 0.7) = (0.7)(0.3)^2 = 0.0630.$

(b) $P(X < 4) = \sum_{x=1}^{3} g(x; 0.7) = \sum_{x=1}^{3} (0.7)(0.3)^{x-1} = 0.9730.$

3.40 (a) $P(X < 4) = P(X \leq 3) = 0.1512.$

(b) $P(6 \leq X \leq 8) = P(X \leq 8) - P(X \leq 5) = 0.4015.$

3.41 (a) $P(X \geq 4) = 1 - P(X \leq 3) = 0.1429.$

(b) $P(X = 0) = p(0; 2) = 0.1353.$

3.42 (a) $\mu = np = (1875)(0.004) = 7.5$, so $P(X < 5) = P(X \leq 4) \approx 0.1321.$

(b) $P(8 \leq X \leq 10) = P(X \leq 10) - P(X \leq 7) \approx 0.8622 - 0.5246 = 0.3376.$

3.43 $\mu = np = (10000)(0.001) = 10$, so

$$P(6 \leq X \leq 8) = P(X \leq 8) - P(X \leq 5) \approx \sum_{x=0}^{8} p(x; 10) - \sum_{x=0}^{5} p(x; 10) = 0.2657.$$

3.44 $\mu = (10000)(0.001) = 10$ and $\sigma^2 = 10.$

3.45 $\mu = 6$ and $\sigma^2 = 6.$

3.46 (a) $P(X = 4|\lambda t = 6) = 0.2851 - 0.1512 = 0.1339.$

(b) $P(X \geq 4|\lambda t = 6) = 1 - 0.1512 = 0.8488.$

(c) $P(X \geq 75|\lambda t = 72) = 1 - \sum_{x=0}^{74} p(x;74) = 0.3773.$

3.47 (a) $P(X \leq 3|\lambda t = 5) = 0.2650.$

(b) $P(X > 1|\lambda t = 5) = 1 - 0.0404 = 0.9596.$

3.48 $\mu = np = (1875)(0.004) = 7.5.$

3.49 (a) $P(X > 10|\lambda t = 14) = 1 - 0.1757 = 0.8243.$

(b) $\lambda t = 14.$

3.50 $\mu = 1$ and $\sigma^2 = 0.99.$

3.51 $\mu = (4000)(0.001) = 4.$

3.52 (a) $P(X \leq 1|\lambda t = 2) = 0.4060.$

(b) $\mu = \lambda t = (2)(5) = 10$ and $P(X \leq 4|\lambda t = 10) = 0.0293.$

3.53 $\mu = \lambda t = (1.5)(5) = 7.5$ and $P(X = 0|\lambda t = 7.5) = e^{-7.5} = 5.53 \times 10^{-4}.$

3.54 $p = 0.03$ with a $g(x; 0.03)$. So, $P(X = 16) = (0.03)(1 - 0.03)^{14} = 0.0196$ and $\mu = \frac{1}{0.03} - 1 = 32.33.$

3.55 (a) $P(X > 10|\lambda t = 5) = 1 - P(X \leq 10|\lambda t = 5) = 1 - 0.9863 = 0.0137.$

(b) $\mu = \lambda t = (5)(3) = 15$, so $P(X > 20|\lambda t = 15) = 1 - P(X \leq 20|\lambda = 15) = 1 - 0.9170 = 0.0830.$

3.56 So, Let Y = number of shifts until it fails. Then Y follows a geometric distribution with $p = 0.10$. So,

$$\begin{aligned} P(Y \leq 6) &= g(1; 0.1) + g(2; 0.1) + \cdots + g(6; 0.1) \\ &= (0.1)[1 + (0.9) + (0.9)^2 + \cdots + (0.9)^5] = 0.4686. \end{aligned}$$

3.57 $f(x) = \frac{1}{B-A}$ for $A \leq x \leq B$.

(a) $\mu = \int_A^B \frac{x}{B-A}\, dx = \frac{B^2-A^2}{2(B-A)} = \frac{A+B}{2}.$

(b) $E(X^2) = \int_A^B \frac{x^2}{B-A}\, dx = \frac{B^3-A^3}{3(B-A)}.$
So, $\sigma^2 = \frac{B^3-A^3}{3(B-A)} - \left(\frac{A+B}{2}\right)^2 = \frac{4(B^2+AB+A^2)-3(B^2+2AB+A^2)}{12} = \frac{B^2-2AB+A^2}{12} = \frac{(B-A)^2}{12}.$

3.58 $P(X > 2.5|X \leq 4) = \frac{P(2.5<X\leq 4)}{P(X\leq 4)} = \frac{4-2.5}{4-1} = \frac{1}{2}.$

3.59 $A = 7$ and $B = 10.$

(a) $P(X \leq 8.8) = \frac{8.8-7}{3} = 0.60.$

(b) $P(7.4 < X < 9.5) = \frac{9.5-7.4}{3} = 0.70.$

(c) $P(X \geq 8.5) = \frac{10-8.5}{3} = 0.50$.

3.60 (a) The area to the left of z is $1 - 0.3622 = 0.6378$ which is closer to the tabled value 0.6368 than to 0.6406. Therefore, we choose $z = 0.35$.

(b) From Table A.3, $z = -1.21$.

(c) The total area to the left of z is 0.5000+0.4838=0.9838. Therefore, from Table A.3, $z = 2.14$.

(d) The distribution contains an area of 0.025 to the left of $-z$ and therefore a total area of 0.025+0.95=0.975 to the left of z. From Table A.3, $z = 1.96$.

3.61 (a) Area=0.0823.

(b) Area=$1 - 0.9750 = 0.0250$.

(c) Area=$0.2578 - 0.0154 = 0.2424$.

(d) Area=0.9236.

(e) Area=$1 - 0.1867 = 0.8133$.

(f) Area=$0.9591 - 0.3156 = 0.6435$.

3.62 (a) Since $P(Z > k) = 0.2946$, then $P(Z < k) = 0.7054$/ From Table A.3, we find $k = 0.54$.

(b) From Table A.3, $k = -1.72$.

(c) The area to the left of $z = -0.93$ is found from Table A.3 to be 0.1762. Therefore, the total area to the left of k is 0.1762+0.7235=0.8997, and hence $k = 1.28$.

3.63 (a) $z = (15 - 18)/2.5 = -1.2$; $P(X < 15) = P(Z < -1.2) = 0.1151$.

(b) $z = -0.76$, $k = (2.5)(-0.76) + 18 = 16.1$.

(c) $z = 0.91$, $k = (2.5)(0.91) + 18 = 20.275$.

(d) $z_1 = (17 - 18)/2.5 = -0.4$, $z_2 = (21 - 18)/2.5 = 1.2$;
$P(17 < X < 21) = P(-0.4 < Z < 1.2) = 0.8849 - 0.3446 = 0.5403$.

3.64 (a) $z = (17 - 30)/6 = -2.17$. Area=$1 - 0.0150 = 0.9850$.

(b) $z = (22 - 30)/6 = -1.33$. Area=0.0918.

(c) $z_1 = (32-30)/6 = 0.33$, $z_2 = (41-30)/6 = 1.83$. Area $= 0.9664-0.6293 = 0.3371$.

(d) $z = 0.84$. Therefore, $x = 30 + (6)(0.84) = 35.04$.

(e) $z_1 = -1.15$, $z_2 = 1.15$. Therefore, $x_1 = 30 + (6)(-1.15) = 23.1$ and $x_2 = 30 + (6)(1.15) = 36.9$.

3.65 (a) $z = (224 - 200)/15 = 1.6$. Fraction of the cups containing more than 224 millimeters is $P(Z > 1.6) = 0.0548$.

(b) $z_1 = (191 - 200)/15 = -0.6$, $Z_2 = (209 - 200)/15 = 0.6$;
$P(191 < X < 209) = P(-0.6 < Z < 0.6) = 0.7257 - 0.2743 = 0.4514$.

(c) $z = (230 - 200)/15 = 2.0$; $P(X > 230) = P(Z > 2.0) = 0.0228$. Therefore, $(1000)(0.0228) = 22.8$ or approximately 23 cups will overflow.

(d) $z = -0.67$, $x = (15)(-0.67) + 200 = 189.95$ millimeters.

3.66 (a) $z = (31.7 - 30)/2 = 0.85$; $P(X > 31.7) = P(Z > 0.85) = 0.1977$. Therefore, 19.77% of the loaves are longer than 31.7 centimeters.

(b) $z_1 = (29.3 - 30)/2 = -0.35$, $z_2 = (33.5 - 30)/2 = 1.75$; $P(29.3 < X < 33.5) = P(-0.35 < Z < 1.75) = 0.9599 - 0.3632 = 0.5967$. Therefore, 59.67% of the loaves are between 29.3 and 33.5 centimeters in length.

(c) $z = (25.5 - 30)/2 = -2.25$; $P(X < 25.5) = P(Z < -2.25) = 0.0122$. Therefore, 1.22% of the loaves are shorter than 25.5 centimeters in length.

3.67 (a) $z = (32 - 40)/6.3 = -1.27$; $P(X > 32) = P(Z > -1.27) = 1 - 0.1020 = 0.8980$.

(b) $z = (28 - 40)/6.3 = -1.90$, $P(X < 28) = P(Z < -1.90) = 0.0287$.

(c) $z_1 = (37 - 40)/6.3 = -0.48$, $z_2 = (49 - 40)/6.3 = 1.43$; So, $P(37 < X < 49) = P(-0.48 < Z < 1.43) = 0.9236 - 0.3156 = 0.6080$.

3.68 (a) $z = (10.075 - 10.000)/0.03 = 2.5$; $P(X > 10.075) = P(Z > 2.5) = 0.0062$. Therefore, 0.62% of the rings have inside diameters exceeding 10.075 cm.

(b) $z_1 = (9.97 - 10)/0.03 = -1.0$, $z_2 = (10.03 - 10)/0.03 = 1.0$; $P(9.97 < X < 10.03) = P(-1.0 < Z < 1.0) = 0.8413 - 0.1587 = 0.6826$.

(c) $z = -1.04$, $x = 10 + (0.03)(-1.04) = 9.969$ cm.

3.69 (a) $z = (30 - 24)/3.8 = 1.58$; $P(X > 30) = P(Z > 1.58) = 0.0571$.

(b) $z = (15 - 24)/3.8 = -2.37$; $P(X > 15) = P(Z > -2.37) = 0.9911$. He is late 99.11% of the time.

(c) $z = (25 - 24)/3.8 = 0.26$; $P(X > 25) = P(Z > 0.26) = 0.3974$.

(d) $z = 1.04$, $x = (3.8)(1.04) + 24 = 27.952$ minutes.

(e) Using the binomial distribution with $p = 0.0571$, we get $b(2; 3, 0.0571) = \binom{3}{2}(0.0571)^2(0.9429) = 0.0092$.

3.70 $\mu = 99.61$ and $\sigma = 0.08$.

(a) $P(99.5 < X < 99.7) = P(-1.375 < Z < 1.125) = 0.8697 - 0.08455 = 0.7852$.

(b) $P(Z > -1.645) = 0.05$; $x = (-1.645)(0.08) + 99.61 = 99.4784$.

3.71 $z = -1.88$, $x = (2)(-1.88) + 10 = 6.24$ years.

3.72 (a) $z = (159.75 - 174.5)/6.9 = -2.14$; $P(X < 159.75) = P(Z < -2.14) = 0.0162$. Therefore, $(1000)(0.0162) = 16$ students.

(b) $z_1 = (171.25 - 174.5)/6.9 = -0.47$, $z_2 = (182.25 - 174.5)/6.9 = 1.12$. $P(171.25 < X < 182.25) = P(-0.47 < Z < 1.12) = 0.8686 - 0.3192 = 0.5494$. Therefore, $(1000)(0.5494) = 549$ students.

(c) $z_1 = (174.75 - 174.5)/6.9 = 0.04$, $z_2 = (175.25 - 174.5)/6.9 = 0.11$.
$P(174.75 < X < 175.25) = P(0.04 < Z < 0.11) = 0.5438 - 0.5160 = 0.0278$.
Therefore, (1000)(0.0278)=28 students.

(d) $z = (187.75 - 174.5)/6.9 = 1.92$; $P(X > 187.75) = P(Z > 1.92) = 0.0274$.
Therefore, $(1000)(0.0274) = 27$ students.

3.73 (a) $z = (10,175 - 10,000)/100 = 1.75$. Proportion of components exceeding 10.150 kilograms in tensile strength$= P(X > 10,175) = P(Z > 1.75) = 0.0401$.

(b) $z_1 = (9,775 - 10,000)/100 = -2.25$ and $z_2 = (10,225 - 10,000)/100 = 2.25$. Proportion of components scrapped$= P(X < 9,775) + P(X > 10,225) = P(Z < -2.25) + P(Z > 2.25) = 2P(Z < -2.25) = 0.0244$.

3.74 (a) $z = (9.55 - 8)/0.9 = 1.72$. Fraction of poodles weighing over 9.5 kilograms $= P(X > 9.55) = P(Z > 1.72) = 0.0427$.

(b) $z = (8.65 - 8)/0.9 = 0.72$. Fraction of poodles weighing at most 8.6 kilograms $= P(X < 8.65) = P(Z < 0.72) = 0.7642$.

(c) $z_1 = (7.25 - 8)/0.9 = -0.83$ and $z_2 = (9.15 - 8)/0.9 = 1.28$.
Fraction of poodles weighing between 7.3 and 9.1 kilograms inclusive
$= P(7.25 < X < 9.15) = P(-0.83 < Z < 1.28) = 0.8997 - 0.2033 = 0.6964$.

3.75 $z = (94.5 - 115)/12 = -1.71$; $P(X < 94.5) = P(Z < -1.71) = 0.0436$. Therefore, $(0.0436)(600) = 26$ students will be rejected.

3.76 (a) $x_1 = \mu + 1.3\sigma$ and $x_2 = \mu - 1.3\sigma$. Then $z_1 = 1.3$ and $z_2 = -1.3$. $P(X > \mu+1.3\sigma)+P(X < 1.3\sigma) = P(Z > 1.3)+P(Z < -1.3) = 2P(Z < -1.3) = 0.1936$. Therefore, 19.36%.

(b) $x_1 = \mu+0.52\sigma$ and $x_2 = \mu-0.52\sigma$. Then $z_1 = 0.52$ and $z_2 = -0.52$. $P(\mu-0.52\sigma < X < \mu + 0.52\sigma) = P(-0.52 < Z < 0.52) = 0.6985 - 0.3015 = 0.3970$. Therefore, 39.70%.

3.77 $n = 100$.

(a) $p = 0.01$ with $\mu = (100)(0.01) = 1$ and $\sigma = \sqrt{(100)(0.01)(0.99)} = 0.995$.
So, $z = (0.5 - 1)/0.995 = -0.503$. $P(X \le 0) \approx P(Z \le -0.503) = 0.3085$.

(b) $p = 0.05$ with $\mu = (100)(0.05) = 5$ and $\sigma = \sqrt{(100)(0.05)(0.95)} = 2.1794$.
So, $z = (0.5 - 5)/2.1794 = -2.06$. $P(X \le 0) \approx P(X \le -2.06) = 0.0197$.

3.78 $\mu = np = (100)(0.1) = 10$ and $\sigma = \sqrt{(100)(0.1)(0.9)} = 3$.

(a) $z = (13.5 - 10)/3 = 1.17$; $P(X > 13.5) = P(Z > 1.17) = 0.1210$.

(b) $z = (7.5 - 10)/3 = -0.83$; $P(X < 7.5) = P(Z < -0.83) = 0.2033$.

3.79 $\mu = (100)(0.9) = 90$ and $\sigma = \sqrt{(100)(0.9)(0.1)} = 3$.

(a) $z_1 = (83.5 - 90)/3 = -2.17$ and $z_2 = (95.5 - 90)/3 = 1.83$.
$P(83.5 < X < 95.5) = P(-2.17 < Z < 1.83) = 0.9664 - 0.0150 = 0.9514$.

(b) $z = (85.5 - 90)/3 = -1.50$; $P(X < 85.5) = P(Z < -1.50) = 0.0668$.

3.80 $\mu = (80)(3/4) = 60$ and $\sigma = \sqrt{(80)(3/4)(1/4)} = 3.873$.

(a) $z = (49.5 - 60)/3.873 = -2.71$; $P(X > 49.5) = P(Z > -2.71) = 1 - 0.0034 = 0.9966$.

(b) $z = (56.5 - 60)/3.873 = -0.90$; $P(X < 56.5) = P(Z < -0.90) = 0.1841$.

3.81 (a) $p = 0.05$, $n = 100$ with $\mu = 5$ and $\sigma = \sqrt{(100)(0.05)(0.95)} = 2.1794$.
So, $z = (2.5 - 5)/2.1794 = -1.147$; $P(X \geq 2) \approx P(Z \geq -1.147) = 0.8749$.

(b) $z = (10.5 - 5)/2.1794 = 2.524$; $P(X \geq 10) \approx P(Z > 2.52) = 0.0059$.

3.82 $\mu = (200)(0.05) = 10$ and $\sigma = \sqrt{(200)(0.05)(0.95)} = 3.082$ with
$z = (9.5 - 10)/3.082 = -0.16$. $P(X < 10) = P(Z < -0.16) = 0.4364$.

3.83 $\mu = (400)(1/10) = 40$ and $\sigma = \sqrt{(400)(1/10)(9/10)} = 6$.

(a) $z = (31.5 - 40)/6 = -1.42$; $P(X < 31.5) = P(Z < -1.42) = 0.0778$.

(b) $z = (49.5 - 40)/6 = 1.58$; $P(X > 49.5) = P(Z > 1.58) = 1 - 0.9429 = 0.0571$.

(c) $z_1 = (34.5 - 40)/6 = -0.92$ and $z_2 = (46.5 - 40)/6 = 1.08$;
$P(34.5 < X < 46.5) = P(-0.92 < Z < 1.08) = 0.8599 - 0.1788 = 0.6811$.

3.84 (a) $\mu = (100)(0.8) = 80$ and $\sigma = \sqrt{(100)(0.8)(0.2)} = 4$ with $z = (74.5 - 80)/4 = -1.38$.
$P(\text{Claim is rejected when } p = 0.8) = P(Z < -1.38) = 0.0838$.

(b) $\mu = (100)(0.7) = 70$ and $\sigma = \sqrt{(100)(0.7)(0.3)} = 4.583$ with $z = (74.5 - 70)/4.583 = 0.98$.
$P(\text{Claim is accepted when } p = 0.7) = P(Z > 0.98) = 1 - 0.8365 = 0.1635$.

3.85 (a) $P(X \geq 230) = P\left(Z > \frac{230-170}{30}\right) = 0.0228$.

(b) Denote by Y the number of students whose serum cholesterol level exceed 230 among the 300. Then $Y \sim b(y; 300, 0.0228$ with

$$\mu = (300)(0.0228) = 6.84, \text{ and } \sigma = \sqrt{(300)(0.0228)(1 - 0.0228)} = 2.5854.$$

So, $z = \frac{8-0.5-6.84}{2.5854} = 0.26$ and
$P(X \geq 8) \approx P(Z > 0.26) = 0.3974$.

3.86 $n = 200$; $X =$ The number of no shows with $p = 0.02$. $z = \frac{3-0.5-4}{\sqrt{(200)(0.02)(0.98)}} = -0.76$.
Therefore, $P(\text{airline overbooks the flight}) = 1 - P(X \geq 3) \approx 1 - P(Z > -0.76) = 0.2236$.

3.87 (a) Denote by X the number of failures among the 20. $X \sim b(x; 20, 0.01)$ and $P(X > 1) = 1 - b(0; 20, 0.01) - b(1; 20, 0.01) = 1 - \binom{20}{0}(0.01)^0(0.99)^{20} - \binom{20}{1}(0.01)(0.99)^{19} = 0.01686$.

(b) $n = 500$ and $p = 0.01$ with $\mu = (500)(0.01) = 5$ and $\sigma = \sqrt{(500)(0.01)(0.99)} = 2.2249$. So, $P(\text{more than 8 failures}) \approx P(Z > (8.5 - 5)/2.2249) = P(Z > 1.57) = 1 - 0.9418 = 0.0582$.

3.88 $P(X > 9) = \frac{1}{9}\int_9^\infty xe^{-x/3}\,dx = \left[-\frac{x}{3}e^{-x/3} - e^{-x/3}\right]\Big|_9^\infty = 4e^{-3} = 0.1992$.

3.89 $P(1.8 < X < 2.4) = \int_{1.8}^{2.4} xe^{-x}\,dx = [-xe^{-x} - e^{-x}]\big|_{1.8}^{2.4} = 2.8e^{-1.8} - 3.4e^{-2.4} = 0.1545$.

3.90 (a) $P(X < 1) = 4\int_0^1 xe^{-2x}\,dx = [-2xe^{-2x} - e^{-2x}]\big|_0^1 = 1 - 3e^{-2} = 0.5940$.

(b) $P(X > 2) = 4\int_0^\infty xe^{-2x}\,dx = [-2xe^{-2x} - e^{-2x}]\big|_2^\infty = 5e^{-4} = 0.0916$.

3.91 $\mu = \alpha\beta = (2)(3) = 6$ million liters; $\sigma^2 = \alpha\beta^2 = (2)(9) = 18$.

3.92 (a) $\mu = \alpha\beta = 6$ and $\sigma^2 = \alpha\beta^2 = 12$. Substituting $\alpha = 6/\beta$ into the variance formula we find $6\beta = 12$ or $\beta = 2$ and then $\alpha = 3$.

(b) $P(X > 12) = \frac{1}{16}\int_{12}^\infty x^2 e^{-x/2}\,dx$. Integrating by parts twice gives

$$P(X > 12) = \frac{1}{16}\left[-2x^2e^{-x/2} - 8xe^{-x/2} - 16e^{-x/2}\right]\Big|_{12}^\infty = 25e^{-6} = 0.0620.$$

3.93 $P(X < 3) = \frac{1}{4}\int_0^3 e^{-x/4}\,dx = -e^{-x/4}\Big|_0^3 = 1 - e^{-3/4} = 0.5276$.
Let Y be the number of days a person is served in less than 3 minutes. Then
$P(Y \geq 4) = \sum_{x=4}^{6} b(y; 6, 1 - e^{-3/4}) = \binom{6}{4}(0.5276)^4(0.4724)^2 + \binom{6}{5}(0.5276)^5(0.4724)$
$+ \binom{6}{6}(0.5276)^6 = 0.3968$.

3.94 $P(X < 1) = \frac{1}{2}\int_0^1 e^{-x/2}\,dx = -e^{-x/2}\Big|_0^1 = 1 - e^{-1/2} = 0.3935$. Let Y be the number of switches that fail during the first year. Using the normal approximation we find $\mu = (100)(0.3935) = 39.35$, $\sigma = \sqrt{(100)(0.3935)(0.6065)} = 4.885$, and $z = (30.5 - 39.35)/4.885 = -1.81$. Therefore, $P(Y \leq 30) = P(Z < -1.81) = 0.0352$.

3.95 $\alpha = 5$; $\beta = 10$;

(a) $\alpha\beta = 50$.

(b) $\sigma^2 = \alpha\beta^2 = 500$; so $\sigma = \sqrt{500} = 22.36$.

(c) $P(X > 30) = \frac{1}{\beta^\alpha\Gamma(\alpha)}\int_{30}^\infty x^{\alpha-1}e^{-x/\beta}\,dx$. Using the incomplete gamma with $y = x/\beta$, then

$$1 - P(X \leq 30) = 1 - P(Y \leq 3) = 1 - \int_0^3 \frac{y^4e^{-y}}{\Gamma(5)}\,dy = 1 - 0.185 = 0.815.$$

3.96 $\alpha\beta = 10$; $\sigma = \sqrt{\alpha\beta^2} = \sqrt{50} = 7.07$.

(a) Using integration by parts,

$$P(X \leq 50) = \frac{1}{\beta^\alpha \Gamma(\alpha)} \int_0^{50} x^{\alpha-1} e^{-x/\beta}\, dx = \frac{1}{25} \int_0^{50} x e^{-x/5}\, dx = 0.9995.$$

(b) $P(X < 10) = \frac{1}{\beta^\alpha \Gamma(\alpha)} \int_0^{10} x^{\alpha-1} e^{-x/\beta}\, dx$. Using the incomplete gamma with $y = x/\beta$, we have

$$P(X < 10) = P(Y < 2) = \int_0^2 y e^{-y}\, dy = 0.5940.$$

3.97 $\mu = 3$ seconds with $f(x) = \frac{1}{3} e^{-x/3}$ for $x > 0$.

(a) $P(X > 5) = \int_5^\infty \frac{1}{3} e^{-x/3}\, dx = \frac{1}{3} \left[-3e^{-x/3}\right]\Big|_5^\infty = e^{-5/3} = 0.1889$.

(b) $P(X > 10) = e^{-10/3} = 0.0357$.

3.98 $\beta = 1/5$ and $\alpha = 10$.

(a) $P(X > 10) = 1 - P(X \leq 10) = 1 - 0.9863 = 0.0137$.

(b) $P(X > 2)$ before 10 cars arrive.

$$P(X \leq 2) = \int_0^2 \frac{1}{\beta^\alpha} \frac{x^{\alpha-1} e^{-x/\beta}}{\Gamma(\alpha)}\, dx.$$

Given $y = x/\beta$, then

$$P(X \leq 2) = P(Y \leq 10) = \int_0^{10} \frac{y^{\alpha-1} e^{-y}}{\Gamma(\alpha)}\, dy = \int_0^{10} \frac{y^{10-1} e^{-y}}{\Gamma(10)}\, dy = 0.542,$$

with $P(X > 2) = 1 - P(X \leq 2) = 1 - 0.542 = 0.458$.

3.99 Let T be the time between two consecutive arrivals

(a) $P(T > 1) = P(\text{no arrivals in 1 minute}) = P(X = 0) = e^{-5} = 0.0067$.

(b) $\mu = \beta = 1/5 = 0.2$.

3.100 $n = 15$ and $p = 0.05$.

(a) $P(X \geq 2) = 1 - P(X \leq 1) = 1 - \sum_{x=0}^{1} \binom{15}{x} (0.05)^x (1 - 0.05)^{15-x} = 1 - 0.8290 = 0.1710$.

(b) $p = 0.07$. So, $P(X \leq 1) = \sum_{x=0}^{1} \binom{15}{x} (0.07)^x (1 - 0.07)^{15-x} = 1 - 0.7168 = 0.2832$.

3.101 $n = 100$ and $p = 0.01$.

(a) $P(X > 3) = 1 - P(X \leq 3) = 1 - \sum_{x=0}^{3} \binom{100}{x}(0.01)^x(1-0.01)^{100-x} = 1 - 0.9816 = 0.0184$.

(b) For $p = 0.05$, $P(X \leq 3) = \sum_{x=0}^{3} \binom{100}{x}(0.05)^x(1-0.05)^{100-x} = 0.2578$.

3.102 $\lambda = 2.7$ call/min.

(a) $P(X \leq 4) = \sum_{x=0}^{4} \frac{e^{-2.7}(2.7)^x}{x!} = 0.8629$.

(b) $P(X \leq 1) = \sum_{x=0}^{1} \frac{e^{-2.7}(2.7)^x}{x!} = 0.2487$.

(c) $\lambda t = 13.5$. So,
$P(X > 10) = 1 - P(X \leq 10) = 1 - \sum_{x=0}^{10} \frac{e^{-13.5}(13.5)^x}{x!} = 1 - 0.2112 = 0.7888$.

3.103 $n = 15$ and $p = 0.05$.

(a) $P(X = 5) = \binom{15}{5}(0.05)^5(1-0.05)^{10} = 0.000562$.

(b) I would not believe the claim of 5% defective.

3.104 $\lambda = 0.2$, so $\lambda t = (0.2)(5) = 1$.

(a) $P(X \leq 1) = \sum_{x=0}^{1} \frac{e^{-1}(1)^x}{x!} = 2e^{-1} = 0.7358$. Hence, $P(X > 1) = 1 - 0.7358 = 0.2642$.

(b) $\lambda = 0.25$, so $\lambda t = 1.25$. $P(X|le1) = \sum_{x=0}^{1} \frac{e^{-1.25}(1.25)^x}{x!} = 0.6446$.

3.105 (a) $1 - P(X \leq 1) = 1 - \sum_{x=0}^{1} \binom{100}{x}(0.01)^x(0.99)^{100-x} = 1 - 0.7358 = 0.2642$.

(b) $P(X \leq 1) = \sum_{x=0}^{1} \binom{100}{x}(0.05)^x(0.95)^{100-x} = 0.0371$.

3.106 (a) $\mu = bp = (200)(0.03) = 6$.

(b) $\sigma^2 = npq = 5.82$.

(c) $P(X = 0) = \frac{e^6(6)^0}{0!} = 0.0025$ (using the Poisson approximation).
$P(X = 0) = (0.97)^{200} = 0.0023$ (using the binomial distribution).

3.107 (a) $p^{10}q^0 = (0.99)^{10} = 0.9044$.

(b) In this case, the last 10 of 12 must be good starts and the 2nd attempt must be a bad one. However, the 1st one can be either bad or good. So, the probability is $p^{10}q = (0.99)^{10}(0.01) = (0.9044)(0.01) = 0.009$.

3.108 $n = 75$ with $p = 0.999$.

(a) $X =$ the number of trials, and $P(X = 75) = (0.999)^{75}(0.001)^0 = 0.9277$.

(b) $Y =$ the number of trials before the first failure (geometric distribution), and $P(Y = 20) = (0.001)(0.999)^{19} = 0.000981$.

(c) $1 - P(\text{no failures}) = 1 - (0.001)^0(0.999)^{10} = 0.01$.

3.109 (a) $\binom{10}{1}pq^9 = (10)(0.25)(0.75)^9 = 0.1877$.

(b) Let X be the number of drills until the first success. X follows a geometric distribution with $p = 0.25$. So, the probability of having the first 10 drills being failure is $q^{10} = (0.75)^{10} = 0.056$. So, there is a small prospects for bankruptcy. Also, the probability that the first success appears in the 11th drill is $pq^{10} = 0.014$ which is even smaller.

3.110 It is a negative binomial distribution. $\binom{x-1}{k-1}p^k q^{x-k} = \binom{6-1}{2-1}(0.25)^2(0.75)^4 = 0.0989$.

3.111 $n = 1000$ and $p = 0.01$, with $\mu = (1000)(0.01) = 10$. $P(X < 7) = P(X \leq 6) = 0.1301$.

3.112 $n = 500$;

(a) If $p = 0.01$,

$$P(X \geq 15) = 1 - P(X \leq 14) = 1 - \sum_{x=0}^{14} \binom{500}{x} (0.01)^x (0.99)^{500-x} = 0.00021.$$

This is a very rare probability and thus the original claim that $p = 0.01$ is questionable.

(b) $P(X = 3) = \binom{500}{3}(0.01)^3(0.99)^{497} = 0.1402$.

(c) For (a), if $p = 0.01$, $\mu = (500)(0.01) = 5$. So

$$P(X \geq 15) = 1 - P(X \leq 14) = 1 - 0.9998 = 0.0002.$$

For (b),

$$P(X = 3) = 0.2650 - 0.1247 = 0.1403.$$

3.113 $N = 50$ and $n = 10$.

(a) $k = 2$; $P(X \geq 1) = 1 - P(X = 0) = 1 - \frac{\binom{2}{0}\binom{48}{10}}{\binom{50}{10}} = 1 - 0.6367 = 0.3633$.

(b) Even though the lot contains 2 defectives, the probability of reject the lot is not very high. Perhaps more items should be sampled.

(c) $\mu = (10)(2/50) = 0.4$.

3.114 Suppose n items need to be sampled. $P(X \geq 1) = 1 - \frac{\binom{2}{0}\binom{48}{n}}{\binom{50}{n}} = 1 - \frac{(50-n)(49-n)}{(50)(49)} \geq 0.9$. The solution is $n = 34$.

3.115 Define $X =$ number of screens will detect. Then $X \sim b(x; 3, 0.8)$.

(a) $P(X = 0) = (1 - 0.8)^3 = 0.008$.

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

(c) $P(X \geq 2) = P(X = 2) + P(X = 3) = (3)(0.8)^2(0.2) + (0.8)^3 = 0.896$.

3.116 (a) $P(X = 0) = (1 - 0.8)^n \leq 0.0001$ implies that $n \geq 6$.

(b) $(1 - p)^3 \leq 0.0001$ implies $p \geq 0.9536$.

3.117 $n = 10$ and $p = \frac{2}{50} = 0.04$.

$$P(X \geq 1) = 1 - P(X = 0) \approx 1 - \binom{10}{0}(0.04)^0(1 - 0.04)^{10} = 1 - 0.6648 = 0.3351.$$

The approximation is not that good due to $\frac{n}{N} = 0.2$ is too large.

3.118 (a) $P = \frac{\binom{2}{2}\binom{3}{0}}{\binom{5}{2}} = 0.1$.

(b) $P = \frac{\binom{2}{1}\binom{1}{1}}{\binom{5}{2}} = 0.2$.

3.119 $n = 200$ with $p = 0.00001$.

(a) $P(X \geq 5) = 1 - P(X \leq 4) = 1 - \sum_{x=0}^{4} \binom{200}{x}(0.0001)^x(1 - 0.0001)^{200-x} \approx 0$. This is a rare event. Therefore, the claim does not seem right.

(b) $\mu = np = (200)(0.0001) = 0.02$. Using Poisson approximation,

$$P(X \geq 5) = 1 - P(X \leq 4) \approx 1 - \sum_{x=0}^{4} e^{-0.02}\frac{(0.02)^x}{x!} = 0.$$

3.121 $\mu = np = (1000)(0.49) = 490$, $\sigma = \sqrt{npq} = \sqrt{(1000)(0.49)(0.51)} = 15.808$.

$$z_1 = \frac{481.5 - 490}{15.808} = -0.54, \qquad z_2 = \frac{510.5 - 490}{15.808} = 1.3.$$

$P(481.5 < X < 510.5) = P(-0.54 < Z < 1.3) = 0.9032 - 0.2946 = 0.6086$.

3.122 $P(X > 1/4) = \int_{1/4}^{\infty} 6e^{-6x}\, dx = -e^{-6x}|_{1/4}^{\infty} = e^{-1.5} = 0.223$.

3.123 Manufacturer A:

$$P(X \geq 10000) = P\left(Z \geq \frac{100000 - 14000}{2000}\right) = P(Z \geq -2) = 0.9772.$$

Manufacturer B:

$$P(X \geq 10000) = P\left(Z \geq \frac{10000 - 13000}{1000}\right) = P(Z \geq -3) = 0.9987.$$

Manufacturer B will produce the fewest number of defective rivets.

3.124 (a) $\mu = \beta = 100$ hours.

(b) $P(X \geq 200) = 0.01 \int_{200}^{\infty} e^{-0.01x}\, dx = e^{-2} = 0.1353$.

3.125 (a) $\mu = 85$ and $\sigma = 4$. So, $P(X < 80) = P(Z < -1.25) = 0.1056$.

(b) $\mu = 79$ and $\sigma = 4$. So, $P(X \geq 80) = P(Z > 0.25) = 0.4013$.

3.126 $1/\beta = 1/5$ hours with $\alpha = 2$ failures and $\beta = 5$ hours.

(a) $\alpha\beta = (2)(5) = 10$.

(b) $P(X \geq 12) = \int_{12}^{\infty} \frac{1}{5^2\Gamma(2)} x e^{-x/5}\, dx = \frac{1}{25}\int_{12}^{\infty} x e^{-x/5}\, dx = \left[-\frac{x}{5}e^{-x/5} - e^{-x/5}\right]_{12}^{\infty}$ $= 0.3084$.

3.127 Denote by X the elongation. We have $\mu = 0.05$ and $\sigma = 0.01$.

(a) $P(X \geq 0.1) = P\left(Z \geq \frac{0.1-0.05}{0.01}\right) = P(Z \geq 5) \approx 0$.

(b) $P(X \leq 0.04) = P\left(Z \leq \frac{0.04-0.05}{0.01}\right) = P(Z \leq -1) = 0.1587$.

(c) $P(0.025 \leq X \leq 0.065) = P(-2.5 \leq Z \leq 1.5) = 0.9332 - 0.0062 = 0.9270$.

3.128 Let X be the error. $X \sim n(x; 0, 4)$. So,

$$P(\text{fails}) = 1 - P(-10 < X < 10) = 1 - P(-2.5 < Z < 2.5) = 2(0.0062) = 0.0124.$$

3.129 Let X be the time to bombing with $\mu = 3$ and $\sigma = 0.5$. Then

$$P(1 \leq X \leq 4) = P\left(\frac{1-3}{0.5} \leq Z \leq \frac{4-3}{0.5}\right) = P(-4 \leq Z \leq 2) = 0.9772.$$

P(of an undesirable product) is $1 - 0.9772 = 0.0228$. Hence a product is undesirable is 2.28% of the time.

3.130 X follows a lognormal distribution.

$$P(X \geq 50{,}000) = 1 - \Phi\left(\frac{\ln 50{,}000 - 5}{2}\right) = 1 - \Phi(2.9099) = 1 - 0.9982 = 0.0018.$$

3.131 The mean of X, which follows a lognormal distribution is $\mu = E(X) = e^{\mu+\sigma^2/2} = e^7$.

3.132 $\mu = 10$ and $\sigma = \sqrt{50}$.

(a) $P(X \leq 50) = P(Z \leq 5.66) \approx 1$.

(b) $P(X \leq 10) = 0.5$.

(c) For the larger values, such as 50, both results are similar. However, for smaller values such as 10, the normal population will give you smaller probabilities.

3.133 (a) $\mu = \frac{1}{10}\int_0^{\infty} z e^{-z/10}\, dz = -\, z e^{-z/10}\Big|_0^{\infty} + \int_0^{\infty} e^{-z/10}\, dz = 10$.

(b) Using integral by parts twice, we get

$$E(Z^2) = \frac{1}{10}\int_0^\infty z^2 e^{-z/10}\, dz = 200.$$

So, $\sigma^2 = E(Z^2) - \mu^2 = 200 - (10)^2 = 100$.

(c) $P(Z > 10) = -\, e^{z/10}\big|_{10}^{\infty} = e^{-1} = 0.3679$.

3.134 The problem has changed to a Poisson distribution: the number of calls per hour. Since the average time between two calls in 6.76 is 10 minutes, the average number of calls per hour should be 6. Therefore, the mean and variance of the number of calls per hour should all be 6.

3.135 $\mu = 0.5$ seconds and $\sigma = 0.4$ seconds.

(a) $P(X > 0.3) = P\left(Z > \frac{0.3-0.5}{0.4}\right) = P(Z > -0.5) = 0.6915$.

(b) $P(Z > -1.645) = 0.95$. So, $-1.645 = \frac{x-0.5}{0.4}$ yields $x = -0.158$ seconds. The negative number in reaction time is not reasonable. So, it means that the normal model may not be accurate enough.

3.136 (a) For an exponential distribution with parameter β,

$$P(X > a+b \mid X > a) = \frac{P(X > a+b)}{P(X > a)} = \frac{e^{-a-b}}{e^{-a}} = e^{-b} = P(X > b).$$

So, $P(\text{it will breakdown in the next 21 days} \mid \text{it just broke down}) = P(X > 21) = e^{-21/15} = e^{-1.4} = 0.2466$.

(b) $P(X > 30) = e^{-30/15} = e^{-2} = 0.1353$.

3.137 Let X be the length of time in seconds. Then $Y = \ln(X)$ follows a normal distribution with $\mu = 1.8$ and $\sigma = 2$.

(a) $P(X > 20) = P(Y > \ln 20) = P(Z > (\ln 20 - 1.8)/2) = P(Z > 0.60) = 0.2743$.
$P(X > 60) = P(Y > \ln 60) = P(Z > (\ln 60 - 1.8)/2) = P(Z > 1.15) = 0.1251$.

(b) The mean of the underlying lognormal distribution is $e^{1.8+4/2} = 44.70$ seconds. So, $P(X < 44.70) = P(Z < (\ln 44.70 - 1.8)/2) = P(Z < 1) = 0.8413$.

Chapter 4

Sampling Distributions and Data Descriptions

4.1 (a) Responses of all people in Richmond who have telephones.

(b) Outcomes for a large or infinite number of tosses of a coin.

(c) Length of life of such tennis shoes when worn on the professional tour.

(d) All possible time intervals for this lawyer to drive from her home to her office.

4.2 (a) $\bar{x} = 8.6$ minutes.

(b) $\tilde{x} = 9.5$ minutes.

(c) Mode are 5 and 10 minutes.

4.3 (a) $\bar{x} = 53.75$.

(b) Modes are 75 and 100.

4.4 (a) $\bar{x} = 35.7$ grams.

(b) $\tilde{x} = 32.5$ grams.

(c) Mode=29 grams.

4.5 (a) Range $= 15 - 5 = 10$.

(b) $s^2 = \frac{n\sum\limits_{i=1}^{n} x_i^2 - (\sum\limits_{i=1}^{n} x_i)^2}{n(n-1)} = \frac{(10)(838)-86^2}{(10)(9)} = 10.933$. Taking the square root, we have $s = 3.307$.

4.6 (a) Number of tickets issued by all state troopers in Montgomery County during the Memorial holiday weekend.

(b) Number of tickets issued by all state troopers in South Carolina during the Memorial holiday weekend.

4.7 (a) $s^2 = \frac{1}{n-1}\sum_{x=1}^{n}(x_i-\bar{x})^2 = \frac{1}{11}[(48-35.7)^2+(47-35.7)^2+\cdots+(26-35.7)^2] = 61.2.$

(b) $s^2 = \frac{n\sum_{i=1}^{n}x_i^2-(\sum_{i=1}^{n}x_i)^2}{n(n-1)} = \frac{(12)(15398)-428^2}{(12)(11)} = 61.2.$

4.8 (a) $\bar{x} = 11.69$ milligrams.

(b) $s^2 = \frac{n\sum_{i=1}^{n}x_i^2-(\sum_{i=1}^{n}x_i)^2}{n(n-1)} = \frac{(8)(1168.21)-93.5^2}{(8)(7)} = 10.776.$

4.9 $s^2 = \frac{n\sum_{i=1}^{n}x_i^2-(\sum_{i=1}^{n}x_i)^2}{n(n-1)} = \frac{(20)(148.55)-53.3^2}{(20)(19)} = 0.342$ and hence $s = 0.585$.

4.10 (a) Replace X_i in S^2 by X_i+c for $i=1,2,\ldots,n$. Then $\bar{X}$ becomes $\bar{X}+c$ and

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}[(X_i+c)-(\bar{X}+c)]^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2.$$

(b) Replace X_i by cX_i in S^2 for $i=1,2,\ldots,n$. Then $\bar{X}$ becomes $c\bar{X}$ and

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(cX_i-c\bar{X})^2 = \frac{c^2}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2.$$

4.11 $s^2 = \frac{n\sum_{i=1}^{n}x_i^2-(\sum_{i=1}^{n}x_i)^2}{n(n-1)} = \frac{(6)(207)-33^2}{(6)(5)} = 5.1.$

(a) Multiplying each observation by 3 gives $s^2 = (9)(5.1) = 45.9$.

(b) Adding 5 to each observation does not change the variance. Hence $s^2 = 5.1$.

4.12 Denote by D the difference in scores.

(a) $\bar{D} = 25.15$.

(b) $\tilde{D} = 31.00$.

4.13 $z_1 = -1.9$, $z_2 = -0.4$. Hence,

$$P(\mu_{\bar{X}} - 1.9\sigma_{\bar{X}} < \bar{X} < \mu_{\bar{X}} - 0.4\sigma_{\bar{X}}) = P(-1.9 < Z < -0.4) = 0.3446 - 0.0287 = 0.3159.$$

4.14 (a) For $n = 64$, $\sigma_{\bar{X}} = 5.6/8 = 0.7$, whereas for $n = 196$, $\sigma_{\bar{X}} = 5.6/14 = 0.4$. Therefore, the variance of the sample mean is reduced from 0.49 to 0.16 when the sample size is increased from 64 to 196.

(b) For $n = 784$, $\sigma_{\bar{X}} = 5.6/28 = 0.2$, whereas for $n = 49$, $\sigma_{\bar{X}} = 5.6/7 = 0.8$. Therefore, the variance of the sample mean is increased from 0.04 to 0.64 when the sample size is decreased from 784 to 49.

4.15 $\mu_{\bar{X}} = \mu = 240$, $\sigma_{\bar{X}} = 15/\sqrt{40} = 2.372$. Therefore, $\mu_{\bar{X}} \pm 2\sigma_{\bar{X}} = 240 \pm (2)(2.372)$ or from 235.256 to 244.744, which indicates that a value of $x = 236$ milliliters is reasonable and hence the machine need not be adjusted.

4.16 (a) $\mu_{\bar{X}} = \mu = 174.5$, $\sigma_{\bar{X}} = \sigma/\sqrt{n} = 6.9/5 = 1.38$.

(b) $z_1 = (172.45 - 174.5)/1.38 = -1.49$, $z_2 = (175.85 - 174.5)/1.38 = 0.98$. So,

$$P(172.45 < \bar{X} < 175.85) = P(-1.49 < Z < 0.98) = 0.8365 - 0.0681 = 0.7684.$$

Therefore, the number of sample means between 172.5 and 175.8 inclusive is $(200)(0.7684) = 154$.

(c) $z = (171.95 - 174.5)/1.38 = -1.85$. So,

$$P(\bar{X} < 171.95) = P(Z < -1.85) = 0.0322.$$

Therefore, about $(200)(0.0322) = 6$ sample means fall below 172.0 centimeters.

4.17 (a) $\mu = \sum xf(x) = (4)(0.2) + (5)(0.4) + (6)(0.3) + (7)(0.1) = 5.3$, and $\sigma^2 = \sum(x-\mu)^2 f(x) = (4-5.3)^2(0.2) + (5-5.3)^2(0.4) + (6-5.3)^2(0.3) + (7-5.3)^2(0.1) = 0.81$.

(b) With $n = 36$, $\mu_{\bar{X}} = \mu = 5.3$ and $\sigma^2_{\bar{X}} = \sigma^2/n = 0.81/36 = 0.0225$.

(c) $n = 36$, $\mu_{\bar{X}} = 5.3$, $\sigma_{\bar{X}} = 0.9/6 = 0.15$, and $z = (5.5 - 5.3)/0.15 = 1.33$. So,

$$P(\bar{X} < 5.5) = P(Z < 1.33) = 0.9082.$$

4.18 $n = 36$, $\mu_{\bar{X}} = 40$, $\sigma_{\bar{X}} = 2/6 = 1/3$ and $z = (40.5 - 40)/(1/3) = 1.5$. So,

$$P\left(\sum_{i=1}^{36} X_i > 1458\right) = P(\bar{X} > 40.5) = P(Z > 1.5) = 1 - 0.9332 = 0.0668.$$

4.19 (a) $P(6.4 < \bar{X} < 7.2) = P(-1.8 < Z < 0.6) = 0.6898$.

(b) $z = 1.04$, $\bar{x} = z(\sigma/\sqrt{n}) + \mu = (1.04)(1/3) + 7 = 7.35$.

4.20 $n = 64$, $\mu_{\bar{X}} = 3.2$, $\sigma_{\bar{X}} = \sigma/\sqrt{n} = 1.6/8 = 0.2$.

(a) $z = (2.7 - 3.2)/0.2 = -2.5$, $P(\bar{X} < 2.7) = P(Z < -2.5) = 0.0062$.

(b) $z = (3.5 - 3.2)/0.2 = 1.5$, $P(\bar{X} > 3.5) = P(Z > 1.5) = 1 - 0.9332 = 0.0668$.

(c) $z_1 = (3.2 - 3.2)/0.2 = 0$, $z_2 = (3.4 - 3.2)/0.2 = 1.0$, $P(3.2 < \bar{X} < 3.4) = P(0 < Z < 1.0) = 0.8413 - 0.5000 = 0.3413$.

4.21 $n = 50$, $\bar{x} = 0.23$ and $\sigma = 0.1$. Now, $z = (0.23 - 0.2)/(0.1/\sqrt{50}) = 2.12$; so

$$P(\bar{X} \geq 0.23) = P(Z \geq 2.12) = 0.0170.$$

Hence the probability of having such observations, given the mean $\mu = 0.20$, is small. Therefore, the mean amount to be 0.20 is not likely to be true.

4.22 $\mu_{\bar{X}_1-\bar{X}_2} = 72 - 28 = 44$, $\sigma_{\bar{X}_1-\bar{X}_2} = \sqrt{100/64 + 25/100} = 1.346$ and $z = (44.2 - 44)/1.346 = 0.15$. So, $P(\bar{X}_1 - \bar{X}_2 < 44.2) = P(Z < 0.15) = 0.5596$.

4.23 (a) If the two population mean drying times are truly equal, the probability that the difference of the two sample means is 1.0 is 0.0013, which is very small. This means that the assumption of the equality of the population means are not reasonable.

(b) If the experiment was run 10,000 times, there would be $(10000)(0.0013) = 13$ experiments where $\bar{X}_A - \bar{X}_B$ would be at least 1.0.

4.24 (a) $n_1 = n_2 = 36$ and $z = 0.2/\sqrt{1/36 + 1/36} = 0.85$. So,

$$P(\bar{X}_B - \bar{X}_A \geq 0.2) = P(Z \geq 0.85) = 0.1977.$$

(b) Since the probability in (a) is not negligible, the experiments do not strongly support the conjecture.

4.25 (a) When the population equals the limit, the probability of a sample mean exceeding the limit would be 1/2 due the symmetry of the approximated normal distribution.

(b) $P(\bar{X} \geq 7960 \mid \mu = 7950) = P(Z \geq (7960 - 7950)/(100/\sqrt{25})) = P(Z \geq 0.5) = 0.3085$. No, this is not very strong evidence that the population mean of the process exceeds the government limit.

4.26 (a) $\sigma_{\bar{X}_A-\bar{X}_B} = \sqrt{\frac{5^2}{30} + \frac{5^2}{30}} = 1.29$ and $z = \frac{4-0}{1.29} = 3.10$. So,

$$P(\bar{X}_A - \bar{X}_B > 4 \mid \mu_A = \mu_B) = P(Z > 3.10) = 0.0010.$$

Such a small probability means that the difference of 4 is not likely if the two population means are equal.

(b) Yes, the data strongly support alloy A.

4.27 Since the probability that $\bar{X} \leq 775$ is 0.0062, given that $\mu = 800$ is true, it suggests that this event is very rare and it is very likely that the claim of $\mu = 800$ is not true. On the other hand, if μ is truly, say, 760, the probability

$$P(\bar{X} \leq 775 \mid \mu = 760) = P(Z \leq (775 - 760)/(40/\sqrt{16})) = P(Z \leq 1.5) = 0.9332,$$

which is very high.

4.28 (a) 16.750.

(b) 30.144.

(c) 26.217.

4.29 (a) 27.488.

(b) 18.475.

(c) 36.415.

4.30 (a) $\chi^2_\alpha = \chi^2_{0.01} = 38.932$.

(b) $\chi^2_\alpha = \chi^2_{0.05} = 12.592$.

(c) $\chi^2_{0.01} = 23.209$ and $\chi^2_{0.025} = 20.483$ with $\alpha = 0.01 + 0.015 = 0.025$.

4.31 (a) $\chi^2_\alpha = \chi^2_{0.99} = 0.297$.

(b) $\chi^2_\alpha = \chi^2_{0.025} = 32.852$.

(c) $\chi^2_{0.05} = 37.652$. Therefore, $\alpha = 0.05 - 0.045 = 0.005$. Hence, $\chi^2_\alpha = \chi^2_{0.005} = 46.928$.

4.32 $\chi^2 = \frac{(19)(20)}{8} = 47.5$ while $\chi^2_{0.01} = 36.191$. Conclusion values are not valid.

4.33 (a) $P(S^2 > 9.1) = P\left(\frac{(n-1)S^2}{\sigma^2} > \frac{(24)(9.1)}{6}\right) = P(\chi^2 > 36.4) = 0.05$.

(b) $P(3.462 < S^2 < 10.745) = P\left(\frac{(24)(3.462)}{6} < \frac{(n-1)S^2}{\sigma^2} < \frac{(24)(10.745)}{6}\right)$
$= P(13.848 < \chi^2 < 42.980) = 0.95 - 0.01 = 0.94$.

4.34 Since $\frac{(n-1)S^2}{\sigma^2}$ is a chi-square statistic, it follows that

$$\sigma^2_{(n-1)S^2/\sigma^2} = \frac{(n-1)^2}{\sigma^4}\sigma^2_{S^2} = 2(n-1).$$

Hence, $\sigma^2_{S^2} = \frac{2\sigma^4}{n-1}$, which decreases as n increases.

4.35 (a) $P(T < 2.365) = 1 - 0.025 = 0.975$.

(b) $P(T > 1.318) = 0.10$.

(c) $P(T < 2.179) = 1 - 0.025 = 0.975$, $P(T < -1.356) = P(T > 1.356) = 0.10$. Therefore, $P(-1.356 < T < 2.179) = 0.975 - 0.010 = 0.875$.

(d) $P(T > -2.567) = 1 - P(T > 2.567) = 1 - 0.01 = 0.99$.

4.36 (a) 2.145.

(b) -1.372.

(c) -3.499.

4.37 (a) From Table A.4 we note that 2.069 corresponds to $t_{0.025}$ when $v = 23$. Therefore, $-t_{0.025} = -2.069$ which means that the total area under the curve to the left of $t = k$ is $0.025 + 0.965 = 0.990$. Hence, $k = t_{0.01} = 2.500$.

(b) From Table A.4 we note that 2.807 corresponds to $t_{0.005}$ when $v = 23$. Therefore the total area under the curve to the right of $t = k$ is $0.095 + 0.005 = 0.10$. Hence, $k = t_{0.10} = 1.319$.

(c) $t_{0.05} = 1.714$ for 23 degrees of freedom.

4.38 (a) Since $t_{0.01}$ leaves an area of 0.01 to the right, and $-t_{0.005}$ an area of 0.005 to the left, we find the total area to be $1 - 0.01 - 0.005 = 0.985$ between $-t_{0.005}$ and $t_{0.01}$. Hence, $P(-t_{0.005} < T < t_{0.01}) = 0.985$.

(b) Since $-t_{0.025}$ leaves an area of 0.025 to the left, the desired area is $1-0.025 = 0.975$. That is, $P(T > -t_{0.025}) = 0.975$.

4.39 From Table A.4 we find $t_{0.025} = 2.131$ for $v = 15$ degrees of freedom. Since the value

$$t = \frac{27.5 - 30}{5/4} = -2.00$$

falls between -2.131 and 2.131, the claim is valid.

4.40 $\bar{x} = 0.475$, $s^2 = 0.0336$ and $t = (0.475 - 0.5)/0.0648 = -0.39$. Hence

$$P(\bar{X} < 0.475) = P(T < -0.39) \approx 0.35.$$

So, the result is inconclusive.

4.41 (a) 2.71.

(b) 3.51.

(c) 2.92.

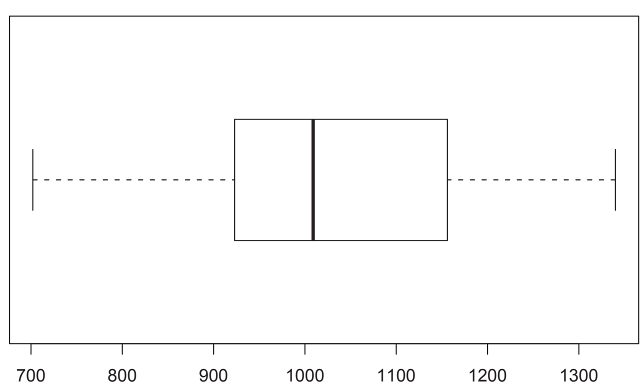
(d) $1/2.11 = 0.47$.

(e) $1/2.90 = 0.34$.

4.42 $s_1^2 = 10.441$ and $s_2^2 = 1.846$ which gives $f = 5.66$. Since, from Table A.6, $f_{0.05}(9, 7) = 3.68$ and $f_{0.01}(9, 7) = 6.72$, the probability of $P(F > 5.66)$ should be between 0.01 and 0.05, which is quite small. Hence the variances may not be equal. Furthermore, if a computer software can be used, the exact probability of $F > 5.66$ can be found 0.0162, or if two sides are considered, $P(F < 1/5.66) + P(F > 5.66) = 0.026$.

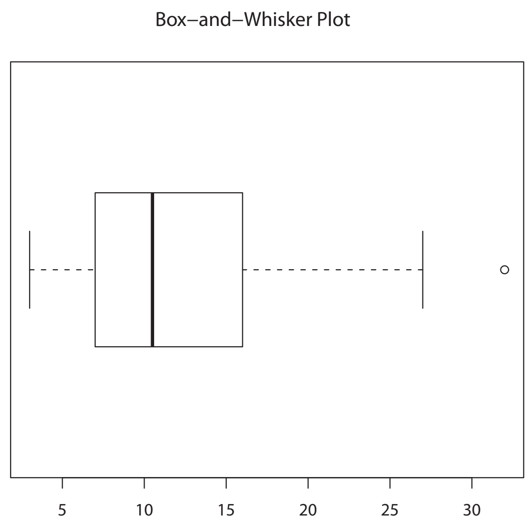
4.43 $s_1^2 = 15750$ and $s_2^2 = 10920$ which gives $f = 1.44$. Since, from Table A.6, $f_{0.05}(4, 5) = 5.19$, the probability of $F > 1.44$ is much bigger than 0.05, which means that the conjecture that the two population variances are equal cannot be rejected. The actual probability of $F > 1.44$ is 0.3442 and $P(F < 1/1.44) + P(F > 1.44) = 0.7170$.

4.44 The box-and-whisker plot is shown below.



The required values are first quartile=925, median=1009, third quartile=1155 and range=$1340 - 702 = 638$.

4.45 The box-and-whisker plot is shown below.



The sample mean = 12.32 and the sample standard deviation = 6.08.

4.46 The variance of the carbon monoxide contents is the same as the variance of the coded measurements. That is, $s^2 = \frac{(15)(199.94)-39^2}{(15)(14)} = 7.039$, which results in $s = 2.653$.

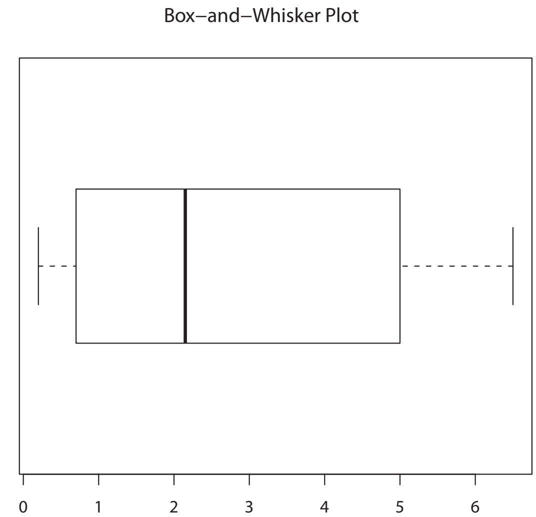
4.47 $P\left(\frac{S_1^2}{S_2^2} < 4.89\right) = P\left(\frac{S_1^2/\sigma^2}{S_2^2/\sigma^2} < 4.89\right) = P(F < 4.89) = 0.99$, where F has 7 and 11 degrees of freedom.

4.48 Let X_1 and X_2 be Poisson variables with parameters $\lambda_1 = 6$ and $\lambda_2 = 6$ representing the number of hurricanes during the first and second years, respectively. Then $Y = X_1 + X_2$ has a Poisson distribution with parameter $\lambda = \lambda_1 + \lambda_2 = 12$.

(a) $P(Y = 15) = \frac{e^{-12}12^{15}}{15!} = 0.0724$.

(b) $P(Y \le 9) = \sum_{y=0}^{9} \frac{e^{-12}12^y}{y!} = 0.2424$.

4.49 The box-and-whisker plot is shown next.



The sample mean is 2.7967 and the sample standard deviation is 2.2273. No outliers are obvious in the data.

4.50 $P\left(\frac{S_1^2}{S_2^2} > 1.26\right) = P\left(\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} > \frac{(15)(1.26)}{10}\right) = P(F > 1.89) \approx 0.05$, where F has 24 and 30 degrees of freedom.

4.51 Probable two low outliers and one high outlier are shown in Figure 4.16.

4.52 $\mu = 5{,}000$ psi, $\sigma = 400$ psi, and $n = 36$.

(a) Using approximate normal distribution (by CLT),

$$P(4800 < \bar{X} < 5200) = P\left(\frac{4800 - 5000}{400/\sqrt{36}} < Z < \frac{5200 - 5000}{400/\sqrt{36}}\right)$$
$$= P(-3 < Z < 3) = 0.9974.$$

(b) To find a z such that $P(-z < Z < z) = 0.99$, we have $P(Z < z) = 0.995$, which results in $z = 2.575$. Hence, by solving $2.575 = \frac{5100-5000}{400/\sqrt{n}}$ we have $n \geq 107$. Note that the value n can be affected by the z values picked (2.57 or 2.58).

4.53 $\bar{x} = 54{,}100$ and $s = 5801.34$. Hence

$$t = \frac{54100 - 53000}{5801.34/\sqrt{10}} = 0.60.$$

So, $P(\bar{X} \geq 54,100) = P(T \geq 0.60)$ is a value between 0.20 and 0.30, which is not a rare event.

4.54 $n_A = n_B = 20$, $\bar{x}_A = 20.50$, $\bar{x}_B = 24.50$, and $\sigma_A = \sigma_B = 5$.

(a) $P(\bar{X}_B - \bar{X}_A \geq 4.0 \mid \mu_A = \mu_B) = P(Z > 4.0/\sqrt{5^2/20 + 5^2/20})$
$= P(Z > 4.0/(5/\sqrt{10})) = P(Z > 2.53) = 0.0059.$

(b) It is extremely unlikely that $\mu_A = \mu_B$.

4.55 (a) $n_A = 30$, $\bar{x}_A = 64.5\%$ and $\sigma_A = 5\%$. Hence,

$$P(\bar{X}_A \le 64.5 \mid \mu_A = 65) = P(Z < (64.5 - 65)/(5/\sqrt{30})) = P(Z < -0.55)$$
$$= 0.2912.$$

There is no evidence that μ_A is less than 65%.

(b) $n_B = 30$, $\bar{x}_B = 70\%$ and $\sigma_B = 5\%$. It turns out $\sigma_{\bar{X}_B - \bar{X}_A} = \sqrt{\frac{5^2}{30} + \frac{5^2}{30}} = 1.29\%$. Hence,

$$P(\bar{X}_B - \bar{X}_A \ge 5.5 \mid \mu_A = \mu_B) = P\left(Z \ge \frac{5.5}{1.29}\right) = P(Z \ge 4.26) \approx 0.$$

It does strongly support that μ_B is greater than μ_A.

(c) i) Since $\sigma_{\bar{X}_B} = \frac{5}{\sqrt{30}} = 0.9129$, $\bar{X}_B \sim n(x; 65\%, 0.9129\%)$.
ii) $\bar{X}_A - \bar{X}_B \sim n(x; 0, 1.29\%)$.
iii) $\frac{\bar{X}_A - \bar{X}_B}{\sigma\sqrt{2/30}} \sim n(z; 0, 1)$.

4.56 $P(\bar{X}_B \ge 70) = P\left(Z \ge \frac{70-65}{0.9129}\right) = P(Z \ge 5.48) \approx 0$.

4.57 It is known, from Table A.3, that $P(-1.96 < Z < 1.96) = 0.95$. Given $\mu = 20$ and $\sigma = \sqrt{9} = 3$, we equate $1.96 = \frac{20.1-20}{3/\sqrt{n}}$ to obtain $n = \left(\frac{(3)(1.96)}{0.1}\right)^2 = 3457.44 \approx 3458$.

4.58 It is known that $P(-2.575 < Z < 2.575) = 0.99$. Hence, by equating $2.575 = \frac{0.05}{1/\sqrt{n}}$, we obtain $n = \left(\frac{2.575}{0.05}\right)^2 = 2652.25 \approx 2653$.

4.59 $\mu = 9$ and $\sigma = 1$. Then

$$P(9 - 1.5 < X < 9 + 1.5) = P(7.5 < X < 10.5) = P(-1.5 < Z < 1.5)$$
$$= 0.9332 - 0.0668 = 0.8664.$$

Thus the proportion of defective is $1 - 0.8654 = 0.1346$. To meet the specifications 99% of the time, we need to equate $2.575 = \frac{1.5}{\sigma}$, since $P(-2.575 < Z < 2.575) = 0.99$. Therefore, $\sigma = \frac{1.5}{2.575} = 0.5825$.

4.60 With the 39 degrees of freedom,

$$P(S^2 \le 0.188 \mid \sigma^2 = 1.0) = P(\chi^2 \le (39)(0.188)) = P(\chi^2 \le 7.332) \approx 0,$$

which means that it is impossible to observe $s^2 = 0.188$ with $n = 40$ for $\sigma^2 = 1$.

Note that Table A.5 does not provide this value. Computer software gives the value of 0.

Chapter 5

One- and Two-Sample Estimation Problems

5.1 $n = [(2.575)(5.8)/2]^2 = 56$ when rounded up.

5.2 $n = 30, \bar{x} = 780$, and $\sigma = 40$. Also, $z_{0.02} = 2.054$. So, a 96% confidence interval for the population mean can be calculated as

$$780 - (2.054)(40/\sqrt{30}) < \mu < 780 + (2.054)(40/\sqrt{30}),$$

or $765 < \mu < 795$.

5.3 $n = 75, \bar{x} = 0.310, \sigma = 0.0015$, and $z_{0.025} = 1.96$. A 95% confidence interval for the population mean is

$$0.310 - (1.96)(0.0015/\sqrt{75}) < \mu < 0.310 + (1.96)(0.0015/\sqrt{75}),$$

or $0.3097 < \mu < 0.3103$.

5.4 $n = 50, \bar{x} = 174.5, \sigma = 6.9$, and $z_{0.01} = 2.33$.

(a) A 98% confidence interval for the population mean is
$174.5 - (2.33)(6.9/\sqrt{50}) < \mu < 174.5 + (2.33)(6.9/\sqrt{50})$, or $172.23 < \mu < 176.77$.

(b) $e < (2.33)(6.9)/\sqrt{50} = 2.27$.

5.5 $n = 100, \bar{x} = 23,500, \sigma = 3900$, and $z_{0.005} = 2.575$.

(a) A 99% confidence interval for the population mean is
$23,500 - (2.575)(3900/10) < \mu < 23,500 + (2.575)(3900/10)$, or
$22,496 < \mu < 24,504$.

(b) $e < (2.575)(3900/10) = 1004$.

5.6 $n = [(2.05)(40)/10]^2 = 68$ when rounded up.

5.7 $n = [(1.96)(0.0015)/0.0005]^2 = 35$ when rounded up.

5.8 $n = [(1.96)(40)/15]^2 = 28$ when rounded up.

5.9 $n = 20, \bar{x} = 11.3, s = 2.45$, and $t_{0.025} = 2.093$ with 19 degrees of freedom. A 95% confidence interval for the population mean is

$$11.3 - (2.093)(2.45/\sqrt{20}) < \mu < 11.3 + (2.093)(2.45/\sqrt{20}),$$

or $10.15 < \mu < 12.45$.

5.10 To show the identity, we have the following:

$$\begin{aligned}
\sum_{i=1}^{n}(x_i - \mu)^2 &= \sum_{i=1}^{n}[(x_i - \bar{x}) + (\bar{x} - \mu)]^2 \\
&= \sum_{i=1}^{n}\left[(x_i - \bar{x})^2 + (\bar{x} - \mu)^2 + 2(x_i - \bar{x})(\bar{x} - \mu)\right] \\
&= \sum_{i=1}^{n}(x_i - \bar{x})^2 + \sum_{i=1}^{n}(\bar{x} - \mu)^2 + 2\sum_{i=1}^{n}(x_i - \bar{x})(\bar{x} - \mu) \\
&= \sum_{i=1}^{n}(x_i - \bar{x})^2 + n(\bar{x} - \mu)^2 + 2(\bar{x} - \mu)\sum_{i=1}^{n}(x_i - \bar{x}) \\
&= \sum_{i=1}^{n}(x_i - \bar{x})^2 + n(\bar{x} - \mu)^2.
\end{aligned}$$

5.11 $n = 12, \bar{x} = 48.50, s = 1.5$, and $t_{0.05} = 1.796$ with 11 degrees of freedom. A 90% confidence interval for the population mean is

$$48.50 - (1.796)(1.5/\sqrt{12}) < \mu < 48.50 + (1.796)(1.5/\sqrt{12}),$$

or $47.722 < \mu < 49.278$.

5.12 $n = 15, \bar{x} = 3.7867, s = 0.9709, 1-\alpha = 95\%$, and $t_{0.025} = 2.145$ with 14 degrees of freedom. So, by calculating $3.7867 \pm (2.145)(0.9709)\sqrt{1 + 1/15}$ we obtain $(1.6358, 5.9376)$ which is a 95% prediction interval the drying times of next paint.

5.13 $n = 100, \bar{x} = 23{,}500, s = 3,900, 1 - \alpha = 0.99$, and $t_{0.005} \approx 2.66$ with 60 degrees of freedom (use table from the book) or $t_{0.005} = 2.626$ if 100 degrees of freedom is used. The prediction interval of next automobile will be driven in Virginia (using 2.66) is $23{,}500 \pm (2.66)(3{,}900)\sqrt{1 + 1/100}$ which yields $13{,}075 < \mu < 33{,}925$ kilometers.

5.14 $n = 20, \bar{x} = 11.3, s = 2.45$, and $t_{0.025} = 2.093$ with 19 degrees of freedom. A 95% prediction interval for a future observation is

$$11.3 \pm (2.093)(2.45)\sqrt{1 + 1/20} = 11.3 \pm 5.25,$$

which yields $(6.05, 16.55)$.

5.15 $n = 25, \bar{x} = 325.05, s = 0.5, \gamma = 5\%$, and $1 - \alpha = 90\%$, with $k = 2.208$. So, $325.05 \pm (2.208)(0.5)$ yields $(323.946, 326.151)$. Thus, we are 95% confident that this tolerance interval will contain 90% of the aspirin contents for this brand of buffered aspirin.

5.16 $n = 12, \bar{x} = 48.50, s = 1.5, 1 - \alpha = 0.90$, and $\gamma = 0.05$, with $k = 2.655$. The tolerance interval is $48.50 \pm (2.655)(1.5)$ which yields $(44.52, 52.48)$.

5.17 $n = 15, \bar{x} = 3.84$, and $s = 3.07$. To calculate an upper 95% prediction limit, we obtain $t_{0.05} = 1.761$ with 14 degrees of freedom. So, the upper limit is $3.84 + (1.761)(3.07)\sqrt{1 + 1/15} = 3.84 + 5.58 = 9.42$. This means that a new observation will have a chance of 95% to fall into the interval $(-\infty, 9.42)$. To obtain an upper 95% tolerance limit, using $1 - \alpha = 0.95$ and $\gamma = 0.05$, with $k = 2.566$, we get $3.84 + (2.566)(3.07) = 11.72$. Hence, we are 95% confident that $(-\infty, 11.72)$ will contain 95% of the orthophosphorous measurements in the river.

5.18 $n = 9, \bar{x} = 1.0056, s = 0.0245, 1 - \alpha = 0.95$, and $\gamma = 0.05$, with $k = 3.532$. The tolerance interval is $1.0056 \pm (3.532)(0.0245)$ which yields $(0.919, 1.092)$.

5.19 (a) $n = 16$, $\bar{x} = 1.0025$, $s = 0.0202$, $1 - \alpha = 0.99$, $t_{0.005} = 2.947$ with 15 degrees of freedom. A 99% confidence interval for the mean diameter is $1.0025 \pm (2.947)(0.0202)/\sqrt{16}$ which yields $(0.9876, 1.0174)$ centimeters.

(b) A 99% prediction interval for the diameter of a new metal piece is

$$1.0025 \pm (2.947)(0.0202)\sqrt{1 + 1/16}$$

which yields $(0.9411, 1.0639)$ centimeters.

(c) For $n = 16$, $1-\gamma = 0.99$ and $1-\alpha = 0.95$, we find $k = 3.421$. Hence, the tolerance limits are $1.0025 \pm (3.421)(0.0202)$ which give $(0.9334, 1.0716)$.

5.20 $n = 50, \bar{x} = 78.3$, and $s = 5.6$. Since $t_{0.05} = 1.677$ with 49 degrees of freedom, the bound of a lower 95% prediction interval for a single new observation is $78.3 - (1.677)(5.6)\sqrt{1 + 1/50} = 68.91$. So, the interval is $(68.91, \infty)$. On the other hand, with $1 - \alpha = 95\%$ and $\gamma = 0.01$, the k value for a one-sided tolerance limit is 2.269 and the bound is $78.3 - (2.269)(5.6) = 65.59$. So, the tolerance interval is $(65.59, \infty)$.

5.21 Since the manufacturer would be more interested in the mean tensile strength for future products, it is conceivable that prediction interval and tolerance interval may be more interesting than just a confidence interval.

5.22 This time $1-\alpha = 0.99$ and $\gamma = 0.05$ with $k = 3.126$. So, the tolerance limit is $78.3 - (3.126)(5.6) = 60.79$. Since 62 exceeds the lower bound of the interval, yes, this is a cause of concern.

5.23 In Exercise 9.14, a 95% prediction interval for a new observation is calculated as $(1.6358, 5.9377)$. Since 6.9 is in the outside range of the prediction interval, this new observation is likely to be an outlier.

5.24 $n = 12, \bar{x} = 48.50, s = 1.5, 1-\alpha = 0.95$, and $\gamma = 0.05$, with $k = 2.815$. The lower bound of the one-sided tolerance interval is $48.50 - (2.815)(1.5) = 44.278$. Their claim is not necessarily correct.

5.25 $n_1 = 100, n_2 = 200, \bar{x}_1 = 12.2, \bar{x}_2 = 9.1, s_1 = 1.1$, and $s_2 = 0.9$. It is known that $z_{0.01} = 2.327$. So

$$(12.2 - 9.1) \pm 2.327\sqrt{1.1^2/100 + 0.9^2/200} = 3.1 \pm 0.30,$$

or $2.80 < \mu_1 - \mu_2 < 3.40$. The treatment appears to reduce the mean amount of metal removed.

5.26 $n_A = 50, n_B = 50, \bar{x}_A = 78.3, \bar{x}_B = 87.2, \sigma_A = 5.6$, and $\sigma_B = 6.3$. It is known that $z_{0.025} = 1.96$. So, a 95% confidence interval for the difference of the population means is

$$(87.2 - 78.3) \pm 1.96\sqrt{5.6^2/50 + 6.3^2/50} = 8.9 \pm 2.34,$$

or $6.56 < \mu_A - \mu_B < 11.24$.

5.27 $n_1 = 12, n_2 = 10, \bar{x}_1 = 85, \bar{x}_2 = 81, s_1 = 4, s_2 = 5$, and $s_p = 4.478$ with $t_{0.05} = 1.725$ with 20 degrees of freedom. So

$$(85 - 81) \pm (1.725)(4.478)\sqrt{1/12 + 1/10} = 4 \pm 3.31,$$

which yields $0.69 < \mu_1 - \mu_2 < 7.31$.

5.28 $n_1 = 10, n_2 = 10, \bar{x}_1 = 0.399, \bar{x}_2 = 0.565, s_1 = 0.07279, s_2 = 0.18674$, and $s_p = 0.14172$ with $t_{0.025} = 2.101$ with 18 degrees of freedom. So,

$$(0.565 - 0.399) \pm (2.101)(0.14172)\sqrt{1/10 + 1/10} = 0.166 \pm 0.133,$$

which yields $0.033 < \mu_1 - \mu_2 < 0.299$.

5.29 $n_1 = 14, n_2 = 16, \bar{x}_1 = 17, \bar{x}_2 = 19, s_1^2 = 1.5, s_2^2 = 1.8$, and $s_p = 1.289$ with $t_{0.005} = 2.763$ with 28 degrees of freedom. So,

$$(19 - 17) \pm (2.763)(1.289)\sqrt{1/16 + 1/14} = 2 \pm 1.30,$$

which yields $0.70 < \mu_1 - \mu_2 < 3.30$.

5.30 $n_1 = 12, n_2 = 10, \bar{x}_1 = 16, \bar{x}_2 = 11, s_1 = 1.0, s_2 = 0.8$, and $s_p = 0.915$ with $t_{0.05} = 1.725$ with 20 degrees of freedom. So,

$$(16 - 11) \pm (1.725)(0.915)\sqrt{1/12 + 1/10} = 5 \pm 0.68,$$

which yields $4.3 < \mu_1 - \mu_2 < 5.7$.

5.31 $n_A = n_B = 12, \bar{x}_A = 36,300, \bar{x}_B = 38,100, s_A = 5,000, s_B = 6,100$, and

$$v = \frac{5000^2/12 + 6100^2/12}{\frac{(5000^2/12)^2}{11} + \frac{(6100^2/12)^2}{11}} = 21,$$

with $t_{0.025} = 2.080$ with 21 degrees of freedom. So,

$$(36,300 - 38,100) \pm (2.080)\sqrt{\frac{5000^2}{12} + \frac{6100^2}{12}} = -1,800 \pm 4,736,$$

which yields $-6,536 < \mu_A - \mu_B < 2,936$.

5.32 $n = 8, \bar{d} = -1112.5, s_d = 1454$, with $t_{0.005} = 3.499$ with 7 degrees of freedom. So,

$$-1112.5 \pm (3.499)\frac{1454}{\sqrt{8}} = -1112.5 \pm 1798.7,$$

which yields $-2911.2 < \mu_D < 686.2$.

5.33 $n = 9, \bar{d} = 2.778, s_d = 4.5765$, with $t_{0.025} = 2.306$ with 8 degrees of freedom. So,

$$2.778 \pm (2.306)\frac{4.5765}{\sqrt{9}} = 2.778 \pm 3.518,$$

which yields $-0.74 < \mu_D < 6.30$.

5.34 $n_I = 5, n_{II} = 7, \bar{x}_I = 98.4, \bar{x}_{II} = 110.7, s_I = 8.375$, and $s_{II} = 32.185$, with

$$v = \frac{(8.735^2/5 + 32.185^2/7)2}{\frac{(8.735^2/5)^2}{4} + \frac{(32.185^2/7)^2}{6}} = 7$$

So, $t_{0.05} = 1.895$ with 7 degrees of freedom.

$$(110.7 - 98.4) \pm 1.895\sqrt{8.735^2/5 + 32.185^2/7} = 12.3 \pm 24.2,$$

which yields $-11.9 < \mu_{II} - \mu_I < 36.5$.

5.35 $n = 10, \bar{d} = 14.89\%$, and $s_d = 30.4868$, with $t_{0.025} = 2.262$ with 9 degrees of freedom. So,

$$14.89 \pm (2.262)\frac{30.4868}{\sqrt{10}} = 14.89 \pm 21.81,$$

which yields $-6.92 < \mu_D < 36.70$.

5.36 $n_A = n_B = 20, \bar{x}_A = 32.91, \bar{x}_B = 30.47, s_A = 1.57, s_B = 1.74$, and $S_p = 1.657$.

(a) $t_{0.025} \approx 2.042$ with 38 degrees of freedom. So,

$$(32.91 - 30.47) \pm (2.042)(1.657)\sqrt{1/20 + 1/20} = 2.44 \pm 1.07,$$

which yields $1.37 < \mu_A - \mu_B < 3.51$.

(b) Since it is apparent that type A battery has longer life, it should be adopted.

5.37 $n_A = n_B = 15, \bar{x}_A = 3.82, \bar{x}_B = 4.94, s_A = 0.7794, s_B = 0.7538$, and $s_p = 0.7667$ with $t_{0.025} = 2.048$ with 28 degrees of freedom. So,

$$(4.94 - 3.82) \pm (2.048)(0.7667)\sqrt{1/15 + 1/15} = 1.12 \pm 0.57,$$

which yields $0.55 < \mu_B - \mu_A < 1.69$.

5.38 $n_1 = 8, n_2 = 13, \bar{x}_1 = 1.98, \bar{x}_2 = 1.30, s_1 = 0.51, s_2 = 0.35$, and $s_p = 0.416$. $t_{0.025} = 2.093$ with 19 degrees of freedom. So,

$$(1.98 - 1.30) \pm (2.093)(0.416)\sqrt{1/8 + 1/13} = 0.68 \pm 0.39,$$

which yields $0.29 < \mu_1 - \mu_2 < 1.07$.

5.39 $n = 1000, \hat{p} = \frac{228}{1000} = 0.228, \hat{q} = 0.772$, and $z_{0.005} = 2.575$. So, using method 1,

$$0.228 \pm (2.575)\sqrt{\frac{(0.228)(0.772)}{1000}} = 0.228 \pm 0.034,$$

which yields $0.194 < p < 0.262$.

When we use method 2, we have

$$\frac{0.228 + 2.575^2/2000}{1 + 2.575^2/1000} \pm \frac{2.575}{1 + 2.575^2/1000}\sqrt{\frac{(0.228)(0.772)}{1000} + \frac{2.575^2}{4(1000)^2}} = 0.2298 \pm 0.0341,$$

which yields an interval of $(0.1957, 0.2639)$.

5.40 $n = 100, \hat{p} = \frac{8}{100} = 0.08, \hat{q} = 0.92$, and $z_{0.01} = 2.33$. So,

$$0.08 \pm (2.33)\sqrt{\frac{(0.08)(0.92)}{100}} = 0.08 \pm 0.063,$$

which yields $0.017 < p < 0.143$, for method 1.

Now for method 2, we have

$$\frac{0.08 + 2.33^2/200}{1 + 2.33^2/100} \pm \frac{2.33}{1 + 2.33^2/100}\sqrt{\frac{(0.08)(0.92)}{100} + \frac{2.33^2}{4(100)^2}} = 0.1016 \pm 0.0652,$$

which yields $(0.0364, 0.1669)$.

5.41 (a) $n = 200, \hat{p} = 0.57, \hat{q} = 0.43$, and $z_{0.02} = 2.05$. So,

$$0.57 \pm (2.05)\sqrt{\frac{(0.57)(0.43)}{200}} = 0.57 \pm 0.072,$$

which yields $0.498 < p < 0.642$.

(b) Error $\leq (2.05)\sqrt{\frac{(0.57)(0.43)}{200}} = 0.072$.

5.42 $n = 500.\hat{p} = \frac{485}{500} = 0.97, \hat{q} = 0.03$, and $z_{0.05} = 1.645$. So,

$$0.97 \pm (1.645)\sqrt{\frac{(0.97)(0.03)}{500}} = 0.97 \pm 0.013,$$

which yields $0.957 < p < 0.983$.

5.43 (a) $n = 40, \hat{p} = \frac{34}{40} = 0.85, \hat{q} = 0.15$, and $z_{0.025} = 1.96$. So,

$$0.85 \pm (1.96)\sqrt{\frac{(0.85)(0.15)}{40}} = 0.85 \pm 0.111,$$

which yields $0.739 < p < 0.961$.

(b) Since $p = 0.8$ falls in the confidence interval, we can not conclude that the new system is better.

5.44 $n = 100, \hat{p} = \frac{24}{100} = 0.24, \hat{q} = 0.76$, and $z_{0.005} = 2.575$.

(a) $0.24 \pm (2.575)\sqrt{\frac{(0.24)(0.76)}{100}} = 0.24 \pm 0.110$, which yields $0.130 < p < 0.350$.

(b) Error $\leq (2.575)\sqrt{\frac{(0.24)(0.76)}{100}} = 0.110$.

5.45 $n = \frac{(2.05)^2(0.57)(0.43)}{(0.02)^2} = 2576$ when round up.

5.46 $n = \frac{(2.575)^2(0.228)(0.772)}{(0.05)^2} = 467$ when round up.

5.47 $n = \frac{(2.33)^2(0.08)(0.92)}{(0.05)^2} = 160$ when round up.

5.48 $n = \frac{(2.575)^2}{(4)(0.01)^2} = 16577$ when round up.

5.49 $n = \frac{(1.96)^2}{(4)(0.04)^2} = 601$ when round up.

5.50 $n_1 = 250, n_2 = 175, \hat{p}_1 = \frac{80}{250} = 0.32, \hat{p}_2 = \frac{40}{175} = 0.2286$, and $z_{0.05} = 1.645$. So,

$$(0.32 - 0.2286) \pm (1.645)\sqrt{\frac{(0.32)(0.68)}{250} + \frac{(0.2286)(0.7714)}{175}} = 0.0914 \pm 0.0713,$$

which yields $0.0201 < p_1 - p_2 < 0.1627$. From this study we conclude that there is a significantly higher proportion of women in electrical engineering than there is in chemical engineering.

5.51 $n_M = n_F = 1000, \hat{p}_M = 0.250, \hat{q}_M = 0.750, \hat{p}_F = 0.275, \hat{q}_F = 0.725$, and $z_{0.025} = 1.96$. So

$$(0.275 - 0.250) \pm (1.96)\sqrt{\frac{(0.250)(0.750)}{1000} + \frac{(0.275)(0.725)}{1000}} = 0.025 \pm 0.039,$$

which yields $-0.0136 < p_F - p_M < 0.0636$.

5.52 $n_{5^\circ\mathrm{C}} = n_{15^\circ\mathrm{C}} = 20, \hat{p}_{5^\circ\mathrm{C}} = 0.50, \hat{p}_{15^\circ\mathrm{C}} = 0.75$, and $z_{0.025} = 1.96$. So,

$$(0.5 - 0.75) \pm (1.96)\sqrt{\frac{(0.50)(0.50)}{20} + \frac{(0.75)(0.25)}{20}} = -0.25 \pm 0.2899,$$

which yields $-0.5399 < p_{5^\circ\mathrm{C}} - p_{15^\circ\mathrm{C}} < 0.0399$. Since this interval includes 0, the significance of the difference cannot be shown at the confidence level of 95%.

5.53 $n_1 = n_2 = 500, \hat{p}_1 = \frac{120}{500} = 0.24, \hat{p}_2 = \frac{98}{500} = 0.196$, and $z_{0.05} = 1.645$. So,

$$(0.24 - 0.196) \pm (1.645)\sqrt{\frac{(0.24)(0.76)}{500} + \frac{(0.196)(0.804)}{500}} = 0.044 \pm 0.0429,$$

which yields $0.0011 < p_1 - p_2 < 0.0869$. Since 0 is not in this confidence interval, we conclude, at the level of 90% confidence, that inoculation has an effect on the incidence of the disease.

5.54 $s^2 = 16$ with $v = 19$ degrees of freedom. It is known $\chi^2_{0.01} = 36.191$ and $\chi^2_{0.99} = 7.633$. Hence

$$\frac{(19)(16)}{36.191} < \sigma^2 < \frac{(19)(16)}{7.633}, \text{ or } 8.400 < \sigma^2 < 39.827.$$

5.55 $s^2 = 0.815$ with $v = 4$ degrees of freedom. Also, $\chi^2_{0.025} = 11.143$ and $\chi^2_{0.975} = 0.484$. So,

$$\frac{(4)(0.815)}{11.143} < \sigma^2 < \frac{(4)(0.815)}{0.484}, \quad \text{which yields} \quad 0.293 < \sigma^2 < 6.736.$$

Since this interval contains 1, the claim that σ^2 seems valid.

5.56 $s^2 = 0.0006$ with $v = 8$ degrees of freedom. Also, $\chi^2_{0.005} = 21.955$ and $\chi^2_{0.995} = 1.344$. Hence,

$$\frac{(8)(0.0006)}{21.955} < \sigma^2 < \frac{(8)(0.0006)}{1.344}, \text{ or } 0.00022 < \sigma^2 < 0.00357.$$

5.57 $s^2 = 6.0025$ with $v = 19$ degrees of freedom. Also, $\chi^2_{0.025} = 32.852$ and $\chi^2_{0.975} = 8.907$. Hence,

$$\frac{(19)(6.0025)}{32.852} < \sigma^2 < \frac{(19)(6.0025)}{8.907}, \text{ or } 3.472 < \sigma^2 < 12.804.$$

5.58 $s^2 = 2.25$ with $v = 11$ degrees of freedom. Also, $\chi^2_{0.05} = 19.675$ and $\chi^2_{0.95} = 4.575$. Hence,

$$\frac{(11)(2.25)}{19.675} < \sigma^2 < \frac{(11)(2.25)}{4.575}, \text{ or } 1.258 < \sigma^2 < 5.410.$$

5.59 $n = 75, x = 28$, hence $\hat{p} = \frac{28}{75} = 0.3733$. Since $z_{0.025} = 1.96$, a 95% confidence interval for p can be calculate as

$$0.3733 \pm (1.96)\sqrt{\frac{(0.3733)(0.6267)}{75}} = 0.3733 \pm 0.1095,$$

which yields $0.2638 < p < 0.4828$. Since the interval contains 0.421, the claim made by the *Roanoke Times* seems reasonable.

5.60 $n = 7, \bar{d} = 3.557, s_d = 2.776$, and $t_{0.025} = 2.447$ with 6 degrees of freedom. So,

$$3.557 \pm (2.447)\frac{2.776}{\sqrt{7}} = 3.557 \pm 2.567,$$

which yields $0.99 < \mu_D < 6.12$. Since 0 is not in the interval, the claim appears valid.

5.61 $n = 6, \bar{d} = 1.5, s_d = 1.543$, and $t_{0.025} = 2.571$ with 5 degrees of freedom. So,

$$1.5 \pm (2.571)\frac{1.543}{\sqrt{6}} = 1.5 \pm 1.62,$$

which yields $-0.12 < \mu_D < 3.12$.

5.62 $n = 12, \bar{d} = 417.5, s_d = 1186.643$, and $t_{0.05} = 1.796$ with 11 degrees of freedom. So,

$$417.5 \pm (1.796)\frac{1186.643}{\sqrt{12}} = 417.5 \pm 615.23,$$

which yields $-197.73 < \mu_D < 1032.73$.

5.63 $n_p = n_u = 8, \bar{x}_p = 86,250.000, \bar{x}_u = 79,837.500, \sigma_p = \sigma_u = 4,000$, and $z_{0.025} = 1.96$. So,

$$(86250 - 79837.5) \pm (1.96)(4000)\sqrt{1/8 + 1/8} = 6412.5 \pm 3920,$$

which yields $2,492.5 < \mu_p - \mu_u < 10,332.5$. Hence, polishing does increase the average endurance limit.

5.64 $n_A = 100, n_B = 120, \hat{p}_A = \frac{24}{100} = 0.24, \hat{p}_B = \frac{36}{120} = 0.30$, and $z_{0.025} = 1.96$. So,

$$(0.30 - 0.24) \pm (1.96)\sqrt{\frac{(0.24)(0.76)}{100} + \frac{(0.30)(0.70)}{120}} = 0.06 \pm 0.117,$$

which yields $-0.057 < p_B - p_A < 0.177$.

5.65 $n_N = n_O = 23, s_N^2 = 105.9271, s_O^2 = 77.4138$, and $f_{0.025}(22, 22) = 2.358$. So,

$$\frac{105.9271}{77.4138}\frac{1}{2.358} < \frac{\sigma_N^2}{\sigma_O^2} < \frac{105.9271}{77.4138}(2.358), \text{ or } 0.58 < \frac{\sigma_N^2}{\sigma_O^2} < 3.23.$$

For the ratio of the standard deviations, the 95% confidence interval is approximately

$$0.76 < \frac{\sigma_N}{\sigma_O} < 1.80.$$

Since the intervals contain 1 we will assume that the variability did not change with the local supplier.

5.66 $n_A = n_B = 6, \bar{x}_A = 0.1407, \bar{x}_B = 0.1385, s_A = 0.002805, s_B = 0.002665$, and $s_p = 002736$. Using a 90% confidence interval for the difference in the population means, $t_{0.05} = 1.812$ with 10 degrees of freedom, we obtain

$$(0.1407 - 0.1385) \pm (1.812)(0.002736)\sqrt{1/6 + 1/6} = 0.0022 \pm 0.0029,$$

which yields $-0.0007 < \mu_A - \mu_B < 0.0051$. Since the 90% confidence interval contains 0, we conclude that wire A was not shown to be better than wire B, with 90% confidence.

5.67 $n_1 = n_2 = 300, \bar{x}_1 = 102300, \bar{x}_2 = 98500, s_1 = 5700$, and $s_2 = 3800$.

(a) $z_{0.005} = 2.575$. Hence,

$$(102300 - 98500) \pm (2.575)\sqrt{\frac{5700^2}{300} + \frac{3800^2}{300}} = 3800 \pm 1018.46,$$

which yields $2781.54 < \mu_1 - \mu_2 < 4818.46$. There is a significant difference in salaries between the two regions.

(b) Since the sample sizes are large enough, it is not necessary to assume the normality due to the Central Limit Theorem.

(c) We assumed that the two variances are not equal. Here we are going to obtain a 95% confidence interval for the ratio of the two variances. It is known that $f_{0.025}(299, 299) = 1.255$. So,

$$\left(\frac{5700}{3800}\right)^2 \frac{1}{1.255} < \frac{\sigma_1^2}{\sigma_2^2} < \left(\frac{5700}{3800}\right)^2 (1.255), \text{ or } 1.793 < \frac{\sigma_1^2}{\sigma_2^2} < 2.824.$$

Since the confidence interval does not contain 1, the difference between the variances is significant.

5.68 The error in estimation, with 95% confidence, is $(1.96)(4000)\sqrt{\frac{2}{n}}$. Equating this quantity to 1000, we obtain

$$n = 2\left(\frac{(1.96)(4000)}{1000}\right)^2 = 123,$$

when round up. Hence, the sample sizes in Review Exercise 9.101 is sufficient to produce a 95% confidence interval on $\mu_1 - \mu_2$ having a width of \$1,000.

5.69 $n = 300, \bar{x} = 6.5$ and $s = 2.5$. Also, $1 - \alpha = 0.99$ and $1 - \gamma = 0.95$. Using Table A.7, $k = 2.522$. So, the limit of the one-sided tolerance interval is $6.5 + (2.522)(2.5) = 12.805$. Since this interval contains 10, the claim by the union leaders appears valid.

5.70 $n = 30, x = 8$, and $z_{0.025} = 1.96$. So,

$$\frac{4}{15} \pm (1.96)\sqrt{\frac{(4/15)(11/15)}{30}} = \frac{4}{15} \pm 0.158,$$

which yields $0.108 < p < 0.425$.

5.71 $n = \frac{(1.96)^2(4/15)(11/15)}{0.05^2} = 301$, when round up.

5.72 $n_1 = n_2 = 100, \hat{p}_1 = 0.1$, and $\hat{p}_2 = 0.06$.

(a) $z_{0.025} = 1.96$. So,

$$(0.1 - 0.06) \pm (1.96)\sqrt{\frac{(0.1)(0.9)}{100} + \frac{(0.06)(0.94)}{100}} = 0.04 \pm 0.075,$$

which yields $-0.035 < p_1 - p_2 < 0.115$.

(b) Since the confidence interval contains 0, it does not show sufficient evidence that $p_1 > p_2$.

5.73 $n = 20$ and $s^2 = 0.045$. It is known that $\chi^2_{0.025} = 32.825$ and $\chi^2_{0.975} = 8.907$ with 19 degrees of freedom. Hence the 95% confidence interval for σ^2 can be expressed as

$$\frac{(19)(0.045)}{32.825} < \sigma^2 < \frac{(19)(0.045)}{8.907}, \text{ or } 0.012 < \sigma^2 < 0.045.$$

Therefore, the 95% confidence interval for σ can be approximated as

$$0.110 < \sigma < 0.212.$$

Since 0.3 falls outside of the confidence interval, there is strong evidence that the process has been improved in variability.

5.74 $n_A = n_B = 15, \bar{y}_A = 87, s_A = 5.99, \bar{y}_B = 75, s_B = 4.85, s_p = 5.450$, and $t_{0.025} = 2.048$ with 28 degrees of freedom. So,

$$(87 - 75) \pm (2.048)(5.450)\sqrt{\frac{1}{15} + \frac{1}{15}} = 12 \pm 4.076,$$

which yields $7.924 < \mu_A - \mu_B < 16.076$. Apparently, the mean operating costs of type A engines are higher than those of type B engines.

5.75 $n = 15, \bar{x} = 3.2$, and $s = 0.6$.

(a) $t_{0.01} = 2.624$ with 14 degrees of freedom. So, a 99% left-sided confidence interval has an upper bound of $3.2 + (2.624)\frac{0.6}{\sqrt{15}} = 3.607$ seconds. We assumed normality in the calculation.

(b) $3.2 + (2.624)(0.6)\sqrt{1 + \frac{1}{15}} = 4.826$. Still, we need to assume normality in the distribution.

(c) $1 - \alpha = 0.99$ and $1 - \gamma = 0.95$. So, $k = 3.520$ with $n = 15$. So, the upper bound is $3.2 + (3.520)(0.6) = 5.312$. Hence, we are 99% confident to claim that 95% of the pilot will have reaction time less than 5.312 seconds.

5.76 $n = 400, x = 17$, so $\hat{p} = \frac{17}{400} = 0.0425$.

(a) $z_{0.025} = 1.96$. So,

$$0.0425 \pm (1.96)\sqrt{\frac{(0.0425)(0.9575)}{400}} = 0.0425 \pm 0.0198,$$

which yields $0.0227 < p < 0.0623$.

(b) $z_{0.05} = 1.645$. So, the upper bound of a left-sided 95% confidence interval is $0.0425 + (1.645)\sqrt{\frac{(0.0425)(0.9575)}{400}} = 0.0591$.

(c) Using both intervals, we do not have evidence to dispute suppliers' claim.

Chapter 6

One- and Two-Sample Tests of Hypotheses

6.1 (a) Conclude that fewer than 30% of the public are allergic to some cheese products when, in fact, 30% or more are allergic.

(b) Conclude that at least 30% of the public are allergic to some cheese products when, in fact, fewer than 30% are allergic.

6.2 (a) The training course is effective.

(b) The training course is effective.

6.3 (a) The firm is not guilty.

(b) The firm is guilty.

6.4 (a) $\alpha = P(X \leq 3 \mid p = 0.6) = 0.0548$.

(b) $\beta = P(X > 3 \mid p = 0.3) = 1 - 0.6496 = 0.3504$.
$\beta = P(X > 3 \mid p = 0.4) = 1 - 0.3823 = 0.6177$.
$\beta = P(X > 3 \mid p = 0.5) = 1 - 0.1719 = 0.8281$.

6.5 (a) $\alpha = P(X \leq 24 \mid p = 0.6) = P(Z < -1.59) = 0.0559$.

(b) $\beta = P(X > 24 \mid p = 0.3) = P(Z > 2.93) = 1 - 0.9983 = 0.0017$.
$\beta = P(X > 24 \mid p = 0.4) = P(Z > 1.30) = 1 - 0.9032 = 0.0968$.
$\beta = P(X > 24 \mid p = 0.5) = P(Z > -0.14) = 1 - 0.4443 = 0.5557$.

6.6 (a) $\alpha = P(X \leq 5 \mid p = 0.6) + P(X \geq 13 \mid p = 0.6) = 0.0338 + (1 - 0.9729) = 0.0609$.

(b) $\beta = P(6 \leq X \leq 12 \mid p = 0.5) = 0.9963 - 0.1509 = 0.8454$.
$\beta = P(6 \leq X \leq 12 \mid p = 0.7) = 0.8732 - 0.0037 = 0.8695$.

(c) This test procedure is not good for detecting differences of 0.1 in p.

6.7 (a) $\alpha = P(X < 110 \mid p = 0.6) + P(X > 130 \mid p = 0.6) = P(Z < -1.52) + P(Z > 1.52) = 2(0.0643) = 0.1286$.

(b) $\beta = P(110 < X < 130 \mid p = 0.5) = P(1.34 < Z < 4.31) = 0.0901$.
$\beta = P(110 < X < 130 \mid p = 0.7) = P(-4.71 < Z < -1.47) = 0.0708$.

(c) The probability of a Type I error is somewhat high for this procedure, although Type II errors are reduced dramatically.

6.8 (a) $n = 7$, $p = 0.4$, $\alpha = P(X \leq 2) = 0.4199$.

(b) $n = 7$, $p = 0.3$, $\beta = P(X \geq 3) = 1 - P(X \leq 2) = 1 - 0.6471 = 0.3529$.

6.9 (a) $n = 12$, $p = 0.7$, and $\alpha = P(X \geq 11) = 0.0712 + 0.0138 = 0.0850$.

(b) $n = 12$, $p = 0.9$, and $\beta = P(X \leq 10) = 0.3410$.

6.10 (a) $n = 100$, $p = 0.7$, $\mu = np = 70$, and $\sigma = \sqrt{npq} = \sqrt{(100)(0.7)(0.3)} = 2.73$. Hence $z = \frac{82.5-70}{4.583} = 0.3410$. Therefore,

$$\alpha = P(X > 82) = P(Z > 2.73) = 1 - 0.9968 = 0.0032.$$

(b) $n = 100$, $p = 0.9$, $\mu = np = 90$, and $\sigma = \sqrt{npq} = \sqrt{(100)(0.9)(0.1)} = 3$. Hence $z = \frac{82.5-90}{3} = -2.5$. So,

$$\beta = P(X \leq 82) = P(Z < -2.5) = 0.0062.$$

6.11 (a) $n = 70$, $p = 0.4$, $\mu = np = 28$, and $\sigma = \sqrt{npq} = 4.099$, with $z = \frac{23.5-28}{4.099} = -1.10$. Then $\alpha = P(X < 24) = P(Z < -1.10) = 0.1357$.

(b) $n = 70$, $p = 0.3$, $\mu = np = 21$, and $\sigma = \sqrt{npq} = 3.834$, with $z = \frac{23.5-21}{3.834} = 0.65$ Then $\beta = P(X \geq 24) = P(Z > 0.65) = 0.2578$.

6.12 (a) $n = 400$, $p = 0.6$, $\mu = np = 240$, and $\sigma = \sqrt{npq} = 9.798$, with

$$z_1 = \frac{259.5 - 240}{9.978} = 1.990, \quad \text{and} \quad z_2 = \frac{220.5 - 240}{9.978} = -1.990.$$

Hence,

$$\alpha = 2P(Z < -1.990) = (2)(0.0233) = 0.0466.$$

(b) When $p = 0.48$, then $\mu = 192$ and $\sigma = 9.992$, with

$$z_1 = \frac{220.5 - 192}{9.992} = 2.852, \quad \text{and} \quad z_2 = \frac{259.5 - 192}{9.992} = 6.755.$$

Therefore,

$$\beta = P(2.852 < Z < 6.755) = 1 - 0.9978 = 0.0022.$$

6.13 From Exercise 6.12(a) we have $\mu = 240$ and $\sigma = 9.798$. We then obtain

$$z_1 = \frac{214.5 - 240}{9.978} = -2.60, \quad \text{and} \quad z_2 = \frac{265.5 - 240}{9.978} = 2.60.$$

So

$$\alpha = 2P(Z < -2.60) = (2)(0.0047) = 0.0094.$$

Also, from Exercise 6.12(b) we have $\mu = 192$ and $\sigma = 9.992$, with

$$z_1 = \frac{214.5 - 192}{9.992} = 2.25, \quad \text{and} \quad z_2 = \frac{265.5 - 192}{9.992} = 7.36.$$

Therefore,

$$\beta = P(2.25 < Z < 7.36) = 1 - 0.9878 = 0.0122.$$

6.14 (a) $n = 50$, $\mu = 15$, $\sigma = 0.5$, and $\sigma_{\bar{X}} = \frac{0.5}{\sqrt{50}} = 0.071$, with $z = \frac{14.9-15}{0.071} = -1.41$. Hence, $\alpha = P(Z < -1.41) = 0.0793$.

(b) If $\mu = 14.8$, $z = \frac{14.9-14.8}{0.071} = 1.41$. So, $\beta = P(Z > 1.41) = 0.0793$.
If $\mu = 14.9$, then $z = 0$ and $\beta = P(Z > 0) = 0.5$.

6.15 (a) $\mu = 200$, $n = 9$, $\sigma = 15$ and $\sigma_{\bar{X}} = \frac{15}{3} = 5$. So,

$$z_1 = \frac{191 - 200}{5} = -1.8, \quad \text{and} \quad z_2 = \frac{209 - 200}{5} = 1.8,$$

with $\alpha = 2P(Z < -1.8) = (2)(0.0359) = 0.0718$.

(b) If $\mu = 215$, then $z_1 = \frac{191-215}{5} = -4.8$ and $z_2 = \frac{209-215}{5} = -1.2$, with

$$\beta = P(-4.8 < Z < -1.2) = 0.1151 - 0 = 0.1151.$$

6.16 (a) When $n = 25$, then $\sigma_{\bar{X}} = \frac{15}{5} = 3$, with $\mu = 200$. Hence

$$z_1 = \frac{191 - 200}{3} = -3, \quad \text{and} \quad z_2 = \frac{209 - 200}{3} = 3,$$

with $\alpha = 2P(Z < -3) = (2)(0.0013) = 0.0026$.

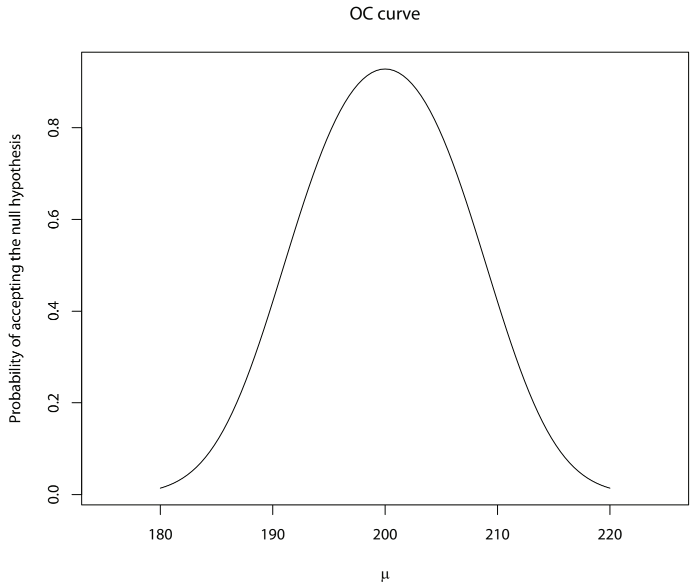
(b) When $\mu = 215$, then $z - 1 = \frac{191-215}{3} = -8$ and $z_2 = \frac{209-215}{3} = -2$, with

$$\beta = P(-8 < Z < -2) = 0.0228 - 0 = 0.0228.$$

6.17 (a) $n = 50$, $\mu = 5000$, $\sigma = 120$, and $\sigma_{\bar{X}} = \frac{120}{\sqrt{50}} = 16.971$, with $z = \frac{4970-5000}{16.971} = -1.77$ and $\alpha = P(Z < -1.77) = 0.0384$.

(b) If $\mu = 4970$, then $z = 0$ and hence $\beta = P(Z > 0) = 0.5$.
If $\mu = 4960$, then $z = \frac{4970-4960}{16.971} = 0.59$ and $\beta = P(Z > 0.59) = 0.2776$.

6.18 The OC curve is shown next.



6.19 The hypotheses are

$$H_0 : \mu = 40 \text{ months},$$
$$H_1 : \mu < 40 \text{ months}.$$

Now, $z = \frac{38-40}{5.8/\sqrt{64}} = -2.76$, and P-value$= P(Z < -2.76) = 0.0029$. Decision: reject H_0.

6.20 The hypotheses are

$$H_0 : \mu = 5.5,$$
$$H_1 : \mu < 5.5.$$

Now, $z = \frac{5.23-5.5}{0.24/\sqrt{64}} = -9.0$, and P-value$= P(Z < -9.0) \approx 0$. The White Cheddar Popcorn, on average, weighs less than 5.5oz.

6.21 The hypotheses are

$$H_0 : \mu = 800,$$
$$H_1 : \mu \neq 800.$$

Now, $z = \frac{788-800}{40/\sqrt{30}} = -1.64$, and P-value$= 2P(Z < -1.64) = (2)(0.0505) = 0.1010$. Hence, the mean is not significantly different from 800 for $\alpha < 0.101$.

6.22 The hypotheses are

$$H_0 : \mu = 8,$$
$$H_1 : \mu > 8.$$

Now, $z = \frac{8.5-8}{2.25/\sqrt{225}} = 3.33$, and P-value$= P(Z > 3.33) = 0.0004$. Decision: Reject H_0 and conclude that men who use TM, on average, mediate more than 8 hours per week.

6.23 The hypotheses are

$$H_0 : \mu = 10,$$
$$H_1 : \mu \neq 10.$$

$\alpha = 0.01$ and $df = 9$.
Critical region: $t < -3.25$ or $t > 3.25$.
Computation: $t = \frac{10.06-10}{0.246/\sqrt{10}} = 0.77$.
Decision: Fail to reject H_0.

6.24 The hypotheses are

$$H_0 : \mu = 162.5 \text{ centimeters},$$
$$H_1 : \mu \neq 162.5 \text{ centimeters}.$$

Now, $z = \frac{165.2-162.5}{6.9/\sqrt{50}} = 2.77$, and P-value$= 2P(Z > 2.77) = (2)(0.0028) = 0.0056$. Decision: reject H_0 and conclude that $\mu \neq 162.5$.

6.25 The hypotheses are

$$H_0 : \mu = 20,000 \text{ kilometers},$$
$$H_1 : \mu > 20,000 \text{ kilometers}.$$

Now, $z = \frac{23,500-20,000}{3900/\sqrt{100}} = 8.97$, and P-value$= P(Z > 8.97) \approx 0$. Decision: reject H_0 and conclude that $\mu \neq 20,000$ kilometers.

6.26 The hypotheses are

$$H_0 : \mu = 220 \text{ milligrams},$$
$$H_1 : \mu > 220 \text{ milligrams}.$$

$\alpha = 0.05$ and $df = 19$.
Critical region: $t > 1.729$.
Computation: $t = \frac{244-220}{24.5/\sqrt{20}} = 4.38$.
Decision: Reject H_0 and claim $\mu > 220$ milligrams.

6.27 The hypotheses are

$$H_0 : \mu_1 = \mu_2,$$
$$H_1 : \mu_1 > \mu_2.$$

Since $s_p = \sqrt{\frac{(29)(10.5)^2+(29)(10.2)^2}{58}} = 10.35$, then

$$P\left[T > \frac{34.0}{10.35\sqrt{1/30 + 1/30}}\right] = P(Z > 12.72) \approx 0.$$

Hence, the conclusion is that running increases the mean RMR in older women.

6.28 The hypotheses are

$$H_0: \mu_C = \mu_A,$$
$$H_1: \mu_C > \mu_A,$$

with $s_p = \sqrt{\frac{(24)(1.5)^2+(24)(1.25)^2}{48}} = 1.3807$. We obtain $t = \frac{20.0-12.0}{1.3807\sqrt{2/25}} = 20.48$. Since $P(T > 20.48) \approx 0$, we conclude that the mean percent absorbency for the cotton fiber is significantly higher than the mean percent absorbency for acetate.

6.29 The hypotheses are

$$H_0: \mu = 35 \text{ minutes},$$
$$H_1: \mu < 35 \text{ minutes}.$$

$\alpha = 0.05$ and $df = 19$.
Critical region: $t < -1.729$.
Computation: $t = \frac{33.1-35}{4.3/\sqrt{20}} = -1.98$.
Decision: Reject H_0 and conclude that it takes less than 35 minutes, on the average, to take the test.

6.30 The hypotheses are

$$H_0: \mu_1 = \mu_2,$$
$$H_1: \mu_1 \neq \mu_2.$$

Since the variances are known, we obtain $z = \frac{81-76}{\sqrt{5.2^2/25+3.4^2/36}} = 4.22$. So, P-value≈ 0 and we conclude that $\mu_1 > \mu_2$.

6.31 The hypotheses are

$$H_0: \mu_A - \mu_B = 12 \text{ kilograms},$$
$$H_1: \mu_A - \mu_B > 12 \text{ kilograms}.$$

$\alpha = 0.05$.
Critical region: $z > 1.645$.
Computation: $z = \frac{(86.7-77.8)-12}{\sqrt{(6.28)^2/50+(5.61)^2/50}} = -2.60$. So, fail to reject H_0 and conclude that the average tensile strength of thread A does not exceed the average tensile strength of thread B by 12 kilograms.

6.32 The hypotheses are

$$H_0: \mu_1 - \mu_2 = \$2,000,$$
$$H_1: \mu_1 - \mu_2 > \$2,000.$$

$\alpha = 0.01$.
Critical region: $z > 2.33$.
Computation: $z = \frac{(70750-65200)-2000}{\sqrt{(6000)^2/200+(5000)^2/200}} = 6.43$, with a P-value$= P(Z > 6.43) \approx 0$. Reject H_0 and conclude that the mean salary for associate professors in research institutions is \$2000 higher than for those in other institutions.

6.33 The hypotheses are

$$H_0: \mu_1 - \mu_2 = 0.5 \text{ micromoles per 30 minutes,}$$
$$H_1: \mu_1 - \mu_2 > 0.5 \text{ micromoles per 30 minutes.}$$

$\alpha = 0.01$.
Critical region: $t > 2.485$ with 25 degrees of freedom.
Computation: $s_p^2 = \frac{(14)(1.5)^2+(11)(1.2)^2}{25} = 1.8936$, and $t = \frac{(8.8-7.5)-0.5}{\sqrt{1.8936}\sqrt{1/15+1/12}} = 1.50$. Do not reject H_0.

6.34 The hypotheses are

$$H_0: \mu_1 - \mu_2 = 8,$$
$$H_1: \mu_1 - \mu_2 < 8.$$

Computation: $s_p^2 = \frac{(10)(4.7)^2+(16)(6.1)^2}{26} = 31.395$, and $t = \frac{(85-79)-8}{\sqrt{31.395}\sqrt{1/11+1/17}} = -0.92$.
Using 28 degrees of freedom and Table A.4, we obtain that $0.15 < P$-value < 0.20.
Decision: Do not reject H_0.

6.35 Let group 1 is the "no treatment" and group the "treatment." The hypotheses are

$$H_0: \mu_1 - \mu_2 = 0,$$
$$H_1: \mu_1 - \mu_2 < 0.$$

$\alpha = 0.05$
Critical region: $t < -1.895$ with 7 degrees of freedom.
Computation: $s_p = \sqrt{\frac{(3)(1.363)+(4)(3.883)}{7}} = 1.674$, and $t = \frac{2.075-2.860}{1.674\sqrt{1/4+1/5}} = -0.70$.
Decision: Do not reject H_0.

6.36 The hypotheses are

$$H_0: \mu_1 = \mu_2,$$
$$H_1: \mu_1 \neq \mu_2.$$

Computation: $s_p = \sqrt{\frac{5100^2+5900^2}{2}} = 5515$, and $t = \frac{37{,}900-39{,}800}{5515\sqrt{1/12+1/12}} = -0.84$.
Using 22 degrees of freedom and since $0.20 < P(T < -0.84) < 0.3$, we obtain $0.4 <$ P-value < 0.6. Decision: Do not reject H_0.

6.37 The hypotheses are

$$H_0 : \mu_1 - \mu_2 = 4 \text{ kilometers},$$
$$H_1 : \mu_1 - \mu_2 \neq 4 \text{ kilometers}.$$

$\alpha = 0.10$ and the critical regions are $t < -1.725$ or $t > 1.725$ with 20 degrees of freedom.
Computation: $t = \frac{(16-11)-4}{(0.915)\sqrt{1/12+1/10}} = 2.55$.
Decision: Reject H_0.

6.38 The hypotheses are

$$H_0 : \mu_1 - \mu_2 = 8,$$
$$H_1 : \mu_1 - \mu_2 < 8.$$

$\alpha = 0.05$ and the critical region is $t < -1.714$ with 23 degrees of freedom.
Computation: $s_p = \sqrt{\frac{(9)(3.2)^2+(14)(2.8)^2}{23}} = 2.963$, and $t = \frac{5.5-8}{2.963\sqrt{1/10+1/15}} = -2.07$.
Decision: Reject H_0 and conclude that $\mu_1 - \mu_2 < 8$ months.

6.39 The hypotheses are

$$H_0 : \mu_{II} - \mu_I = 10,$$
$$H_1 : \mu_{II} - \mu_I > 10.$$

$\alpha = 0.1$.
Degrees of freedom is calculated as

$$v = \frac{(78.8/5 + 913.333/7)^2}{(78.8/5)^2/4 + (913/333/7)^2/6} = 7.38,$$

hence we use 7 degrees of freedom with the critical region $t > 1.415$.
Computation: $t = \frac{(110-97.4)-10}{\sqrt{78.800/5+913.333/7}} = 0.22$.
Decision: Fail to reject H_0.

6.40 The hypotheses are

$$H_0 : \mu_S = \mu_N,$$
$$H_1 : \mu_S \neq \mu_N.$$

Degrees of freedom is calculated as

$$v = \frac{(0.391478^2/8 + 0.214414^2/24)^2}{(0.391478^2/8)^2/7 + (0.214414^2/24)^2/23} = 8.44,$$

which is rounded down to 8.
Computation: $t = \frac{0.97625-0.91583}{\sqrt{0.391478^2/8+0.214414^2/24}} = -0.42$. Since $0.3 < P(T < -0.42) < 0.4$, we obtain $0.6 < P\text{-value} < 0.8$.
Decision: Fail to reject H_0.

6.41 The hypotheses are

$$H_0 : \mu_1 = \mu_2,$$
$$H_1 : \mu_1 \neq \mu_2.$$

$\alpha = 0.05$.
Degrees of freedom is calculated as

$$v = \frac{(7874.329^2/16 + 2479.503^2/12)^2}{(7874.329^2/16)^2/15 + (2479.503^2/12)^2/11} = 19 \text{ degrees of freedom.}$$

Critical regions $t < -2.093$ or $t > 2.093$.
Computation: $t = \frac{9897.500-4120.833}{\sqrt{7874.329^2/16+2479.503^2/12}} = 2.76$.
Decision: Reject H_0 and conclude that $\mu_1 > \mu_2$.

6.42 The hypotheses are

$$H_0 : \mu_1 = \mu_2,$$
$$H_1 : \mu_1 \neq \mu_2.$$

$\alpha = 0.05$.
Critical regions $t < -2.776$ or $t > 2.776$, with 4 degrees of freedom.
Computation: $\bar{d} = -0.1$, $s_d = 0.1414$, so $t = \frac{-0.1}{0.1414/\sqrt{5}} = -1.58$.
Decision: Do not reject H_0 and conclude that the two methods are not significantly different.

6.43 The hypotheses are

$$H_0 : \mu_1 = \mu_2,$$
$$H_1 : \mu_1 < \mu_2.$$

Computation: $\bar{d} = -54.13$, $s_d = 83.002$, $t = \frac{-54.13}{83.002/\sqrt{15}} = -2.53$, and $0.01 < P\text{-value} < 0.015$ with 14 degrees of freedom.
Decision: Reject H_0.

6.44 The hypotheses are

$$H_0 : \mu_1 = \mu_2,$$
$$H_1 : \mu_1 \neq \mu_2.$$

$\alpha = 0.05$.
Critical regions are $t < -2.365$ or $t > 2.365$ with 7 degrees of freedom.
Computation: $\bar{d} = 198.625$, $s_d = 210.165$, $t = \frac{198.625}{210.165/\sqrt{8}} = 2.67$.
Decision: Reject H_0; length of storage influences sorbic acid residual concentrations.

6.45 The hypotheses are

$$H_0: \mu_1 = \mu_2,$$
$$H_1: \mu_1 > \mu_2.$$

Computation: $\bar{d} = 0.1417$, $s_d = 0.198$, $t = \frac{0.1417}{0.198/\sqrt{12}} = 2.48$ and $0.015 < P\text{-value} < 0.02$ with 11 degrees of freedom.
Decision: Reject H_0 when a significance level is above 0.02.

6.46 The hypotheses are

$$H_0: \mu_1 - \mu_2 = 4.5 \text{ kilograms},$$
$$H_1: \mu_1 - \mu_2 < 4.5 \text{ kilograms}.$$

Computation: $\bar{d} = 3.557$, $s_d = 2.776$, $t = \frac{3.557-4.5}{2.778/\sqrt{7}} = -0.896$, and $0.2 < P\text{-value} < 0.3$ with 6 degrees of freedom.
Decision: Do not reject H_0.

6.47 $n = \frac{(1.645+1.282)^2(0.24)^2}{0.3^2} = 5.48$. The sample size needed is 6.

6.48 $\beta = 0.1$, $\sigma = 5.8$, $\delta = 35.9 - 40 = -4.1$. Assume $\alpha = 0.05$ then $z_{0.05} = 1.645$, $z_{0.10} = 1.28$. Therefore,

$$n = \frac{(1.645 + 1.28)^2(5.8)^2}{(-4.1)^2} = 17.12 \approx 18 \text{ due to round up.}$$

6.49 $1 - \beta = 0.95$ so $\beta = 0.05$, $\delta = 3.1$ and $z_{\alpha/2} = z_{0.01} = 2.33$. Therefore,

$$n = \frac{(1.645 + 2.33)^2(6.9)^2}{3.1^2} = 78.28 \approx 79 \text{ due to round up.}$$

6.50 $\beta = 0.05$, $\delta = 8$, $\alpha = 0.05$, $z_{0.05} = 1.645$, $\sigma_1 = 6.28$ and $\sigma_2 = 5.61$. Therefore,

$$n = \frac{(1.645 + 1.645)^2(6.28^2 + 5.61^2)}{8^2} = 11.99 \approx 12 \text{ due to round up.}$$

6.51 To calculate the sample size, we have

$$n = \frac{(1.645 + 0.842)^2(2.25)^2}{[(1.2)(2.25)]^2} = 4.29.$$

The sample size would be 5.

6.52 Using paired T-test, we find out $t = 2.4$ with 8 degrees of freedom. So, $0.02 < P\text{-value} < 0.025$. Reject H_0; breathing frequency significantly higher in the presence of CO.

6.53 (a) The hypotheses are

$$H_0 : M_{\text{hot}} - M_{\text{cold}} = 0,$$
$$H_1 : M_{\text{hot}} - M_{\text{cold}} \neq 0.$$

(b) Use paired T-test we have $\bar{d} = 726$, $s_d = 207$ with $df = -7$. We obtain $t = 0.99$ with $0.3 <$ 2-sided P-value < 0.4. Hence, fail to reject H_0.

6.54 $\sigma = 1.25$, $\alpha = 0.05$, $\beta = 0.1$, $\delta = 0.5$, with $z_{\alpha/2} = z_{0.025} = 1.96$ and $z_\beta = z_{0.1} = 1.282$. Hence,

$$n \approx \frac{(z_{\alpha/2} + z_\beta)^2 \sigma^2}{\delta^2} = \frac{(1.96 + 1.282)^2 (1.25^2)}{0.5^2} = 65.7.$$

We round up the value to $n = 66$.

6.55 The hypotheses are

$$H_0 : p = 0.40,$$
$$H_1 : p > 0.40.$$

Denote by X for those who choose lasagna.

$$P\text{-value} = P(X \geq 9 \mid p = 0.40) = 0.4044.$$

The claim that $p = 0.40$ is not refuted.

6.56 The hypotheses are

$$H_0 : p = 0.40,$$
$$H_1 : p > 0.40.$$

$\alpha = 0.05$.
Test statistic: binomial variable X with $p = 0.4$ and $n = 15$.
Computation: $x = 8$. Therefore, from Table A.1,

$$P\text{-value} = P(X \geq 8 \mid p = 0.4) = 1 - P(X \leq 7 \mid p = 0.4) = 0.2131,$$

which is larger than 0.05.
Decision: Do not reject H_0.

6.57 The hypotheses are

$$H_0 : p = 0.8,$$
$$H_1 : p > 0.8.$$

$\alpha = 0.04$.
Critical region: $z > 1.75$.
Computation: $z = \frac{250-(300)(0.8)}{\sqrt{(300)(0.8)(0.2)}} = 1.44$.
Decision: Fail to reject H_0; it cannot conclude that the new missile system is more accurate.

6.58 The hypotheses are

$$H_0: p = 0.6,$$
$$H_1: p < 0.6.$$

So

$$P\text{-value} \approx P\left(Z < \frac{110 - (200)(0.6)}{\sqrt{(200)(0.6)(0.4)}}\right) = P(Z < -1.44) = 0.0749.$$

Decision: Fail to reject H_0.

6.59 The hypotheses are

$$H_0: p = 0.2,$$
$$H_1: p < 0.2.$$

Then

$$P\text{-value} \approx P\left(Z < \frac{136 - (1000)(0.2)}{\sqrt{(1000)(0.2)(0.8)}}\right) = P(Z < -5.06) \approx 0.$$

Decision: Reject H_0; less than 1/5 of the homes in the city are heated by oil.

6.60 The hypotheses are

$$H_0: p = 0.25,$$
$$H_1: p > 0.25.$$

$\alpha = 0.05$.
Computation:

$$P\text{-value} \approx P\left(Z > \frac{28 - (90)(0.25)}{\sqrt{(90)(0.25)(0.75)}}\right) = P(Z > 1.34) = 0.091.$$

Decision: Fail to reject H_0; No sufficient evidence to conclude that $p > 0.25$.

6.61 The hypotheses are

$$H_0: p_1 = p_2,$$
$$H_1: p_1 > p_2.$$

Computation: $\hat{p} = \frac{29+56}{120+280} = 0.2125$, $z = \frac{(29/120)-(56/280))}{\sqrt{(0.2125)(0.7875)(1/120+1/280)}} = 0.93$, with P-value $= P(Z > 0.93) = 0.1762$.
Decision: Fail to reject H_0. There is no significant evidence to conclude that the new medicine is more effective.

6.62 The hypotheses are

$$H_0: p = 0.25,$$
$$H_1: p > 0.25.$$

$\alpha = 0.05$.
Critical region: $z > 1.645$.
Computation: $z = \frac{16-(48)(0.25)}{\sqrt{(48)(0.25)(0.75)}} = 1.333$.
Decision: Fail to reject H_0. On the other hand, we can calculate

$$P\text{-value} = P(Z > 1.33) = 0.0918.$$

6.63 The hypotheses are

$$H_0: p_1 = p_2,$$
$$H_1: p_1 \neq p_2.$$

Computation: $\hat{p} = \frac{63+59}{100+125} = 0.5422$, $z = \frac{(63/100)-(59/125)}{\sqrt{(0.5422)(0.4578)(1/100+1/125)}} = 2.36$, with P-value $= 2P(Z > 2.36) = 0.0182$.
Decision: Reject H_0 at level 0.0182. The proportion of urban residents who favor the nuclear plant is larger than the proportion of suburban residents who favor the nuclear plant.

6.64 The hypotheses are

$$H_0: p_1 = p_2,$$
$$H_1: p_1 > p_2.$$

Computation: $\hat{p} = \frac{240+288}{300+400} = 0.7543$, $z = \frac{(240/300)-(288/400)}{\sqrt{(0.7543)(0.2457)(1/300+1/400)}} = 2.44$, with P-value $= P(Z > 2.44) = 0.0073$.
Decision: Reject H_0. The proportion of couples married less than 2 years and planning to have children is significantly higher than that of couples married 5 years and planning to have children.

6.65 The hypotheses are

$$H_0: p_U = p_R,$$
$$H_1: p_U > p_R.$$

Computation: $\hat{p} = \frac{20+10}{200+150} = 0.085714$, $z = \frac{(20/200)-(10/150)}{\sqrt{(0.085714)(0.914286)(1/200+1/150)}} = 1.10$, with P-value $= P(Z > 1.10) = 0.1357$.
Decision: Fail to reject H_0. It cannot be shown that breast cancer is more prevalent in the urban community.

6.67 The hypotheses are

$$H_0 : \text{nuts are mixed in the ratio 5:2:2:1},$$
$$H_1 : \text{nuts are not mixed in the ratio 5:2:2:1}.$$

$\alpha = 0.05$.
Critical region: $\chi^2 > 7.815$ with 3 degrees of freedom.
Computation:

Observed	269	112	74	45
Expected	250	100	100	50

$$\chi^2 = \frac{(269-250)^2}{250} + \frac{(112-100)^2}{100} + \frac{(74-100)^2}{100} + \frac{(45-50)^2}{50} = 10.14.$$

Decision: Reject H_0; the nuts are not mixed in the ratio 5:2:2:1.

6.68 The hypotheses are

$$H_0 : \text{Distribution of grades is uniform},$$
$$H_1 : \text{Distribution of grades is not uniform}.$$

$\alpha = 0.05$.
Critical region: $\chi^2 > 9.488$ with 4 degrees of freedom.
Computation: Since $e_i = 20$, for $i = 1, 2, \dots, 5$, then

$$\chi^2 = \frac{(14-20)^2}{20} + \frac{(18-20)^2}{20} + \cdots \frac{(16-20)^2}{20} = 10.0.$$

Decision: Reject H_0; the distribution of grades is not uniform.

6.69 The hypotheses are

$$H_0 : \text{die is balanced},$$
$$H_1 : \text{die is unbalanced}.$$

$\alpha = 0.01$.
Critical region: $\chi^2 > 15.086$ with 5 degrees of freedom.
Computation: Since $e_i = 30$, for $i = 1, 2, \dots, 6$, then

$$\chi^2 = \frac{(28-30)^2}{30} + \frac{(36-30)^2}{30} + \cdots + \frac{(23-30)^2}{30} = 4.47.$$

Decision: Fail to reject H_0; the die is balanced.

6.70 The hypotheses are

$$H_0: \text{Data follows the hypergeometric distribution } h(x; 8, 3, 5),$$

$$H_1: \text{Data does not follows the hypergeometric distribution.}$$

$\alpha = 0.05$.
Computation: $h(0; 8, 3, 5) = 1/56$, $b(1; 8, 3, 5) = 15/56$, $b(2; 8, 3, 5) = 30/56$, and $b(3; 8, 3, 5) = 10/56$. Hence $e_1 = 2$, $e_2 = 30$, $e_3 = 60$ and $e_4 = 20$. Combining the first two classes together, we obtain

$$\chi^2 = \frac{(32-32)^2}{32} + \frac{(55-60)^2}{60} + \frac{(25-20)^2}{20} = 1.67.$$

Critical region: $\chi^2 > 5.991$ with 2 degrees of freedom.
Decision: Fail to reject H_0; the data is from a distribution not significantly different from $h(y; 8, 3, 5)$.

6.71 The hypotheses are

$$H_0: f(x) = g(x; 1/2) \text{ for } x = 1, 2, \ldots,$$
$$H_1: f(x) \neq g(x; 1/2).$$

$\alpha = 0.05$.
Computation: $g(x; 1/2) = \frac{1}{2^x}$, for $x = 1, 2, \ldots, 7$ and $P(X \geq 8) = \frac{1}{2^7}$. Hence $e_1 = 128$, $e_2 = 64$, $e_3 = 32$, $e_4 = 16$, $e_5 = 8$, $e_6 = 4$, $e_7 = 2$ and $e_8 = 2$. Combining the last three classes together, we obtain

$$\chi^2 = \frac{(136-128)^2}{128} + \frac{(60-64)^2}{64} + \frac{(34-32)^2}{32} + \frac{(12-16)^2}{16} + \frac{(9-8)^2}{8} + \frac{(5-8)^2}{8} = 3.125$$

Critical region: $\chi^2 > 11.070$ with 5 degrees of freedom.
Decision: Fail to reject H_0; $f(x) = g(x; 1/2)$, for $x = 1, 2, \ldots$

6.72 The hypotheses are

$$H_0: \text{Distribution of grades is normal } n(x; 65, 21),$$

$$H_1: \text{Distribution of grades is not normal.}$$

$\alpha = 0.05$.
Computation:

z values	$P(Z < z)$	$P(z_{i-1} < Z < z_i)$	e_i		o_i	
$z_1 = \frac{19.5-65}{21} = -2.17$	0.0150	0.0150	0.9		3	
$z_2 = \frac{29.5-65}{21} = -1.69$	0.0454	0.0305	1.8	6.8	2	8
$z_3 = \frac{39.5-65}{21} = -1.21$	0.1131	0.0676	4.1		3	
$z_4 = \frac{49.5-65}{21} = -0.74$	0.2296	0.1165	7.0		4	
$z_5 = \frac{59.5-65}{21} = -0.26$	0.3974	0.1678	10.1		5	
$z_6 = \frac{69.5-65}{21} = 0.21$	0.5832	0.1858	11.1		11	
$z_7 = \frac{79.5-65}{21} = 0.69$	0.7549	0.1717	10.3		14	
$z_8 = \frac{89.5-65}{21} = 1.17$	0.8790	0.1241	7.4		14	
$z_9 = \infty$	1.0000	0.1210	7.3		4	

A goodness-of-fit test with 6 degrees of freedom is based on the following data:

o_i	8	4	5	11	14	14	4
e_i	6.8	7.0	10.1	11.1	10.3	7.4	7.3

Critical region: $\chi^2 > 12.592$.

$$\chi^2 = \frac{(8-6.8)^2}{6.8} + \frac{(4-7.0)^2}{7.0} + \cdots + \frac{(4-7.3)^2}{7.3} = 12.78.$$

Decision: Reject H_0; distribution of grades is not normal.

6.73 The hypotheses are

H_0 : Distribution of nicotine contents is normal $n(x; 1.8, 0.4)$,

H_1 : Distribution of nicotine contents is not normal.

$\alpha = 0.01$.

For the total of 30 observations, let's use 6 equal probability intervals, which would lead to $o_i = 5$ for $i = 1, \ldots, 6$. Using $x = 1.8 + (0.4)(z_{1-i/6})$ for $i = 1, \ldots, 6$, we have the following table.

Range	e_i	o_i	$(e_i - o_i)^2/e_i$
$(-\infty, 1.41]$	5	11	7.2
$(1.41, 1.63]$	5	2	1.8
(1.63,1.8]	5	1	3.2
(1.8,1.97]	5	0	5
(1.8,1.97]	5	0	5
(1.97,2.19]	5	1	3.2
$(2.19, \infty)$	5	15	20
Total	30	30	40.4

$\chi^2 = 40.4$ with critical region: $\chi^2 > 15.086$.
Decision: Reject H_0 and conclude that the distribution is different from $n(x; 1.8, 0.4)$.

6.74 The hypotheses are

H_0 : Presence or absence of hypertension is independent of smoking habits,

H_1 : Presence or absence of hypertension is not independent of smoking habits.

$\alpha = 0.05$.
Critical region: $\chi^2 > 5.991$ with 2 degrees of freedom.
Computation:

Observed and expected frequencies				
	Nonsmokers	Moderate Smokers	Heavy Smokers	Total
Hypertension	21 (33.4)	36 (30.0)	30 (23.6)	87
No Hypertension	48 (35.6)	26 (32.0)	19 (25.4)	93
Total	69	62	49	180

$$\chi^2 = \frac{(21 - 33.4)^2}{33.4} + \cdots + \frac{(19 - 25.4)^2}{25.4} = 14.60.$$

Decision: Reject H_0; presence or absence of hypertension and smoking habits are not independent.

6.75 The hypotheses are

H_0 : A person's gender and time spent watching television are independent,

H_1 : A person's gender and time spent watching television are not independent.

$\alpha = 0.01$.
Critical region: $\chi^2 > 6.635$ with 1 degrees of freedom.
Computation:

Observed and expected frequencies			
	Male	Female	Total
Over 25 hours	15 (20.5)	29 (23.5)	44
Under 25 hours	27 (21.5)	19 (24.5)	46
Total	42	48	90

$$\chi^2 = \frac{(15 - 20.5)^2}{20.5} + \frac{(29 - 23.5)^2}{23.5} + \frac{(27 - 21.5)^2}{21.5} + \frac{(19 - 24.5)^2}{24.5} = 5.4.$$

Decision: Fail to reject H_0; a person's gender and time spent watching television are independent.

6.76 The hypotheses are

H_0 : Size of family is independent of level of education of father,

H_1 : Size of family and the education level of father are not independent.

$\alpha = 0.05$.
Critical region: $\chi^2 > 9.488$ with 4 degrees of freedom.
Computation:

Observed and expected frequencies				
	Number of Children			
Education	0–1	2–3	Over 3	Total
Elementary	14 (18.7)	37 (39.8)	32 (24.5)	83
Secondary	19 (17.6)	42 (37.4)	17 (23.0)	78
College	12 (8.7)	17 (18.8)	10 (11.5)	39
Total	45	96	59	200

$$\chi^2 = \frac{(14-18.7)^2}{18.7} + \frac{(37-39.8)^2}{39.8} + \cdots + \frac{(10-11.5)^2}{11.5} = 7.54.$$

Decision: Fail to reject H_0; size of family is independent of level of education of father.

6.77 The hypotheses are

H_0 : Occurrence of types of crime is independent of city district,
H_1 : Occurrence of types of crime is dependent upon city district.

$\alpha = 0.01$.
Critical region: $\chi^2 > 21.666$ with 9 degrees of freedom.
Computation:

Observed and expected frequencies					
District	Assault	Burglary	Larceny	Homicide	Total
1	162 (186.4)	118 (125.8)	451 (423.5)	18 (13.3)	749
2	310 (380.0)	196 (256.6)	996 (863.4)	25 (27.1)	1527
3	258 (228.7)	193 (154.4)	458 (519.6)	10 (16.3)	919
4	280 (214.9)	175 (145.2)	390 (488.5)	19 (15.3)	864
Total	1010	682	2295	72	4059

$$\chi^2 = \frac{(162-186.4)^2}{186.4} + \frac{(118-125.8)^2}{125.8} + \cdots + \frac{(19-15.3)^2}{15.3} = 124.59.$$

Decision: Reject H_0; occurrence of types of crime is dependent upon city district.

6.78 The hypotheses are

H_0 : The proportions of widows and widowers are equal with respect to the different time period,
H_1 : The proportions of widows and widowers are not equal with respect to the different time period.

$\alpha = 0.05$.
Critical region: $\chi^2 > 5.991$ with 2 degrees of freedom.
Computation:

Observed and expected frequencies			
Years Lived	Widow	Widower	Total
Less than 5	25 (32)	39 (32)	64
5 to 10	42 (41)	40 (41)	82
More than 10	33 (26)	21 (26)	54
Total	100	100	200

$$\chi^2 = \frac{(25-32)^2}{32} + \frac{(39-32)^2}{32} + \cdots + \frac{(21-26)^2}{26} = 5.96.$$

Decision: Fail to reject H_0; the proportions of widows and widowers are equal with respect to the different time period.

6.79 The hypotheses are

H_0 : Proportions of household within each standard of living category are equal,

H_1 : Proportions of household within each standard of living category are not equal.

$\alpha = 0.05$.
Critical region: $\chi^2 > 12.592$ with 6 degrees of freedom.
Computation:

Observed and expected frequencies				
Period	Somewhat Better	Same	Not as Good	Total
1980: Jan.	72 (66.6)	144 (145.2)	84 (88.2)	300
May.	63 (66.6)	135 (145.2)	102 (88.2)	300
Sept.	47 (44.4)	100 (96.8)	53 (58.8)	200
1981: Jan.	40 (44.4)	105 (96.8)	55 (58.8)	200
Total	222	484	294	1000

$$\chi^2 = \frac{(72-66.6)^2}{66.6} + \frac{(144-145.2)^2}{145.2} + \cdots + \frac{(55-58.8)^2}{58.8} = 5.92.$$

From Table A.5, $\chi^2 = 5.9$ with 6 degrees of freedom has a P-value between 0.3 and 0.5.

Decision: Fail to reject H_0; proportions of household within each standard of living category are equal.

6.80 The hypotheses are

H_0 : The three cough remedies are equally effective,

H_1 : The three cough remedies are not equally effective.

$\alpha = 0.05$.
Critical region: $\chi^2 > 9.488$ with 4 degrees of freedom.
Computation:

Observed and expected frequencies				
	NyQuil	Robitussin	Triaminic	Total
No Relief	11 (11)	13 (11)	9 (11)	33
Some Relief	32 (29)	28 (29)	27 (29)	87
Total Relief	7 (10)	9 (10)	14 (10)	30
Total	50	50	50	150

$$\chi^2 = \frac{(11-11)^2}{11} + \frac{(13-11)^2}{11} + \cdots + \frac{(14-10)^2}{10} = 3.81.$$

From Table A.5, $\chi^2 = 3.8$ with 4 degrees of freedom has a P-value between 0.3 and 0.5.

Decision: Fail to reject H_0; the three cough remedies are equally effective.

6.81 The hypotheses are

H_0 : The attitudes among the four counties are homogeneous,
H_1 : The attitudes among the four counties are not homogeneous.

Computation:

Observed and expected frequencies					
	County				
Attitude	Craig	Giles	Franklin	Montgomery	Total
Favor	65 (74.5)	66 (55.9)	40 (37.3)	34 (37.3)	205
Oppose	42 (53.5)	30 (40.1)	33 (26.7)	42 (26.7)	147
No Opinion	93 (72.0)	54 (54.0)	27 (36.0)	24 (36.0)	198
Total	200	150	100	100	550

$$\chi^2 = \frac{(65-74.5)^2}{74.5} + \frac{(66-55.9)^2}{55.9} + \cdots + \frac{(24-36.0)^2}{36.0} = 31.17.$$

Since P-value $= P(\chi^2 > 31.17) < 0.001$ with 6 degrees of freedom, we reject H_0 and conclude that the attitudes among the four counties are not homogeneous.

6.82 The hypotheses are

H_0 : Proportions of voters within each attitude category are the same for each of the three states,
H_1 : Proportions of voters within each attitude category are not the same for each of the three states.

$\alpha = 0.05$.
Critical region: $\chi^2 > 9.488$ with 4 degrees of freedom.
Computation:

Observed and expected frequencies				
	Support	Do not Support	Undecided	Total
Indiana	82 (94)	97 (79)	21 (27)	200
Kentucky	107 (94)	66 (79)	27 (27)	200
Ohio	93 (94)	74 (79)	33 (27)	200
Total	282	237	81	600

$$\chi^2 = \frac{(82-94)^2}{94} + \frac{(97-79)^2}{79} + \cdots + \frac{(33-27)^2}{27} = 12.56.$$

Decision: Reject H_0; the proportions of voters within each attitude category are not the same for each of the three states.

6.83 The hypotheses are

H_0 : Proportions of voters favoring candidate A, candidate B, or undecided are the same for each city,

H_1 : Proportions of voters favoring candidate A, candidate B, or undecided are not the same for each city.

$\alpha = 0.05$.
Critical region: $\chi^2 > 5.991$ with 2 degrees of freedom.
Computation:

Observed and expected frequencies			
	Richmond	Norfolk	Total
Favor A	204 (214.5)	225 (214.5)	429
Favor B	211 (204.5)	198 (204.5)	409
Undecided	85 (81)	77 (81)	162
Total	500	500	1000

$$\chi^2 = \frac{(204-214.5)^2}{214.5} + \frac{(225-214.5)^2}{214.5} + \cdots + \frac{(77-81)^2}{81} = 1.84.$$

Decision: Fail to reject H_0; the proportions of voters favoring candidate A, candidate B, or undecided are not the same for each city.

6.84 The hypotheses are

$$H_0 : p_1 = p_2 = p_3,$$
$$H_1 : p_1, p_2, \text{ and } p_3 \text{ are not all equal.}$$

$\alpha = 0.05$.
Critical region: $\chi^2 > 5.991$ with 2 degrees of freedom.
Computation:

Observed and expected frequencies				
	Denver	Phoenix	Rochester	Total
Watch Soap Operas	52 (48)	31 (36)	37 (36)	120
Do not Watch	148 (152)	119 (114)	113 (114)	380
Total	200	150	150	500

$$\chi^2 = \frac{(52-48)^2}{48} + \frac{(31-36)^2}{36} + \cdots + \frac{(113-114)^2}{114} = 1.39.$$

Decision: Fail to reject H_0; no difference among the proportions.

6.85 (a) $H_0: \ \mu = 21.8$, $H_1: \ \mu \neq 21.8$; critical region in both tails.

(b) $H_0: \ p = 0.2$, $H_1: \ p > 0.2$; critical region in right tail.

(c) $H_0: \ \mu = 6.2$, $H_1: \ \mu > 6.2$; critical region in right tail.

(d) $H_0: \ p = 0.7$, $H_1: \ p < 0.7$; critical region in left tail.

(e) $H_0: \ p = 0.58$, $H_1: \ p \neq 0.58$; critical region in both tails.

(f) $H_0: \ \mu = 340$, $H_1: \ \mu < 340$; critical region in left tail.

6.86 The hypotheses are

$$H_0: p_1 = p_2,$$
$$H_1: p_1 > p_2.$$

$\alpha = 0.01$.
Critical region: $z > 2.33$.
Computation: $\hat{p}_1 = 0.31$, $\hat{p}_2 = 0.24$, $\hat{p} = 0.275$, and

$$z = \frac{0.31 - 0.24}{\sqrt{(0.275)(0.725)(1/100 + 1/100)}} = 1.11.$$

Decision: Fail to reject H_0; proportions are the same.

6.87 The hypotheses are

$$H_0: p_1 = p_2,$$
$$H_1: p_1 > p_2.$$

$\alpha = 0.05$.
Critical region: $z > 1.645$.
Computation: $\hat{p}_1 = 0.24$, $\hat{p}_2 = 0.175$, $\hat{p} = 0.203$, and

$$z = \frac{0.24 - 0.175}{\sqrt{(0.203)(0.797)(1/300 + 1/400)}} = 2.12.$$

Decision: Reject H_0; there is statistical evidence to conclude that more Italians prefer white champagne at weddings.

6.88 From the data, $n = 9$, $\bar{d} = 1.58$, and $s_d = 3.07$. Using paired t-test, we observe that $t = 1.55$ with P-value > 0.05. Hence, the data was not sufficient to show that the oxygen consumptions was higher when there was little or not CO.

6.89 $n_1 = n_2 = 5$, $\bar{x}_1 = 165.0$, $s_1 = 6.442$, $\bar{x}_2 = 139.8$, $s_2 = 12.617$, and $s_p = 10.02$. Hence

$$t = \frac{165 - 139.8}{(10.02)\sqrt{1/5 + 1/5}} = 3.98.$$

This is a one-sided test. Therefore, $0.0025 < P$-value < 0.005 with 8 degrees of freedom. Reject H_0; the speed is increased by using the facilitation tools.

6.90 (a) $H_0: \ p = 0.2$, $H_1: \ p > 0.2$; critical region in right tail.

(b) $H_0: \ \mu = 3$, $H_1: \ \mu \neq 3$; critical region in both tails.

(c) $H_0: \ p = 0.15$, $H_1: \ p < 0.15$; critical region in left tail.

(d) $H_0: \ \mu = \$10$, $H_1: \ \mu > \$10$; critical region in right tail.

(e) $H_0: \ \mu = 9$, $H_1: \ \mu \neq 9$; critical region in both tails.

6.91 The hypotheses are

$$H_0: p_1 = p_2 = p_3,$$
$$H_1: p_1, p_2, \text{ and } p_3 \text{ are not all equal.}$$

$\alpha = 0.01$.
Critical region: $\chi^2 > 9.210$ with 2 degrees of freedom.
Computation:

Observed and expected frequencies				
	Distributor			
Nuts	1	2	3	Total
Peanuts	345 (339)	313 (339)	359 (339)	1017
Other	155 (161)	187 (161)	141 (161)	483
Total	500	500	500	1500

$$\chi^2 = \frac{(345-339)^2}{339} + \frac{(313-339)^2}{339} + \cdots + \frac{(141-161)^2}{161} = 10.19.$$

Decision: Reject H_0; the proportions of peanuts for the three distributors are not equal.

6.92 The hypotheses are

$$H_0: p_1 = p_2 = p_3 = p_4,$$
$$H_1: p_1, p_2, p_3, \text{ and } p_4 \text{ are not all equal.}$$

$\alpha = 0.01$.
Critical region: $\chi^2 > 11.345$ with 3 degrees of freedom.
Computation:

Observed and expected frequencies					
Preference	Maryland	Virginia	Georgia	Alabama	Total
Yes	65 (74)	71 (74)	78 (74)	82 (74)	296
No	35 (26)	29 (26)	22 (26)	18 (26)	104
Total	100	100	100	100	400

$$\chi^2 = \frac{(65-74)^2}{74} + \frac{(71-74)^2}{74} + \cdots + \frac{(18-26)^2}{26} = 8.84.$$

Decision: Fail to reject H_0; the proportions of parents favoring Bibles in elementary schools are the same across states.

6.93 $\bar{d} = -2.905$, $s_d = 3.3557$, and $t = \frac{\bar{d}}{s_d/\sqrt{n}} = -2.12$. Since $0.025 < P(T > 2.12) < 0.05$ with 5 degrees of freedom, we have $0.05 < P$-value < 0.10. There is no significant change in WBC leukograms.

6.94 $n_1 = 15$, $\bar{x}_1 = 156.33$, $s_1 = 33.09$, $n_2 = 18$, $\bar{x}_2 = 170.00$ and $s_2 = 30.79$. First we do the f-test to test equality of the variances. Since $f = \frac{s_1^2}{s_2^2} = 1.16$ and $f_{0.05}(14, 17) \approx 2.31$, we conclude that the two variances are equal.
To test the difference of the means, we first calculate $s_p = 31.85$. Therefore, $t = \frac{156.33-170.00}{(31.85)\sqrt{1/15+1/18}} = -1.23$ with a P-value > 0.10 from Table A.4 with 31 degrees of freedom.
Decision: H_0 cannot be rejected at 0.05 level of significance.

6.95 $n_1 = n_2 = 10$, $\bar{x}_1 = 7.95$, $s_1 = 1.10$, $\bar{x}_2 = 10.26$ and $s_2 = 0.57$. First we do the f-test to test equality of the variances. Since $f = \frac{s_1^2}{s_2^2} = 3.72$ and $f_{0.05}(9, 9) = 3.18$, we conclude that the two variances are not equal at level 0.10.
To test the difference of the means, we first find the degrees of freedom $v = 13$ when round up. Also, $t = \frac{7.95-10.26}{\sqrt{1.10^2/10+0.57^2/10}} = -5.90$ with a P-value < 0.0005.
Decision: Reject H_0; there is a significant difference in the steel rods.

6.96 $n_1 = n_2 = 10$, $\bar{x}_1 = 21.5$, $s_1 = 5.3177$, $\bar{x}_2 = 28.3$ and $s_2 = 5.8699$. Since $f = \frac{s_1^2}{s_2^2} = 0.8207$ and $f_{0.05}(9, 9) = 3.18$, we conclude that the two variances are equal.
$s_p = 5.6001$ and hence $t = \frac{21.5-28.3}{(5.6001)\sqrt{1/10+1/10}} = -2.71$ with $0.005 < P$-value < 0.0075.
Decision: Reject H_0; the high income neighborhood produces significantly more wastewater to be treated.

6.97 $n_1 = n_2 = 16$, $\bar{x}_1 = 48.1875$, $s_1 = 4.9962$, $\bar{x}_2 = 43.7500$ and $s_2 = 4.6833$. Since $f = \frac{s_1^2}{s_2^2} = 1.1381$ and $f_{0.05}(15, 15) = 2.40$, we conclude that the two variances are equal.
$s_p = 4.8423$ and hence $t = \frac{48.1875-43.7500}{(4.8423)\sqrt{1/16+1/16}} = 2.59$. This is a two-sided test. Since $0.005 < P(T > 2.59) < 0.0075$, we have $0.01 < P$-value < 0.015.
Decision: Reject H_0; there is a significant difference in the number of defects.

6.98 The hypotheses are:

$$H_0 : \mu = 24 \times 10^{-4} \text{ gm},$$
$$H_1 : \mu < 24 \times 10^{-4} \text{ gm}.$$

$t = \frac{22.8-24}{4.8/\sqrt{50}} = -1.77$ with $0.025 < P$-value < 0.05. Hence, at significance level of $\alpha = 0.05$, the mean concentration of PCB in malignant breast tissue is less than 24×10^{-4} gm.

Chapter 7

Linear Regression

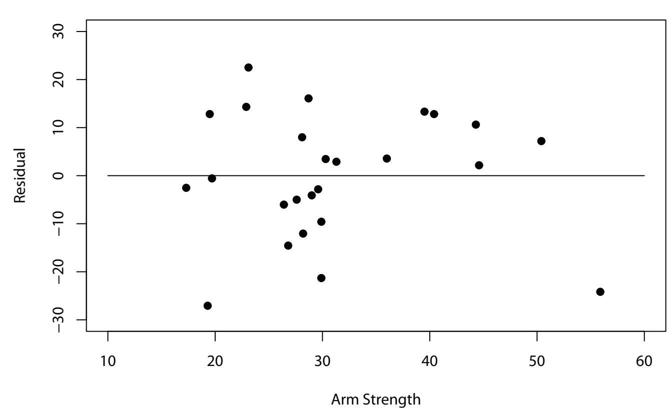
In this chapter, you may find out that many numbers do not really match the numbers in the computing formulas due to rounding issues in parameter estimations. Many of the final result of each question may come from computer output, which was computed with better precision.

7.1 (a) $\sum_i x_i = 778.7$, $\sum_i y_i = 2050.0$, $\sum_i x_i^2 = 26,591.63$, $\sum_i x_i y_i = 65,164.04$, $n = 25$. Therefore,

$$b_1 = \frac{(25)(65,164.04) - (778.7)(2050.0)}{(25)(26,591.63) - (778.7)^2} = 0.5609,$$
$$b_0 = \frac{2050 - (0.5609)(778.7)}{25} = 64.53.$$

(b) Using the equation $\hat{y} = 64.53 + 0.5609x$ with $x = 30$, we find $\hat{y} = 64.53 + (0.5609)(30) = 81.40$.

(c) Residuals appear to be random as desired.



7.2 (a) $\sum_i x_i = 707$, $\sum_i y_i = 658$, $\sum_i x_i^2 = 57,557$, $\sum_i x_i y_i = 53,258$, $n = 9$.

$$b_1 = \frac{(9)(53,258) - (707)(658)}{(9)(57,557) - (707)^2} = 0.7771,$$
$$b_0 = \frac{658 - (0.7771)(707)}{9} = 12.0623.$$

Hence $\hat{y} = 12.0623 + 0.7771x$.

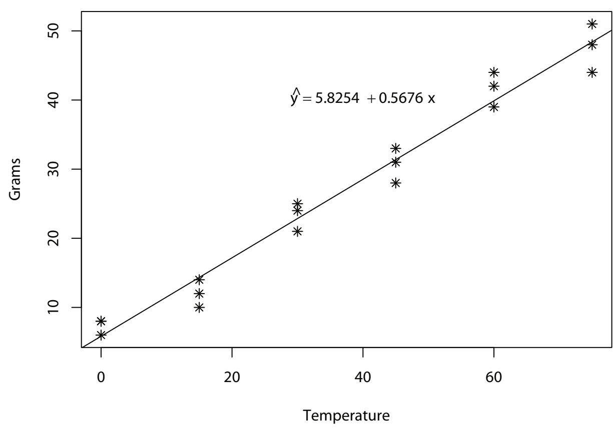
(b) For $x = 85$, $\hat{y} = 12.0623 + (0.7771)(85) = 78$.

7.3 (a) $\sum_i x_i = 675$, $\sum_i y_i = 488$, $\sum_i x_i^2 = 37,125$, $\sum_i x_i y_i = 25,005$, $n = 18$. Therefore,

$$b_1 = \frac{(18)(25,005) - (675)(488)}{(18)(37,125) - (675)^2} = 0.5676,$$
$$b_0 = \frac{488 - (0.5676)(675)}{18} = 5.8261.$$

Hence $\hat{y} = 5.8261 + 0.5676x$

(b) The scatter plot and the regression line are shown below.



(c) For $x = 50$, $\hat{y} = 5.8261 + (0.5676)(50) = 34.21$ grams.

7.4 (a) $\hat{y} = -1.70 + 1.81x$.

(b) $\hat{x} = (54 + 1.71)/1.81 = 30.78$.

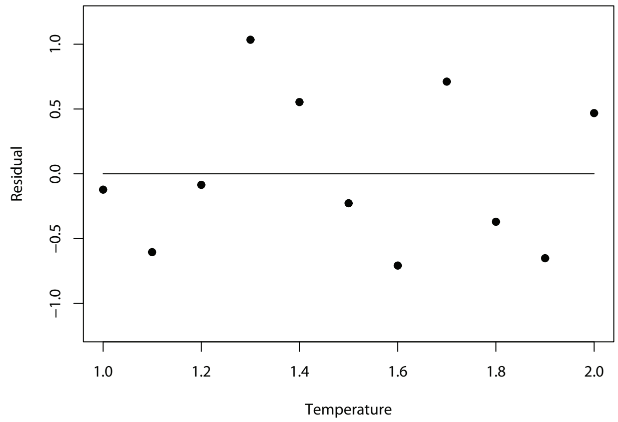
7.5 (a) $\sum_i x_i = 16.5$, $\sum_i y_i = 100.4$, $\sum_i x_i^2 = 25.85$, $\sum_i x_i y_i = 152.59$, $n = 11$. Therefore,

$$b_1 = \frac{(11)(152.59) - (16.5)(100.4)}{(11)(25.85) - (16.5)^2} = 1.8091,$$
$$b_0 = \frac{100.4 - (1.8091)(16.5)}{11} = 6.4136.$$

Hence $\hat{y} = 6.4136 + 1.8091x$

(b) For $x = 1.75$, $\hat{y} = 6.4136 + (1.8091)(1.75) = 9.580$.

(c) Residuals appear to be random as desired.



7.6 (a) $\sum_i x_i = 311.6$, $\sum_i y_i = 297.2$, $\sum_i x_i^2 = 8134.26$, $\sum_i x_i y_i = 7687.76$, $n = 12$.

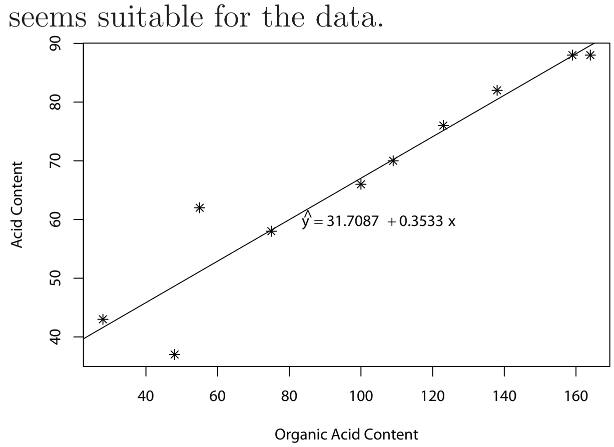
$$b_1 = \frac{(12)(7687.26) - (311.6)(297.2)^2 12}{=} - 0.6861,$$

$$b_0 = \frac{297.2 - (-0.6861)(311.6)}{12} = 42.582.$$

Hence $\hat{y} = 42.582 - 0.6861x$.

(b) At $x = 24.5$, $\hat{y} = 42.582 - (0.6861)(24.5) = 25.772$.

7.7 (a) The scatter plot and the regression line are shown here. A simple linear model seems suitable for the data.



(b) $\sum_i x_i = 999$, $\sum_i y_i = 670$, $\sum_i x_i^2 = 119,969$, $\sum_i x_i y_i = 74,058$, $n = 10$. Therefore,

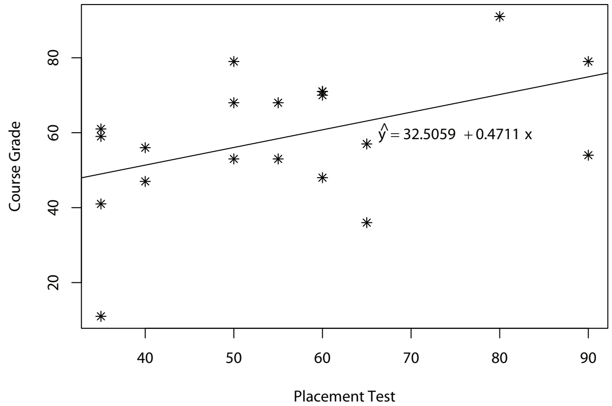
$$b_1 = \frac{(10)(74,058) - (999)(670)}{(10)(119,969) - (999)^2} = 0.3533,$$

$$b_0 = \frac{670 - (0.3533)(999)}{10} = 31.71.$$

Hence $\hat{y} = 31.71 + 0.3533x$.

(c) See (a).

7.8 (a) The scatter plot and the regression line are shown below.



(b) $\sum_i x_i = 1110$, $\sum_i y_i = 1173$, $\sum_i x_i^2 = 67,100$, $\sum_i x_i y_i = 67,690$, $n = 20$. Therefore,

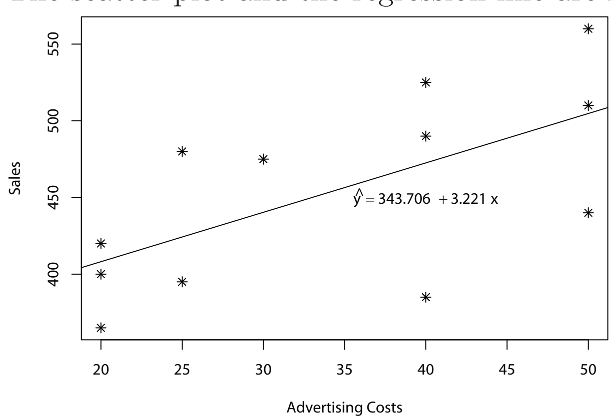
$$b_1 = \frac{(20)(67,690) - (1110)(1173)}{(20)(67,100) - (1110)^2} = 0.4711,$$
$$b_0 = \frac{1173 - (0.4711)(1110)}{20} = 32.5059.$$

Hence $\hat{y} = 32.5059 + 0.4711x$

(c) See part (a).

(d) For $\hat{y} = 60$, we solve $60 = 32.5059 + 0.4711x$ to obtain $x = 58.361$. Therefore, students scoring below 59 should be denied admission.

7.9 (a) The scatter plot and the regression line are shown here.



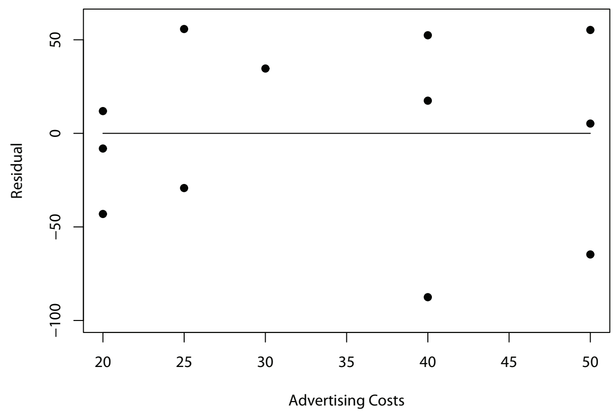
(b) $\sum_i x_i = 410$, $\sum_i y_i = 5445$, $\sum_i x_i^2 = 15,650$, $\sum_i x_i y_i = 191,325$, $n = 12$. Therefore,

$$b_1 = \frac{(12)(191,325) - (410)(5445)}{(12)(15,650) - (410)^2} = 3.2208,$$
$$b_0 = \frac{5445 - (3.2208)(410)}{12} = 343.7056.$$

Hence $\hat{y} = 343.7056 + 3.2208x$

(c) When $x = \$35$, $\hat{y} = 343.7056 + (3.2208)(35) = \$456.43$.

(d) Residuals appear to be random as desired.



7.10 $\hat{z} = cd^w$, $\ln \hat{z} = \ln c + (\ln d)w$; setting $\hat{y} = \ln z$, $a = \ln c$, $b = \ln d$, and $\hat{y} = a + bx$, we have

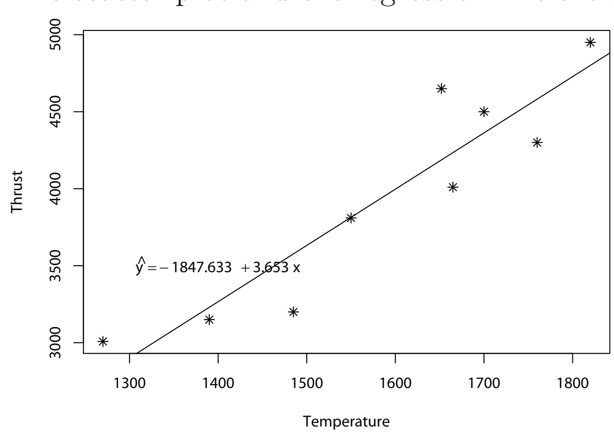
$x = w$	1	2	2	3	5	5
$y = \ln z$	8.7562	8.6473	8.6570	8.5932	8.5142	8.4960

$\sum_i x_i = 18$, $\sum_i y_i = 51.6639$, $\sum_i x_i^2 = 68$, $\sum_i x_i y_i = 154.1954$, $n = 6$.

$$b = \ln d = \frac{(6)(154.1954) - (18)(51.6639)}{(6)(68) - (18)^2} = -0.0569,$$
$$a = \ln c = \frac{51.6639 - (-0.0569)(18)}{6} = 8.7813.$$

Now $c = e^{8.7813} = 6511.3364$, $d = e^{-0.0569} = 0.9447$, and $\hat{z} = 6511.3364 \times 0.9447^w$.

7.11 (a) The scatter plot and the regression line are shown here.

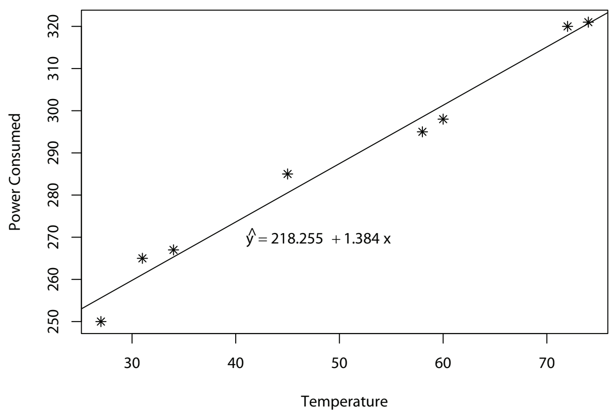


(b) $\sum_i x_i = 14,292$, $\sum_i y_i = 35,578$, $\sum_i x_i^2 = 22,954,054$, $\sum_i x_i y_i = 57,441,610$, $n = 9$. Therefore,

$$b = \frac{(9)(57,441,610) - (14,292)(35,578)}{(9)(22,954,054) - (14,292)^2} = 3.6529,$$
$$a = \frac{35,578 - (3.6529)(14,292)}{9} = -1847.69.$$

Hence $\hat{y} = -1847.69 + 3.6529x$.

7.12 (a) The scatter plot and the regression line are shown here.



(b) $\sum_i x_i = 401$, $\sum_i y_i = 2301$, $\sum_i x_i^2 = 22,495$, $\sum_i x_i y_i = 118,652$, $n = 8$. Therefore,

$$b = \frac{(8)(118,652) - (401)(2301))}{(8)(22,495) - (401)^2} = 1.3839,$$
$$a = \frac{2301 - (1.3839)(401)}{8} = 218.26.$$

Hence $\hat{y} = 218.26 + 1.3839x$.

(c) For $x = 65°\text{F}$, $\hat{y} = 218.26 + (1.3839)(65) = 308.21$.

7.13 (a) $\sum_i x_i = 45$, $\sum_i y_i = 1094$, $\sum_i x_i^2 = 244.26$, $\sum_i x_i y_i = 5348.2$, $n = 9$.

$$b = \frac{(9)(5348.2) - (45)(1094)}{(9)(244.26) - (45)^2} = -6.3240,$$
$$a = \frac{1094 - (-6.3240)(45)}{9} = 153.1755.$$

Hence $\hat{y} = 153.1755 - 6.3240x$.

(b) For $x = 4.8$, $\hat{y} = 153.1755 - (6.3240)(4.8) = 123$.

7.14 From the data summary, we obtain

$$b = \frac{(12)(318) - [(4)(12)][(12)(12)]}{(12)(232) - [(4)(12)]^2} = -6.45,$$
$$a = 12 - (-6.45)(4) = 37.8.$$

Hence, $\hat{y} = 37.8 - 6.45x$. It appears that attending professional meetings would not result in publishing more papers.

7.15 $S_{xx} = 26,591.63 - 778.7^2/25 = 2336.6824$, $S_{yy} = 172,891.46 - 2050^2/25 = 4791.46$, $S_{xy} = 65,164.04 - (778.7)(2050)/25 = 1310.64$, and $b = 0.5609$.

(a) $s^2 = \frac{4791.46 - (0.5609)(1310.64)}{23} = 176.362$.

(b) The hypotheses are

$$H_0 : \beta_1 = 0,$$
$$H_1 : \beta_1 \neq 0.$$

$\alpha = 0.05$.
Critical region: $t < -2.069$ or $t > 2.069$.
Computation: $t = \frac{0.5609}{\sqrt{176.362/2336.6824}} = 2.04$.
Decision: Do not reject H_0.

7.16 $S_{xx} = 57,557 - 707^2/9 = 2018.2222$, $S_{yy} = 51,980 - 658^2/9 = 3872.8889$, $S_{xy} = 53,258 - (707)(658)/9 = 1568.4444$, $b_0 = 12.0623$ and $b_1 = 0.7771$.

(a) $s^2 = \frac{3872.8889 - (0.7771)(1568.4444)}{7} = 379.150$.

(b) Since $s = 19.472$ and $t_{0.025} = 2.365$ for 7 degrees of freedom, then a 95% confidence interval is

$$12.0623 \pm (2.365)\sqrt{\frac{(379.150)(57,557)}{(9)(2018.222)}} = 12.0623 \pm 81.975,$$

which implies $-69.91 < \alpha < 94.04$.

(c) $0.7771 \pm (2.365)\sqrt{\frac{379.150}{2018.2222}}$ implies $-0.25 < \beta_1 < 1.80$.

7.17 $S_{xx} = 25.85 - 16.5^2/11 = 1.1$, $S_{yy} = 923.58 - 100.4^2/11 = 7.2018$, $S_{xy} = 152.59 - (165)(100.4)/11 = 1.99$, $b_0 = 6.4136$ and $b_1 = 1.8091$.

(a) $s^2 = \frac{7.2018 - (1.8091)(1.99)}{9} = 0.40$.

(b) Since $s = 0.632$ and $t_{0.025} = 2.262$ for 9 degrees of freedom, then a 95% confidence interval is

$$6.4136 \pm (2.262)(0.632)\sqrt{\frac{25.85}{(11)(1.1)}} = 6.4136 \pm 2.0895,$$

which implies $4.324 < \alpha < 8.503$.

(c) $1.8091 \pm (2.262)(0.632)/\sqrt{1.1}$ implies $0.446 < \beta_1 < 3.172$.

7.18 $S_{xx} = 8134.26 - 311.6^2/12 = 43.0467$, $S_{yy} = 7407.80 - 297.2^2/12 = 47.1467$, $S_{xy} = 7687.76 - (311.6)(297.2)/12 = -29.5333$, $b_0 = 42.5818$ and $b_1 = -0.6861$.

(a) $s^2 = \frac{47.1467-(-0.6861)(-29.5333)}{10} = 2.688$.

(b) Since $s = 1.640$ and $t_{0.005} = 3.169$ for 10 degrees of freedom, then a 99% confidence interval is

$$42.5818 \pm (3.169)(1.640)\sqrt{\frac{8134.26}{(12)(43.0467)}} = 42.5818 \pm 20.6236,$$

which implies $21.958 < \alpha < 63.205$.

(c) $-0.6861 \pm (3.169)(1.640)/\sqrt{43.0467}$ implies $-1.478 < \beta_1 < 0.106$.

7.19 $S_{xx} = 37,125 - 675^2/18 = 11,812.5$, $S_{yy} = 17,142 - 488^2/18 = 3911.7778$, $S_{xy} = 25,005 - (675)(488)/18 = 6705$, $b_0 = 5.8261$ and $b_1 = 0.5676$.

(a) $s^2 = \frac{3911.7778-(0.5676)(6705)}{16} = 6.626$.

(b) Since $s = 2.574$ and $t_{0.005} = 2.921$ for 16 degrees of freedom, then a 99% confidence interval is

$$5.8261 \pm (2.921)(2.574)\sqrt{\frac{37,125}{(18)(11,812.5)}} = 5.8261 \pm 3.1417,$$

which implies $2.686 < \alpha < 8.968$.

(c) $0.5676 \pm (2.921)(2.574)/\sqrt{11,812.5}$ implies $0.498 < \beta_1 < 0.637$.

7.20 The hypotheses are

$$H_0 : \beta_0 = 10,$$
$$H_1 : \beta_0 > 10.$$

$\alpha = 0.05$.
Critical region: $t > 1.734$.
Computations: $S_{xx} = 67,100 - 1110^2/20 = 5495$, $S_{yy} = 74,725 - 1173^2/20 = 5928.55$, $S_{xy} = 67,690 - (1110)(1173)/20 = 2588.5$, $s^2 = \frac{5928.55-(0.4711)(2588.5)}{18} = 261.617$ and then $s = 16.175$. Now

$$t = \frac{32.51 - 10}{16.175\sqrt{67,100/(20)(5495)}} = 1.78.$$

Decision: Reject H_0 and claim $\beta_0 > 10$.

7.21 The hypotheses are

$$H_0 : \beta_1 = 6,$$
$$H_1 : \beta_1 < 6.$$

$\alpha = 0.025$.
Critical region: $t = -2.228$.
Computations: $S_{xx} = 15,650 - 410^2/12 = 1641.667$, $S_{yy} = 2,512.925 - 5445^2/12 = 42,256.25$, $S_{xy} = 191,325 - (410)(5445)/12 = 5,287.5$, $s^2 = \frac{42,256.25-(3,221)(5,287.5)}{10} = 2,522.521$ and then $s = 50.225$. Now

$$t = \frac{3.221 - 6}{50.225/\sqrt{1641.667}} = -2.24.$$

Decision: Reject H_0 and claim $\beta_1 < 6$.

7.22 Using the value $s = 19.472$ from Exercise 7.16(a) and the fact that $\bar{y}_0 = 74.230$ when $x_0 = 85$, and $\bar{x} = 78.556$, we have

$$74.230 \pm (2.365)(19.472)\sqrt{\frac{1}{9} + \frac{1.444^2}{2018.222}} = 74.230 \pm 15.4216.$$

Simplifying it we get $58.808 < \mu_{Y \mid 80} < 89.652$.

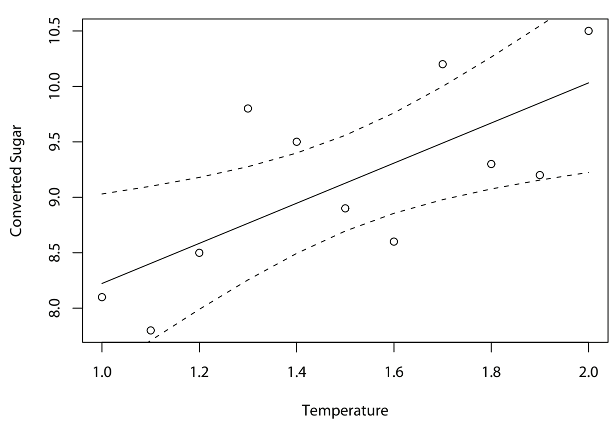
7.23 Using the value $s = 1.64$ from Exercise 11.20(a) and the fact that $y_0 = 25.7724$ when $x_0 = 24.5$, and $\bar{x} = 25.9667$, we have

(a) $25.7724 \pm (2.228)(1.640)\sqrt{\frac{1}{12} + \frac{(-1.4667)^2}{43.0467}} = 25.7724 \pm 1.3341$ implies $24.438 < \mu_{Y \mid 24.5} < 27.106$.

(b) $25.7724 \pm (2.228)(1.640)\sqrt{1 + \frac{1}{12} + \frac{(-1.4667)^2}{43.0467}} = 25.7724 \pm 3.8898$ implies $21.883 < y_0 < 29.662$.

7.24 95% confidence bands are obtained by plotting the limits

$$(6.4136 + 1.809x) \pm (2.262)(0.632)\sqrt{\frac{1}{11} + \frac{(x - 1.5)^2}{1.1}}.$$



7.25 Using the value $s = 0.632$ from Exercise 7.17(a) and the fact that $\hat{y}_0 = 9.308$ when $x_0 = 1.6$, and $\bar{x} = 1.5$, we have

$$9.308 \pm (2.262)(0.632)\sqrt{1 + \frac{1}{11} + \frac{0.1^2}{1.1}} = 9.308 \pm 1.4994$$

implies $7.809 < y_0 < 10.807$.

7.26 Using the value $s = 2.574$ from Exercise 7.19(a) and the fact that $\hat{y}_0 = 34.205$ when $x_0 = 50$, and $\bar{x} = 37.5$, we have

(a) $34.205 \pm (2.921)(2.574)\sqrt{\frac{1}{18} + \frac{12.5^2}{11{,}812.5}} = 34.205 \pm 1.9719$ implies $32.23 < \mu_{Y \mid 50} < 36.18$.

(b) $34.205 \pm (2.921)(2.574)\sqrt{1 + \frac{1}{18} + \frac{12.5^2}{11{,}812.5}} = 34.205 \pm 7.7729$ implies $26.43 < y_0 < 41.98$.

7.27 (a) 17.34.

(b) The goal of 30 mpg is unlikely based on the confidence interval for mean mpg, $(26.41, 29.67)$.

(c) Based on the prediction interval, the mpg of the Lexus ES300 would more likely exceed 18.

7.28 (a) Suppose that the fitted model is $\hat{y} = bx$. Then

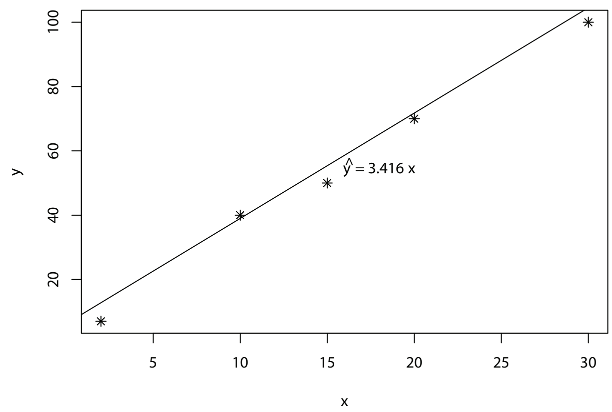
$$SSE = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = \sum_{i=1}^{n}(y_i - bx_i)^2.$$

Taking derivative of the above with respect to b and setting the derivative to zero, we have $-2\sum\limits_{i=1}^{n} x_i(y_i - bx_i) = 0$, which implies $b = \frac{\sum\limits_{i=1}^{n} x_i y_i}{\sum\limits_{i=1}^{n} x_i^2}$.

(b) $\sigma_B^2 = \frac{Var\left(\sum\limits_{i=1}^{n} x_i Y_i\right)}{(\sum\limits_{i=1}^{n} x_i^2)^2} = \frac{\sum\limits_{i=1}^{n} x_i^2 \sigma_{Y_i}^2}{(\sum\limits_{i=1}^{n} x_i^2)^2} = \frac{\sigma^2}{\sum\limits_{i=1}^{n} x_i^2}$, since Y_i's are independent.

(c) $E(B) = \frac{E\left(\sum\limits_{i=1}^{n} x_i Y_i\right)}{\sum\limits_{i=1}^{n} x_i^2} = \frac{\sum\limits_{i=1}^{n} x_i(\beta x_i)}{\sum\limits_{i=1}^{n} x_i^2} = \beta$.

7.29 (a) The scatter plot of the data is shown next.



(b) $\sum_{i=1}^{n} x_i^2 = 1629$ and $\sum_{i=1}^{n} x_i y_i = 5564$. Hence $b = \frac{5564}{1629} = 3.4156$. So, $\hat{y} = 3.4156x$.

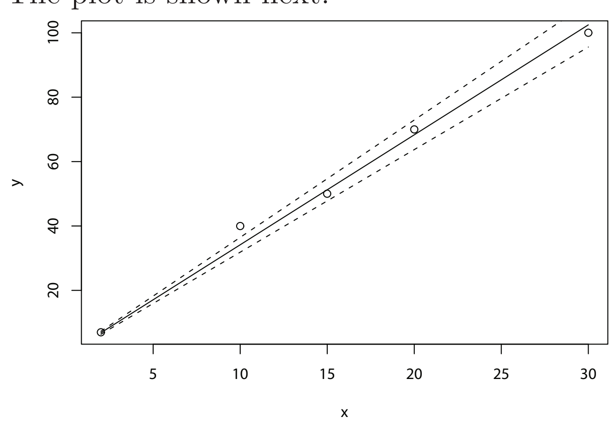
(c) See (a).

(d) Since there is only one regression coefficient, β, to be estimated, the degrees of freedom in estimating σ^2 is $n-1$. So,

$$\hat{\sigma}^2 = s^2 = \frac{SSE}{n-1} = \frac{\sum_{i=1}^{n}(y_i - bx_i)^2}{n-1}.$$

(e) $Var(\hat{y}_i) = Var(Bx_i) = x_i^2 Var(B) = \frac{x_i^2 \sigma^2}{\sum_{i=1}^{n} x_i^2}$.

(f) The plot is shown next.



7.30 Using part (e) of Exercise 7.29, we can see that the variance of a prediction y_0 at x_0 is $\sigma_{y_0}^2 = \sigma^2 \left(1 + \frac{x_0^2}{\sum_{i=1}^{n} x_i^2}\right)$. Hence the 95% prediction limits are given as

$$(3.4145)(25) \pm (2.776)\sqrt{11.16132}\sqrt{1 + \frac{25^2}{1629}} = 85.3625 \pm 10.9092,$$

which implies $74.45 < y_0 < 96.27$.

7.31 The hypotheses are

$$H_0: \text{ The regression is linear in } x,$$
$$H_1: \text{ The regression is nonlinear in } x.$$

$\alpha = 0.05$.
Critical regions: $f > 3.26$ with 4 and 12 degrees of freedom.
Computations: $SST = S_{yy} = 3911.78$, $SSR = bS_{xy} = 3805.89$ and $SSE = S_{yy} - SSR = 105.89$. $SSE(\text{pure}) = \sum_{i=1}^{6}\sum_{j=1}^{3} y_{ij}^2 - \sum_{i=1}^{6} \frac{T_{i.}^2}{3} = 69.33$, and the "Lack-of-fit SS" is $105.89 - 69.33 = 36.56$.

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f
Regression	3805.89	1	3805.89	
Error	105.89	16	6.62	
Lack of fit	36.56	4	9.14	1.58
Pure error	69.33	12	5.78	
Total	3911.78	17		

Decision: Do not reject H_0; the lack-of-fit test is insignificant.

7.32 The hypotheses are

$$H_0: \text{ The regression is linear in } x,$$
$$H_1: \text{ The regression is nonlinear in } x.$$

$\alpha = 0.05$.
Critical regions: $f > 3.00$ with 6 and 12 degrees of freedom.
Computations: $SST = S_{yy} = 5928.55$, $SSR = bS_{xy} = 1219.35$ and $SSE = S_{yy} - SSR = 4709.20$. $SSE(\text{pure}) = \sum_{i=1}^{8}\sum_{j=1}^{n_i} y_{ij}^2 - \sum_{i=1}^{8} \frac{T_{i.}^2}{n_i} = 3020.67$, and the "Lack-of-fit SS" is $4709.20 - 3020.67 = 1688.53$.

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f
Regression	1219.35	1	1219.35	
Error	4709.20	18	261.62	
Lack of fit	1688.53	6	281.42	1.12
Pure error	3020.67	12	251.72	
Total	5928.55	19		

Decision: Do not reject H_0; the lack-of-fit test is insignificant.

7.33 (a) Since $\sum_{i=1}^{n} x_i y_i = 197.59$, and $\sum_{i=1}^{n} x_i^2 = 98.64$, then $b = \frac{197.59}{98.64} = 2.003$ and $\hat{y} = 2.003x$.

(b) It can be calculated that $b_1 = 1.929$ and $b_0 = 0.349$ and hence $\hat{y} = 0.349 + 1.929x$ when intercept is in the model. To test the hypotheses

$$H_0 : \beta_0 = 0,$$
$$H_1 : \beta_1 \neq 0,$$

with 0.10 level of significance, we have the critical regions as $t < -2.132$ or $t > 2.132$.
Computations: $s^2 = 0.0957$ and $t = \frac{0.349}{\sqrt{(0.0957)(98.64)/(6)(25.14)}} = 1.40$.
Decision: Fail to reject H_0; the intercept appears to be zero.

7.34 The hypotheses are

$$H_0 : \beta = 0,$$
$$H_1 : \beta \neq 0.$$

Level of significance: 0.05.
Critical regions: $f > 5.12$.
Computations: $SSR = bS_{xy} = (1.8091)(1.99) = 3.60$ and $SSE = S_{yy} - SSR = 7.20 - 3.60 = 3.60$.

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f
Regression	3.60	1	3.60	9.00
Error	3.60	9	0.40	
Total	7.20	10		

Decision: Reject H_0.

7.35 (a) $S_{xx} = 1058$, $S_{yy} = 198.76$, $S_{xy} = -363.63$, $b = \frac{S_{xy}}{S_{xx}} = -0.34370$, and $a = \frac{210-(-0.34370)(172.5)}{25} = 10.81153$.

(b) The hypotheses are

$$H_0 : \text{ The regression is linear in } x,$$
$$H_1 : \text{ The regression is nonlinear in } x.$$

$\alpha = 0.05$.
Critical regions: $f > 3.10$ with 3 and 20 degrees of freedom.
Computations: $SST = S_{yy} = 198.76$, $SSR = bS_{xy} = 124.98$ and $SSE = S_{yy} - SSR = 73.98$. Since

$$T_{1.} = 51.1, T_{2.} = 51.5, T_{3.} = 49.3, T_{4.} = 37.0 \text{ and } T_{5.} = 22.1,$$

then

$$SSE(\text{pure}) = \sum_{i=1}^{5}\sum_{j=1}^{5} y_{ij}^2 - \sum_{i=1}^{5} \frac{T_{i.}^2}{5} = 1979.60 - 1910.272 = 69.33.$$

Hence the "Lack-of-fit SS" is $73.78 - 69.33 = 4.45$.

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f
Regression	124.98	1	124.98	
Error	73.98	23	3.22	
Lack of fit	4.45	3	1.48	0.43
Pure error	69.33	20	3.47	
Total	198.76	24		

Decision: Do not reject H_0.

7.36 (a) $t = 2.679$ and $0.01 < P(T > 2.679) < 0.015$, hence $0.02 < P$-value < 0.03. There is a strong evidence that the slope is not 0. Hence emitter drive-in time influences gain in a positive linear fashion.

(b) $f = 7.48$ with a P-value of 0.019, which results in a strong evidence that the lack-of-fit test is significant. Hence the linear model is not adequate.

(c) Emitter does does not influence gain in a linear fashion. A better model is a quadratic one using emitter drive-in time to explain the variability in gain.

7.37 $\hat{y} = -21.0280 + 0.4072x$; $f_{\text{LOF}} = 1.71$ with a P-value $= 0.2517$. Hence, lack-of-fit test is insignificant and the linear model is adequate.

7.38 (a) $\hat{y} = 0.011571 + 0.006462x$ with $t = 7.532$ and $P(T > 7.532) < 0.0005$ Hence P-value < 0.001; the slope is significantly different from 0 in the linear regression model.

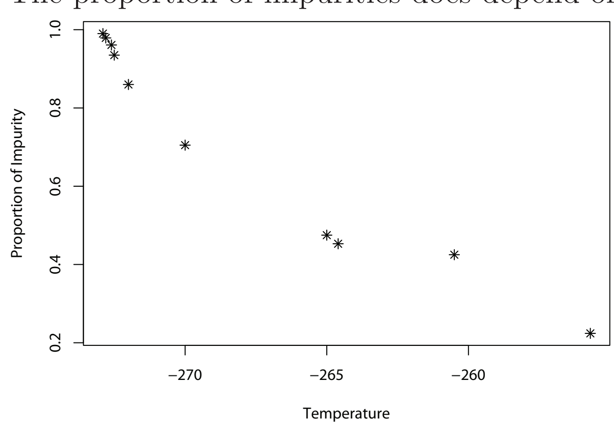
(b) $f_{\text{LOF}} = 14.02$ with P-value < 0.0001. The lack-of-fit test is significant and the linear model does not appear to be the best model.

7.39 (a) $\hat{y} = -11.3251 - 0.0449$ temperature.

(b) Yes.

(c) 0.9355.

(d) The proportion of impurities does depend on temperature.



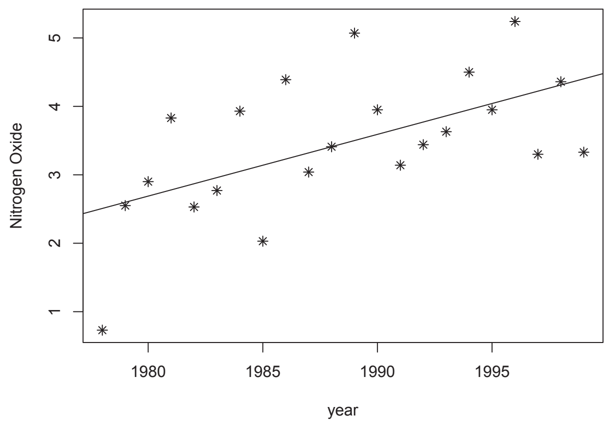
However, based on the plot, it does not appear that the dependence is in linear fashion. If there were replicates, a lack-of-fit test could be performed.

7.40 (a) $\hat{y} = 125.9729 + 1.7337$ population; P-value for the regression is 0.0023.

(b) $f_{6,2} = 0.49$ and P-value $= 0.7912$; the linear model appears to be adequate based on the lack-of-fit test.

(c) $f_{1,2} = 11.96$ and P-value $= 0.0744$. The results do not change. The pure error test is not as sensitive because the loss of error degrees of freedom.

7.41 (a) The figure is shown next.



(b) $\hat{y} = -175.9025 + 0.0902$ year; $R^2 = 0.3322$.

(c) There is definitely a relationship between year and nitrogen oxide. It does not appear to be linear.

7.42 The ANOVA model is:

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f
Regression	135.2000	1	135.2000	
Error	10.4700	14	0.7479	
Lack of fit	6.5150	2	3.2575	9.88
Pure error	3.9550	12	0.3296	
Total	145.6700	15		

The P-value $= 0.0029$ with $f = 9.88$.
Decision: Reject H_0; the lack-of-fit test is significant.

7.43 $S_{xx} = 36,354 - 35,882.667 = 471.333$, $S_{yy} = 38,254 - 37,762.667 = 491.333$, and $S_{xy} = 36,926 - 36,810.667 = 115.333$. So, $r = \frac{115.333}{\sqrt{(471.333)(491.333)}} = 0.240$.

7.44 (a) From the data of Exercise 7.1 we can calculate

$$S_{xx} = 26,591.63 - (778.7)^2/25 = 2336.6824,$$
$$S_{yy} = 172,891.46 - (2050)^2/25 = 4791.46,$$
$$S_{xy} = 65,164.04 - (778.7)(2050)/25 = 1310.64.$$

Therefore, $r = \frac{1310.64}{\sqrt{(2236.6824)(4791.46)}} = 0.40$.

(b) The hypotheses are

$$H_0 : \rho = 0,$$
$$H_1 : \rho \neq 0.$$

$\alpha = 0.05$.
Critical regions: $t < -2.069$ or $t > 2.069$.
Computations: $t = \frac{0.392\sqrt{23}}{\sqrt{1-0.392^2}} = 2.04$.
Decision: Fail to reject H_0 at level 0.05. However, the P-value $= 0.0530$ which is marginal.

7.45 (a) From the data of Exercise 7.13 we find $S_{xx} = 244.26 - 45^2/9 = 19.26$, $S_{yy} = 133,786 - 1094^2/9 = 804.2222$, and $S_{xy} = 5348.2 - (45)(1094)/9 = -121.8$. So, $r = \frac{-121.8}{\sqrt{(19.26)(804.2222)}} = -0.979$.

(b) The hypotheses are

$$H_0 : \rho = -0.5,$$
$$H_1 : \rho < -0.5.$$

$\alpha = 0.025$.
Critical regions: $z < -1.96$.
Computations: $z = \frac{\sqrt{6}}{2} \ln\left[\frac{(0.021)(1.5)}{(1.979)(0.5)}\right] = -4.22$.
Decision: Reject H_0: $\rho < -0.5$.

(c) $(-0.979)^2(100\%) = 95.8\%$.

7.46 The hypotheses are

$$H_0 : \rho = 0,$$
$$H_1 : \rho \neq 0.$$

$\alpha = 0.05$.
Critical regions: $t < -2.776$ or $t > 2.776$.
Computations: $t = \frac{0.240\sqrt{4}}{\sqrt{1-0.240^2}} = 0.49$.
Decision: Do not reject H_0.

7.47 (a) $S_{xx} = 128.6602 - 32.68^2/9 = 9.9955$, $S_{yy} = 7980.83 - 266.7^2/9 = 77.62$, and $S_{xy} = 990.268 - (32.68)(266.7)/9 = 21.8507$. So, $r = \frac{21.8507}{\sqrt{(9.9955)(77.62)}} = 0.784$.

(b) The hypotheses are

$$H_0 : \rho = 0,$$
$$H_1 : \rho > 0.$$

$\alpha = 0.01$.
Critical regions: $t > 2.998$.
Computations: $t = \frac{0.784\sqrt{7}}{\sqrt{1-0.784^2}} = 3.34$.
Decision: Reject H_0; $\rho > 0$.

(c) $(0.784)^2(100\%) = 61.5\%$.

7.48 $\hat{y} = -3.3727 + 0.0036x_1 + 0.9476x_2$.

7.49 $\hat{y} = 0.5800 + 2.7122x_1 + 2.0497x_2$.

7.50 (a) $\hat{y} = -22.99316 + 1.39567x_1 + 0.21761x_2$.

(b) $\hat{y} = -22.99316 + (1.39567)(35) + (0.21761)(250) = 80.25874$.

7.51 (a) $\widehat{y} = 27.5467 + 0.9217x_1 + 0.2842x_2$.

(b) When $x_1 = 60$ and $x_2 = 4$, the predicted value of the chemistry grade is $\hat{y} = 27.5467 + (0.9217)(60) + (0.2842)(4) = 84$.

7.52 (a) $\hat{d} = 13.35875 - 0.33944v - 0.01183v^2$.

(b) $\hat{d} = 13.35875 - (-0.33944)(70) - (0.01183)(70)^2 = 47.54206$.

7.53 (a) $\hat{y} = -102.71324 + 0.60537x_1 + 8.92364x_2 + 1.43746x_3 + 0.01361x_4$.

(b) $\hat{y} = -102.71324 + (0.60537)(75) + (8.92364)(24) + (1.43746)(90) + (0.01361)(98) = 287.56183$.

7.54 (a) $\hat{y} = 19.03333 + 1.0086x - 0.02038x^2$.

(b) $SSE = 24.47619$ with 12 degrees of freedom and $SS(\text{pure error}) = 24.36667$ with 10 degrees of freedom. So, $SSLOF = 24.47619 - 24.36667 = 0.10952$ with 2 degrees of freedom. Hence $f = \frac{0.10952/2}{24.36667/10} = 0.02$ with a P-value of 0.9778. Therefore, there is no lack of fit and the quadratic model fits the data well.

7.55 $\hat{y} = 141.61178 - 0.28193x + 0.00031x^2$.

7.56 (a) $\hat{y} = 1.07143 + 4.60317x - 1.84524x^2 + 0.19444x^3$.

(b) $\hat{y} = 1.07143 + (4.60317)(2) - (1.84524)(2)^2 + (0.19444)(2)^3 = 4.45238$.

7.57 (a) $\hat{y} = 56.46333 + 0.15253x - 0.00008x^2$.

(b) $\hat{y} = 56.46333 + (0.15253)(225) - (0.00008)(225)^2 = 86.73333\%$.

7.58 $\hat{y} = 1,962.94816 - 15.85168x_1 + 0.05593x_2 + 1.58962x_3 - 4.21867x_4 - 394.31412x_5$.

7.59 $\hat{y} = -6.51221 + 1.99941x_1 - 3.67510x_2 + 2.52449x_3 + 5.15808x_4 + 14.40116x_5$.

7.60 (a) $\hat{y} = -21.46964 - 3.32434x_1 + 0.24649x_2 + 20.34481x_3$.

(b) $\hat{y} = -21.46964 - (3.32434)(14) + (0.24649)(220) + (20.34481)(5) = 87.94123$.

7.61 (a) $\hat{y} = 350.99427 - 1.27199x_1 - 0.15390x_2$.

(b) $\hat{y} = 350.99427 - (1.27199)(20) - (0.15390)(1200) = 140.86930$.

7.62 $\hat{y} = -884.667 - 0.83813x_1 + 4.90661x_2 + 1.33113x_3 + 11.93129x_4$.

7.63 $\hat{y} = 3.3205 + 0.42105x_1 - 0.29578x_2 + 0.01638x_3 + 0.12465x_4$.

7.64 $s^2 = 0.43161$.

7.65 $s^2 = 0.16508$.

7.66 Using *SAS* output, we obtain

$$\hat{\sigma}^2_{b_1} = 3.747 \times 10^{-7}, \quad \hat{\sigma}^2_{b_2} = 0.13024, \quad \hat{\sigma}_{b_1 b_2} = -4.165 \times 10^{-7}.$$

7.67 $s^2 = 242.71561$.

7.68 The hypotheses are

$$H_0 : \beta_2 = 0,$$
$$H_1 : \beta_2 \neq 0.$$

The test statistic value is $t = 2.86$ with a P-value $= 0.0145$. Hence, we reject H_0 and conclude $\beta_2 \neq 0$.

7.69 Using *SAS* output, we obtain

(a) $\hat{\sigma}^2_{b_2} = 28.09554$.

(b) $\hat{\sigma}_{b_1 b_4} = -0.00958$.

7.70 The hypotheses are

$$H_0 : \beta_1 = 2,$$
$$H_1 : \beta_1 \neq 2.$$

The test statistics is $t = \frac{2.71224-2}{0.20209} = 3.524$ with P-value $= 0.0097$. Reject H_0 and conclude that $\beta_1 \neq 2$.

7.71 The test statistic is $t = \frac{0.00362}{0.000612} = 5.91$ with P-value $= 0.0002$. Reject H_0 and claim that $\beta_1 \neq 0$.

7.72 Using *SAS* output, we obtain a 90% confidence interval for the mean response when $x = 19.5$ as $29.9284 < \mu_{Y|x=19.5} < 31.9729$.

7.73 Using *SAS* output, we obtain
$0.4516 < \mu_{Y|x_1=900,x_2=1} < 1.2083$, and $-0.1640 < y_0 < 1.8239$.

7.74 Using *SAS* output, we obtain

(a) $s^2 = 650.1408$.

(b) $\hat{y} = 171.6501$, $135.8735 < \mu_{Y|x_1=20,x_2=1000} < 207.4268$, and
$82.9677 < y_0 < 260.3326$.

7.75 Using *SAS* output, we obtain
$263.7879 < \mu_{Y|x_1=75,x_2=24,x_3=90,x_4=98} < 311.3357$, and $243.7175 < y_0 < 331.4062$.

7.76 (a) $s^2 = 17.22858$.

(b) $\hat{y} = 104.9617$ and $95.5660 < y_0 < 114.3574$.

7.77 (a) P-value $= 0.3562$. Hence, fail to reject H_0.

(b) P-value $= 0.1841$. Again, fail to reject H_0.

(c) There is not sufficient evidence that the regressors x_1 and x_2 significantly influence the response with the described linear model.

7.78 Using the value $s = 16.175$ from Exercise 7.8 and the fact that $\hat{y}_0 = 48.994$ when $x_0 = 35$, and $\bar{x} = 55.5$, we have

(a) $48.994 \pm (2.101)(16.175)\sqrt{1/20 + (-20.5)^2/5495}$ which implies to $36.908 < \mu_{Y\ |\ 35} < 61.080$.

(b) $48.994 \pm (2.101)(16.175)\sqrt{1 + 1/20 + (-20.5)^2/5495}$ which implies to $12.925 < y_0 < 85.063$.

7.79 The fitted model can be derived as $\hat{y} = 3667.3968 - 47.3289x$.
The hypotheses are

$$H_0 : \beta = 0,$$
$$H_1 : \beta \neq 0.$$

$t = -0.30$ with P-value $= 0.77$. Hence H_0 cannot be rejected.

7.80 (a) $S_{xx} = 729.18 - 118.6^2/20 = 25.882$, $S_{xy} = 1714.62 - (118.6)(281.1)/20 = 47.697$, so $b = \frac{S_{xy}}{S_{xx}} = 1.8429$, and $a = \bar{y} - b\bar{x} = 3.1266$. Hence $\hat{y} = 3.1266 + 1.8429x$.

(b) The hypotheses are

$$H_0 : \text{ the regression is linear in } x,$$
$$H_1 : \text{ the regression is not linear in } x.$$

$\alpha = 0.05$.
Critical region: $f > 3.07$ with 8 and 10 degrees of freedom.
Computations: $SST = 13.3695$, $SSR = 87.9008$, $SSE = 50.4687$, $SSE(\text{pure}) = 16.375$, and Lack-of-fit SS $= 34.0937$.

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f
Regression	87.9008	1	87.9008	
Error	50.4687	18	2.8038	
Lack of fit	34.0937	8	4.2617	2.60
Pure error	16.375	10	1.6375	
Total	138.3695	19		

The P-value $= 0.0791$. The linear model is adequate at the level 0.05.

7.81 Using the value $s = 50.225$ and the fact that $\hat{y}_0 = \$488.642$ when $x_0 = \$45$, and $\bar{x} = \$34.167$, we have

(a) $488.642 \pm (2.228)(50.225)\sqrt{\frac{1}{12} + \frac{10.833^2}{1641.667}}$, which implies $444.61 < \mu_{Y \mid 45} < 532.68$.

(b) $488.642 \pm (2.228)(50.225)\sqrt{1 + \frac{1}{12} + \frac{10.833^2}{1641.667}}$, which implies $368.38 < y_0 < 608.90$.

7.82 (a) $\hat{y} = 7.3598 + 135.4034x$.

(b) $SS(\text{Pure Error}) = 52,941.06$; $f_{\text{LOF}} = 0.46$ with P-value $= 0.64$. The lack-of-fit test is insignificant.

(c) No.

7.83 (a) $S_{xx} = 672.9167$, $S_{yy} = 728.25$, $S_{xy} = 603.75$ and $r = \frac{603.75}{\sqrt{(672.9167)(728.25)}} = 0.862$, which means that $(0.862)^2(100\%) = 74.3\%$ of the total variation of the values of Y in our sample is accounted for by a linear relationship with the values of X.

(b) To estimate and test hypotheses on ρ, X and Y are assumed to be random variables from a bivariate normal distribution.

(c) The hypotheses are

$$H_0 : \rho = 0.5,$$
$$H_1 : \rho > 0.5.$$

$\alpha = 0.01$.
Critical regions: $z > 2.33$.
Computations: $z = \frac{\sqrt{9}}{2} \ln\left[\frac{(1.862)(0.5)}{(0.138)(1.5)}\right] = 2.26$.
Decision: Fail to reject H_0; $\rho > 0.5$.

7.84 (a) The confidence interval is an interval on the mean sale price for a given buyer's bid. The prediction interval is an interval on a future observed sale price for a given buyer's bid.

(b) The standard errors of the prediction of sale price depend on the value of the buyer's bid.

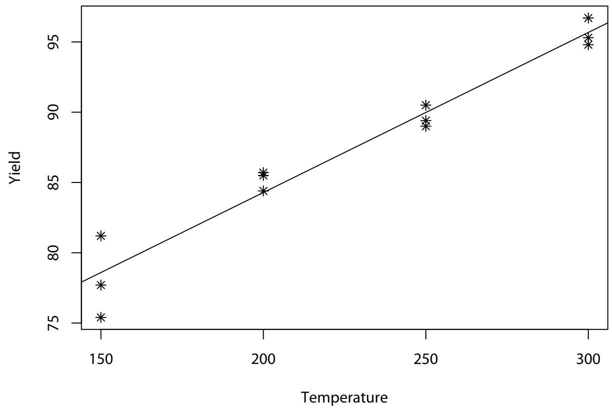
(c) Observations 4, 9, 10, and 17 have the lowest standard errors of prediction. These observations have buyer's bids very close to the mean.

7.85 (a) The residual plot appears to have a pattern and not random scatter. The R^2 is only 0.82.

(b) The log model has an R^2 of 0.84. There is still a pattern in the residuals.

(c) The model using gallons per 100 miles has the best R^2 with a 0.85. The residuals appear to be more random. This model is the best of the three models attempted. Perhaps a better model could be found.

7.86 (a) The plot of the data and an added least squares fitted line are given here.



(b) Yes.

(c) $\hat{y} = 61.5133 + 0.1139x$.

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f
Regression	486.21	1	486.21	
Error	24.80	10	2.48	
Lack of fit	3.61	2	1.81	0.68
Pure error	21.19	8	2.65	
Total	511.01	11		

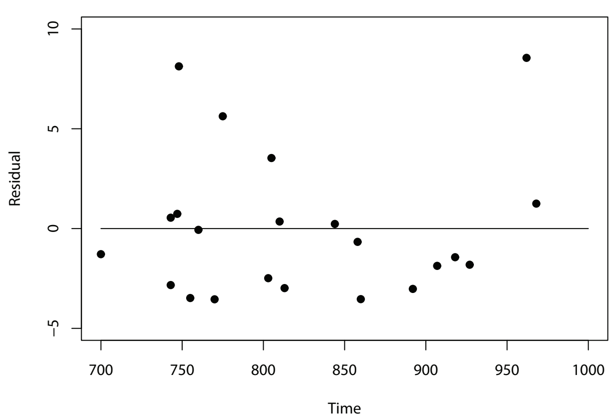
The P-value $= 0.533$.

(d) The results in (c) show that the linear model is adequate.

7.87 (a) $\hat{y} = 90.8904 - 0.0513x$.

(b) The t-value in testing $H_0: \ \beta = 0$ is -6.533 which results in a P-value < 0.0001. Hence, the time it takes to run two miles has a significant influence on maximum oxygen uptake.

(c) The residual graph shows that there may be some systematic behavior of the residuals and hence the residuals are not completely random



7.89 The fitted model is $\hat{y} = 45.4878+4.0917x_1+6.8933x_2-2.2617x_1^2-0.7267x_2^2+2.7800x_1x_2$. The t-tests for each coefficient show that x_1^2 and x_2^2 may be eliminated. So, we ran a test for $\beta_{11} = \beta_{22} = 0$ which yields P-value $= 0.2817$. Therefore, both x_1^2 and x_2^2 may be dropped out from the model.

7.90 Denote by $Z_1 = 1$ when Group=1, and $Z_1 = 0$ otherwise;
Denote by $Z_2 = 1$ when Group=2, and $Z_2 = 0$ otherwise;
Denote by $Z_3 = 1$ when Group=3, and $Z_3 = 0$ otherwise;

(a) The parameter estimates are:

Variable	DF	Parameter Estimate	P-value
Intercept	1	46.34694	0.0525
BMI	1	-1.79090	0.0515
z_1	1	-23.84705	0.0018
z_2	1	-17.46248	0.0109

Yes, Group I has a mean change in blood pressure that was significantly lower than the control group. It is about 23.85 points lower.

(b) The parameter estimates are:

Variable	DF	Parameter Estimate	P-value
Intercept	1	28.88446	0.1732
BMI	1	-1.79090	0.0515
z_1	1	-6.38457	0.2660
z_3	1	17.46248	0.0109

Although Group I has a mean change in blood pressure that was 6.38 points lower than that of Group II, the difference is not very significant due to a high P-value.

7.91 Using the formula of R^2_{adj} on page 467, we have

$$R^2_{\text{adj}} = 1 - \frac{SSE/(n-k-1)}{SST/(n-1)} = 1 - \frac{MSE}{MST}.$$

Since MST is fixed, maximizing R^2_{adj} is thus equivalent to minimizing MSE.

7.92 (b) All possible regressions should be run. $R^2 = 0.9908$ and there is only one significant variable.

(c) The model including x_2, x_3 and x_5 is the best in terms of C_p, PRESS and has all variables significant.

Chapter 8

One-Factor Experiments: General

8.1 The hypotheses are

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_6,$$
$$H_1 : \text{At least two of the means are not equal.}$$

$\alpha = 0.05$.
Critical region: $f > 2.77$ with $v_1 = 5$ and $v_2 = 18$ degrees of freedom.
Computation:

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f
Treatment	5.34	5	1.07	0.31
Error	62.64	18	3.48	
Total	67.98	23		

with P-value=0.9024.
Decision: The treatment means do not differ significantly.

8.2 The hypotheses are

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_5,$$
$$H_1 : \text{At least two of the means are not equal.}$$

$\alpha = 0.05$.
Critical region: $f > 2.87$ with $v_1 = 4$ and $v_2 = 20$ degrees of freedom.
Computation:

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f
Tablets	78.422	4	19.605	6.59
Error	59.532	20	2.977	
Total	137.954	24		

with P-value=0.0015.
Decision: Reject H_0. The mean number of hours of relief differ significantly.

8.3 The hypotheses are

$$H_0: \mu_1 = \mu_2 = \mu_3,$$
$$H_1: \text{At least two of the means are not equal.}$$

$\alpha = 0.01$.
Computation:

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f
Shelf Height	399.3	2	199.63	14.52
Error	288.8	21	13.75	
Total	688.0	23		

with P-value=0.0001.
Decision: Reject H_0. The amount of money spent on dog food differs with the shelf height of the display.

8.4 The hypotheses are

$$H_0: \mu_A = \mu_B = \mu_C,$$
$$H_1: \text{At least two of the means are not equal.}$$

$\alpha = 0.01$.
Computation:

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f
Drugs	158.867	2	79.433	5.46
Error	393.000	27	14.556	
Total	551.867	29		

with P-value=0.0102.
Decision: Since $\alpha = 0.01$, we fail to reject H_0. However, this decision is very marginal since the P-value is very close to the significance level.

8.5 The hypotheses are

$$H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4,$$
$$H_1: \text{At least two of the means are not equal.}$$

$\alpha = 0.01$.
Computation:

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f
Treatments	27.5506	3	9.1835	8.38
Error	18.6360	17	1.0962	
Total	46.1865	20		

with P-value= 0.0012.
Decision: Reject H_0. Average specific activities differ.

8.6 The hypotheses are

$$H_0 : \mu_1 = \mu_2 = \mu_3,$$
$$H_1 : \text{At least two of the means are not equal.}$$

Computation:

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f
Solvents	3.3054	2	1.6527	24.51
Error	1.9553	29	0.0674	
Total	5.2608	31		

with P-value< 0.0001.
Decision: There is significant difference in the mean sorption rate for the three solvents. The mean sorption for the solvent Chloroalkanes is the highest. We know that it is significantly higher than the rate of Esters.

8.7 The hypotheses are

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4,$$
$$H_1 : \text{At least two of the means are not equal.}$$

$\alpha = 0.05$.
Computation:

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f
Treatments	119.787	3	39.929	2.25
Error	638.248	36	17.729	
Total	758.035	39		

with P-value=0.0989.
Decision: Fail to reject H_0 at level $\alpha = 0.05$.

8.8 $s_{50} = 3.2098, s_{100} = 4.5253, s_{200} = 5.1788$, and $s_{400} = 3.6490$. Since the sample sizes are all the same,

$$s_p^2 = \frac{1}{4}\sum_{i=1}^{4} s_i^2 = 17.7291.$$

Therefore, the Bartlett's statistic is

$$b = \frac{\left(\prod_{i=1}^{4} s_i^2\right)^{1/4}}{s_p^2} = 0.9335.$$

Using Table A.8, the critical value of the Bartlett's test with $k = 4$ and $\alpha = 0.05$ is 0.7970. Since $b > 0.7970$, we fail to reject H_0 and hence the variances can be assumed equal.

8.9 The hypotheses for the Bartlett's test are

$$H_0 : \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma_4^2,$$
$$H_1 : \text{The variances are not all equal.}$$

$\alpha = 0.01$.
Critical region: We have $n_1 = n_2 = n_3 = 4$, $n_4 = 9$ $N = 21$, and $k = 4$. Therefore, we reject H_0 when

$$\begin{aligned} b &< b_4(0.01, 4, 4, 4, 9) \\ &= \frac{(4)(0.3475) + (4)(0.3475) + (4)(0.3475) + (9)(0.6892)}{21} = 0.4939. \end{aligned}$$

Computation: $s_1^2 = 0.41709$, $s_2^2 = 0.93857$, $s_3^2 = 0.25673$, $s_4^2 = 1.72451$ and hence $s_p^2 = 1.0962$. Therefore,

$$b = \frac{[(0.41709)^3(0.93857)^3(0.25763)^3(1.72451)^8]^{1/17}}{1.0962} = 0.79.$$

Decision: Do not reject H_0; the variances are not significantly different.

8.10 The hypotheses for the Cochran's test are

$$H_0 : \sigma_A^2 = \sigma_B^2 = \sigma_C^2,$$
$$H_1 : \text{The variances are not all equal.}$$

$\alpha = 0.01$.
Critical region: $g > 0.6912$.
Computation: $s_A^2 = 29.5667$, $s_B^2 = 10.8889$, $s_C^2 = 3.2111$, and hence $\sum s_i^2 = 43.6667$.
Now, $g = \frac{29.5667}{43.6667} = 0.6771$.
Decision: Do not reject H_0; the variances are not significantly different.

8.11 The hypotheses for the Bartlett's test are

$$H_0 : \sigma_1^2 = \sigma_2^2 = \sigma_3^2,$$
$$H_1 : \text{The variances are not all equal.}$$

$\alpha = 0.05$.
Critical region: reject H_0 when

$$b < b_4(0.05, 9, 8, 15) = \frac{(9)(0.7686) + (8)(0.7387) + (15)(0.8632)}{32} = 0.8055.$$

Computation: $b = \frac{[(0.02832)^8(0.16077)^7(0.04310)^{14}]^{1/29}}{0.067426} = 0.7822$.
Decision: Reject H_0; the variances are significantly different.

8.12 (a) The hypotheses are

$$H_0 : \mu_{29} = \mu_{54} = \mu_{84},$$
$$H_1 : \text{At least two of the means are not equal.}$$

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f
Protein Levels	32, 974.87	2	16, 487.43	5.15
Error	28, 815.80	9	3, 201.76	
Total	61, 790.67	11		

with P-value$= 0.0323$.
Decision: Reject H_0. The mean nitrogen loss was significantly different for the three protein levels.

(b) For testing the contrast $L = 2\mu_{29} - \mu_{54} - \mu_{84}$ at level $\alpha = 0.05$, we have $SSw = 31,576.42$ and $f = 9.86$, with P-value=0.0119. Hence, the mean nitrogen loss for 29 grams of protein was different from the average of the two higher protein levels.

8.13 (a) The hypotheses are

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4,$$
$$H_1 : \text{At least two of the means are not equal.}$$

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f
Treatments	1083.60	3	361.20	13.50
Error	1177.68	44	26.77	
Total	2261.28	47		

with P-value< 0.0001.
Decision: Reject H_0. The treatment means are different.

(b) The treatment means are

$$\bar{y}_{1.} = 16.68, \bar{y}_{2.} = 5.24, \bar{y}_{3.} = 17.07, \bar{y}_{4.} = 13.07.$$

Since $q(0.05; 4, 14) \approx 3.79$ from Table A.10, $s^2 = 26.77$ and $n = 12$, the critical difference is 5.66.
Therefore, the results of Tukey's test are

$\bar{y}_{2.}$	$\bar{y}_{4.}$	$\bar{y}_{1.}$	$\bar{y}_{3}$
5.24	13.07	16.68	17.07

8.14 (a) The hypotheses are

$$H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4,$$
$$H_1: \text{At least two of the means are not equal.}$$

$\alpha = 0.05$.
Computation:

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f
Blends	119.649	3	39.883	7.10
Error	44.920	8	5.615	
Total	164.569	11		

with P-value$= 0.0121$.
Decision: Reject H_0. There is a significant difference in mean yield reduction for the 4 preselected blends.

(b) Since $\sqrt{s^2/3} = 1.368$ we get

p	2	3	4
r_p	3.261	3.399	3.475
R_p	4.46	4.65	4.75

Therefore,

$\bar{y}_{3.}$	$\bar{y}_{1.}$	$\bar{y}_{2.}$	$\bar{y}_{4.}$
23.23	25.93	26.17	31.90

(c) Since $q(0.05, 4, 8) = 4.53$, the critical difference is 6.20. Hence

$\bar{y}_{3.}$	$\bar{y}_{1.}$	$\bar{y}_{2.}$	$\bar{y}_{4.}$
23.23	25.93	26.17	31.90

8.15 The means of the treatments are:

$$\bar{y}_{1.} = 5.44, \bar{y}_{2.} = 7.90, \bar{y}_{3.} = 4.30, \bar{y}_{4.} = 2.98, \text{ and } \bar{y}_{5.} = 6.96.$$

Since $q(0.05, 5, 20) = 4.24$, the critical difference is $(4.24)\sqrt{\frac{2.9766}{5}} = 3.27$. Therefore, the Tukey's result may be summarized as follows:

$\bar{y}_{4.}$	$\bar{y}_{3.}$	$\bar{y}_{1.}$	$\bar{y}_{5.}$	$\bar{y}_{2.}$
2.98	4.30	5.44	6.96	7.90

8.16 The hypotheses are

$$H_0: \mu_1 = \mu_2 = \cdots = \mu_5,$$
$$H_1: \text{At least two of the means are not equal.}$$

Computation:

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f
Angles	99.024	4	24.756	21.40
Error	23.136	20	1.157	
Total	122.160	24		

with P-value< 0.0001.
Decision: Reject H_0. There is a significant difference in mean pressure for the different angles.

8.17 (a) The hypotheses are

$$H_0: \mu_1 = \mu_2 = \cdots = \mu_5,$$
$$H_1: \text{At least two of the means are not equal.}$$

Computation:

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f
Procedures	7828.30	4	1957.08	9.01
Error	3256.50	15	217.10	
Total	11084.80	19		

with P-value$= 0.0006$.
Decision: Reject H_0. There is a significant difference in the average species count for the different procedures.

(b) Since $q(0.05, 5, 15) = 4.373$ and $\sqrt{\frac{217.10}{4}} = 7.367$, the critical difference is 32.2. Hence

$\bar{y}_K$	$\bar{y}_S$	$\bar{y}_{Sub}$	$\bar{y}_M$	$\bar{y}_D$
12.50	24.25	26.50	55.50	64.25

8.18 The ANOVA table can be obtained as follows:

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f
Cables	1924.296	8	240.537	9.07
Error	2626.917	99	26.535	
Total	4551.213	107		

with P-value< 0.0001.
The results from Tukey's procedure can be obtained as follows:

$\bar{y}_{2.}$	$\bar{y}_{3.}$	$\bar{y}_{1.}$	$\bar{y}_{4.}$	$\bar{y}_{6.}$	$\bar{y}_{7.}$	$\bar{y}_{5.}$	$\bar{y}_{8.}$	$\bar{y}_{9.}$
-7.000	-6.083	-4.083	-2.667	0.833	0.917	1.917	3.333	6.250

The cables are significantly different:
9 is different from 4, 1, 2, 3
8 is different from 1, 3, 2
5, 7, 6 are different from 3, 2.

8.19 The ANOVA table can be obtained as follows:

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f
Temperatures	1268.5333	4	317.1333	70.27
Error	112.8333	25	4.5133	
Total	1381.3667	29		

with P-value< 0.0001.
The results from Tukey's procedure can be obtained as follows:

$\bar{y}_{0}$	$\bar{y}_{25}$	$\bar{y}_{100}$	$\bar{y}_{75}$	$\bar{y}_{50}$
55.167	60.167	64.167	70.500	72.833

The batteries activated at temperature 50 and 75 have significantly longer activated life.

8.20 The Duncan's procedure shows the following results:

$\bar{y}_{E}$	$\bar{y}_{A}$	$\bar{y}_{C}$
0.3300	0.9422	1.0063

Hence, the sorption rate using the Esters is significantly lower than the sorption rate using the Aromatics or the Chloroalkanes.

8.21 Aggregate 4 has a significantly lower absorption rate than the other aggregates.

8.22 The hypotheses are

$$H_0 : \alpha_1 = \alpha_2 = \alpha_3 = 0, \text{ no differences in the varieties}$$
$$H_1 : \text{At least one of the } \alpha_i\text{'s is not equal to zero.}$$

$\alpha = 0.05$.
Critical region: $f > 5.14$.
Computation:

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f
Treatments	24.500	2	12.250	1.74
Blocks	171.333	3	57.111	
Error	42.167	6	7.028	
Total	238.000	11		

P-value=0.2535. Decision: Do not reject H_0; could not show that the varieties of potatoes differ in field.

8.23 The hypotheses are

$$H_0: \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0, \text{ fertilizer effects are zero}$$
$$H_1: \text{At least one of the } \alpha_i\text{'s is not equal to zero.}$$

$\alpha = 0.05$.
Critical region: $f > 4.76$.
Computation:

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f
Fertilizers	218.1933	3	72.7311	6.11
Blocks	197.6317	2	98.8158	
Error	71.4017	6	11.9003	
Total	487.2267	11		

P-value= 0.0296. Decision: Reject H_0. The means are not all equal.

8.24 The hypotheses are

$$H_0: \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0, \text{ courses are equal difficulty}$$
$$H_1: \text{At least one of the } \alpha_i\text{'s is not equal to zero.}$$

$\alpha = 0.05$.
Computation:

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f
Subjects	42.150	3	14.050	0.15
Students	1618.700	4	404.675	
Error	1112.100	12	92.675	
Total	2772.950	19		

P-value=0.9267. Decision: Fail to reject H_0; there is no significant evidence to conclude that courses are of different difficulty.

8.25 The hypotheses are

$$H_0: \alpha_1 = \alpha_2 = \alpha_3 = 0, \text{ brand effects are zero}$$
$$H_1: \text{At least one of the } \alpha_i\text{'s is not equal to zero.}$$

$\alpha = 0.05$.
Critical region: $f > 3.84$.
Computation:

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f
Treatments	27.797	2	13.899	5.99
Blocks	16.536	4	4.134	
Error	18.556	8	2.320	
Total	62.889	14		

P-value=0.0257. Decision: Reject H_0; mean percent of foreign additives is not the same for all three brand of jam. The means are:

Jam A: 2.36, Jam B: 3.48, Jam C: 5.64.

Based on the means, Jam A appears to have the smallest amount of foreign additives.

8.26 The hypotheses are

$$H_0 : \alpha_1 = \alpha_2 = \alpha_3 = 0, \text{ station effects are zero}$$
$$H_1 : \text{At least one of the } \alpha_i\text{'s is not equal to zero.}$$

$\alpha = 0.05$.
Computation:

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f
Stations	10.115	2	5.057	0.15
Months	537.030	11	48.821	
Error	744.416	22	33.837	
Total	1291.561	35		

P-value= 0.8620. Decision: Do not reject H_0; the treatment means do not differ significantly.

8.27 The hypotheses are

$$H_0 : \alpha_1 = \alpha_2 = \cdots = \alpha_6 = 0, \text{ station effects are zero}$$
$$H_1 : \text{At least one of the } \alpha_i\text{'s is not equal to zero.}$$

$\alpha = 0.01$.
Computation:

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f
Stations	230.127	5	46.025	26.14
Dates	3.259	5	0.652	
Error	44.018	25	1.761	
Total	277.405	35		

P-value< 0.0001. Decision: Reject H_0; the mean concentration is different at the different stations.

8.28 The hypotheses are

$$H_0 : \alpha_1 = \alpha_2 = \alpha_3 = 0, \text{ analyst effects are zero}$$
$$H_1 : \text{At least one of the } \alpha_i\text{'s is not equal to zero.}$$

$\alpha = 0.05$.
Computation:

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f
Analysts	0.001400	2	0.000700	3.00
Individuals	0.021225	3	0.007075	
Error	0.001400	6	0.000233	
Total	0.024025	11		

P-value$= 0.1250$. Decision: Do not reject H_0; cannot show that the analysts differ significantly.

8.29 The hypotheses are

$$H_0 : \alpha_1 = \alpha_2 = \alpha_3 = 0, \text{ diet effects are zero}$$
$$H_1 : \text{At least one of the } \alpha_i\text{'s is not equal to zero.}$$

$\alpha = 0.01$.
Computation:

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f
Diets	4297.000	2	2148.500	11.86
Subjects	6033.333	5	1206.667	
Error	1811.667	10	181.167	
Total	12142.000	17		

P-value$= 0.0023$. Decision: Reject H_0; differences among the diets are significant.

8.30 The hypotheses are

$$H_0 : \alpha_1 = \alpha_2 = \alpha_3 = 0, \text{ treatment effects are zero}$$
$$H_1 : \text{At least one of the } \alpha_i\text{'s is not equal to zero.}$$

$\alpha = 0.01$.
Computation:

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f
Treatments	203.2792	2	101.6396	8.60
Subjects	188.2271	9	20.9141	
Error	212.8042	18	11.8225	
Total	604.3104	29		

P-value= 0.0024. Decision: Reject H_0; the mean weight losses are different for different treatments and the therapists had the greatest effect on the weight loss.

8.31 The hypotheses are

$$H_0 : \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = \alpha_5 = 0, \text{ treatment effects are zero}$$
$$H_1 : \text{At least one of the } \alpha_i\text{'s is not equal to zero.}$$

$\alpha = 0.01$.
Computation:

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f
Treatments	79630.133	4	19907.533	0.58
Locations	634334.667	5	126866.933	
Error	689106.667	20	34455.333	
Total	1403071.467	29		

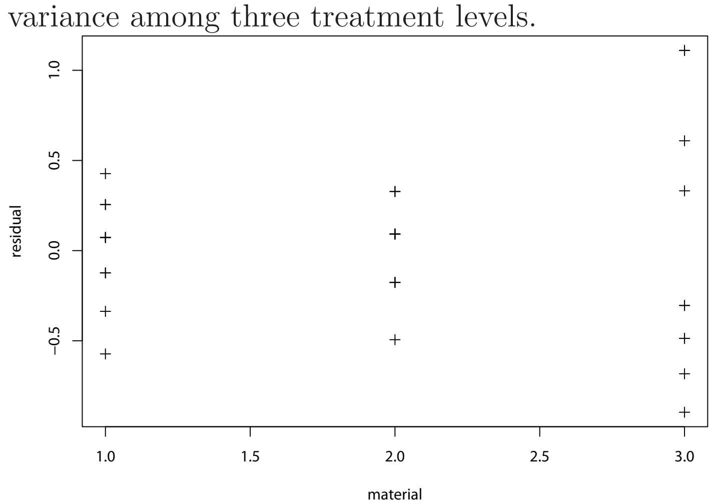
P-value= 0.6821. Decision: Do not reject H_0; the treatment means do not differ significantly.

8.32 (a) After a transformation $g(y) = \sqrt{y}$, we come up with an ANOVA table as follows.

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f
Materials	7.5123	2	3.7561	16.20
Error	6.2616	27	0.2319	
Total	13.7739	29		

(b) The P-value< 0.0001. Hence, there is significant difference in flaws among three materials.

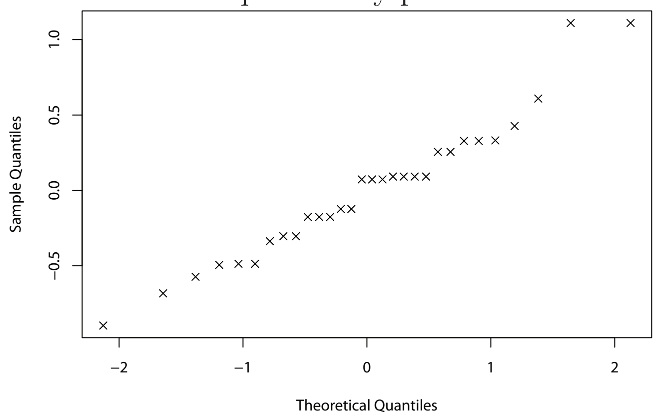
(c) A residual plot is given below and it does show some heterogeneity of the variance among three treatment levels.



(d) The purpose of the transformation is to stabilize the variances.

(e) One could be the distribution assumption itself. Once the data is transformed, it is not necessary that the data would follow a normal distribution.

(f) Here the normal probability plot on residuals is shown.



It appears to be close to a straight line. So, it is likely that the transformed data are normally distributed.

8.33 The hypotheses are

$$H_0: \alpha_1 = \alpha_2 = \alpha_3 = 0, \text{ dye effects are zero}$$

$$H_1: \text{At least one of the } \alpha_i\text{'s is not equal to zero.}$$

$\alpha = 0.05$.
Computation:

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f
Amounts	1238.8825	2	619.4413	122.37
Plants	53.7004	1	53.7004	
Error	101.2433	20	5.0622	
Total	1393.8263	23		

P-value< 0.0001. Decision: Reject H_0; color densities of fabric differ significantly for three levels of dyes.

8.34 (a) The hypotheses are

$$H_0: \sigma_\alpha^2 = 0,$$
$$H_1: \sigma_\alpha^2 \neq 0$$

$\alpha = 0.05$.
Computation:

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f
Treatments	23.238	3	7.746	3.33
Blocks	45.283	4	11.321	
Error	27.937	12	2.328	
Total	96.458	19		

P-value$= 0.0565$. Decision: Not able to show a significant difference in the random treatments at 0.05 level, although the P-value shows marginal significance.

(b) $\sigma_\alpha^2 = \frac{7.746-2.328}{5} = 1.084$, and $\sigma_\beta^2 = \frac{11.321-2.328}{4} = 2.248$.

8.35 (a) The model is $y_{ij} = \mu + \alpha_i + \epsilon_{ij}$, where $\alpha_i \sim n(0, \sigma_\alpha^2)$.

(b) Since $s^2 = 0.02056$ and $s_1^2 = 0.01791$, we have $\hat{\sigma}^2 = 0.02056$ and $\frac{s_1^2 - s^2}{10} = \frac{0.01791-0.02056}{10} = -0.00027$, which implies $\hat{\sigma}_\alpha^2 = 0$.

8.36 (a) The P-value of the test result is 0.7830. Hence, the variance component of pour is not significantly different from 0.

(b) We have the ANOVA table as follows:

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f
Pours	0.08906	4	0.02227	0.43
Error	1.02788	20	0.05139	
Total	1.11694	24		

Since $\frac{s_1^2 - s^2}{5} = \frac{0.02227-0.05139}{5} < 0$, we have $\hat{\sigma}_\alpha^2 = 0$.

8.37 (a) The hypotheses are

$$H_0 : \sigma_\alpha^2 = 0,$$
$$H_1 : \sigma_\alpha^2 \neq 0$$

$\alpha = 0.05$.

Computation:

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f
Operators	371.8719	3	123.9573	14.91
Error	99.7925	12	8.3160	
Total	471.6644	15		

P-value$= 0.0002$. Decision: Reject H_0; operators are different.

(b) $\hat{\sigma}^2 = 8.316$ and $\hat{\sigma}_\alpha^2 = \frac{123.9573-8.3160}{4} = 28.910$.

8.38 (a) $y_{ij} = \mu + \alpha_i + \epsilon_{ij}$, where $\alpha_i \sim n(x; 0, \sigma_\alpha^2)$.

(b) Running an ANOVA analysis, we obtain the P-value as 0.0121. Hence, the loom variance component is significantly different from 0 at level 0.05.

(c) The suspicion is supported by the data.

8.39 The hypotheses are

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4,$$
$$H_1 : \text{At least two of the } \mu_i\text{'s are not equal.}$$

$\alpha = 0.05$.

Computation:

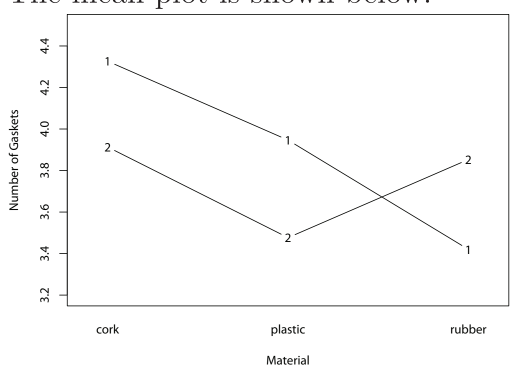
Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f
Garlon levels	3.7289	3	1.2430	1.57
Error	9.5213	12	0.7934	
Total	13.2502	15		

P-value= 0.2487. Decision: Do not reject H_0; there is insufficient evidence to claim that the concentration levels of Garlon would impact the heights of shoots.

8.40 Bartlett's statistic is $b = 0.8254$. Conclusion: do not reject homogeneous variance assumption.

8.41 (a) The process would include more than one stamping machine and the results might differ with different machines.

(b) The mean plot is shown below.



(c) Cork appears to be the best.

(d) Yes, there is interaction. Cork and plastic have better results with Machine 1, but rubber has better results with Machine 2. This interaction does not obscure the conclusion.

8.42 (a) The hypotheses for the Bartlett's test are

$$H_0 : \sigma_A^2 = \sigma_B^2 = \sigma_C^2 = \sigma_D^2,$$
$$H_1 : \text{The variances are not all equal.}$$

$\alpha = 0.05$.
Critical region: We have $n_1 = n_2 = n_3 = n_4 = 5$, $N = 20$, and $k = 4$. Therefore, we reject H_0 when $b < b_4(0.05, 5) = 0.5850$.
Computation: $s_A = 1.40819$, $s_B = 2.16056$, $s_C = 1.16276$, $s_D = 0.76942$ and hence $s_p = 1.46586$. From these, we can obtain that $b = 0.7678$.
Decision: Do not reject H_0; there is no sufficient evidence to conclude that the variances are not equal.

(b) The hypotheses are

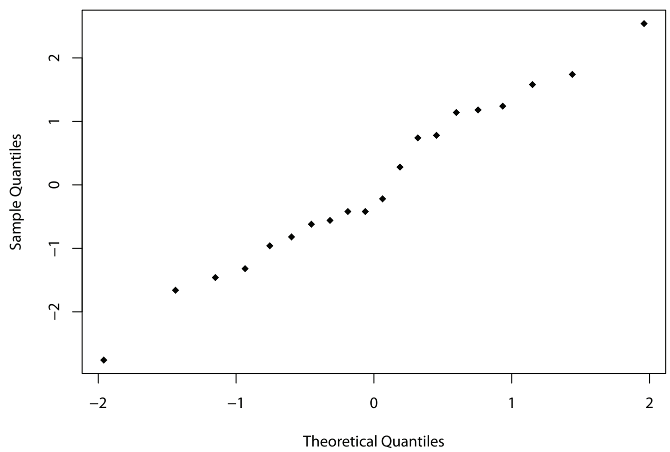
$$H_0 : \mu_A = \mu_B = \mu_C = \mu_D,$$
$$H_1 : \text{At least two of the } \mu_i\text{'s are not equal.}$$

$\alpha = 0.05$.
Computation:

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f
Laboratories	85.9255	3	28.6418	13.33
Error	34.3800	16	2.1488	
Total	120.3055	19		

P-value= 0.0001. Decision: Reject H_0; the laboratory means are significantly different.

(c) The normal probability plot is given as follows:



8.43 (a) The gasoline manufacturers would want to apply their results to more than one model of car.

(b) Yes, there is a significant difference in the miles per gallon for the three brands of gasoline.

(c) I would choose brand C for the best miles per gallon.

8.44 The hypotheses are

$$H_0 : \alpha_1 = \alpha_2 = \alpha_3 = 0,$$
$$H_1 : \text{At least one of the } \alpha_i\text{'s is not zero.}$$

$\alpha = 0.05$.
Computation:

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f
Diet	0.32356	2	0.16178	9.33
Error	0.20808	12	0.01734	
Total	0.53164	14		

with P-value= 0.0036.
Decision: Reject H_0; zinc is significantly different among the diets.

8.45 (a) The hypotheses are

$$H_0 : \alpha_1 = \alpha_2 = \alpha_3 = 0,$$
$$H_1 : \text{At least one of the } \alpha_i\text{'s is not zero.}$$

Computation:

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f
Paint Types	227875.11	2	113937.57	5.08
Error	336361.83	15	22424.12	
Total	564236.94	17		

with P-value= 0.0207.

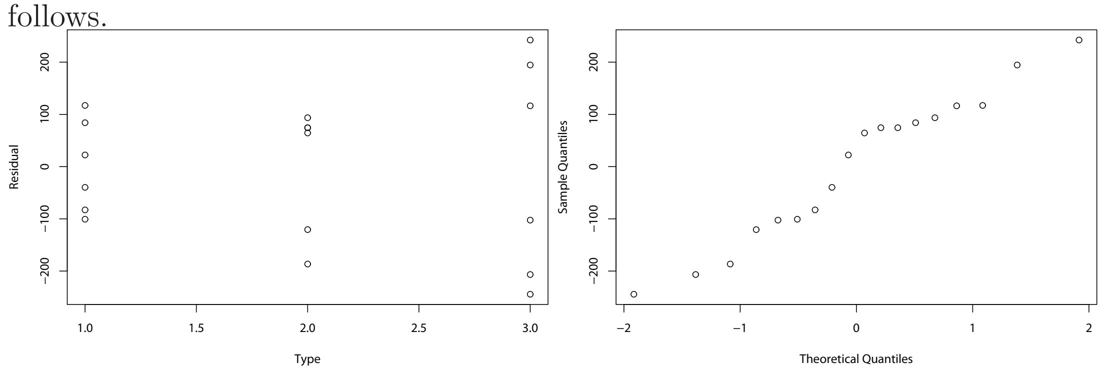
Decision: Reject H_0 at level 0.05; the average wearing quality differs significantly for three paints.

(b) Using Tukey's test, it turns out the following.

$\bar{y}_{1.}$	$\bar{y}_{3.}$	$\bar{y}_{2.}$
197.83	419.50	450.50

Types 2 and 3 are not significantly different, while Type 1 is significantly different from Type 2.

(c) We plot the residual plot and the normal probability plot for the residuals as follows.



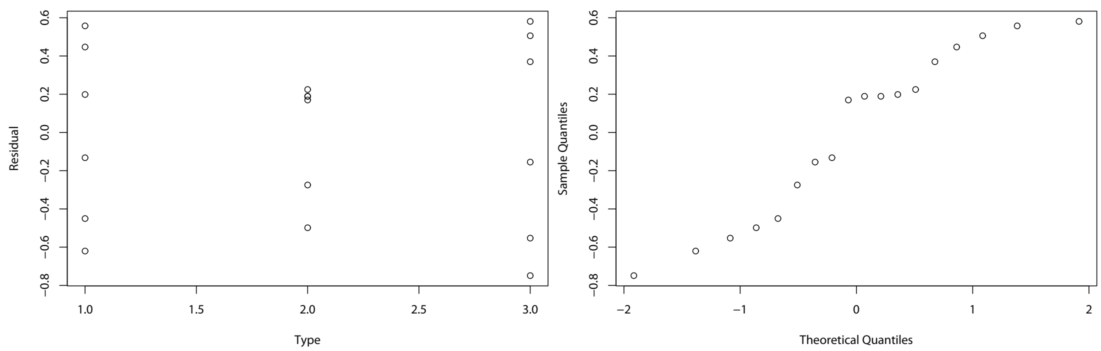
It appears that the homogeneity in variances may be violated, as is the normality assumption.

(d) We do a natural log transformation of the data, i.e., $y' = \log(y)$. The ANOVA result has changed as follows.

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f
Paint Types	2.6308	2	1.3154	6.07
Error	3.2516	15	0.2168	
Total	5.8824	17		

with P-value= 0.0117.

Decision: Reject H_0 at level 0.05; the average wearing quality differ significantly for three paints. The residual and normal probability plots are shown here:



While the homogeneity of the variances seem to be a little better, the normality assumption may still be invalid.

8.46 (a) The hypotheses are

$$H_0 : \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0,$$
$$H_1 : \text{At least one of the } \alpha_i\text{'s is not zero.}$$

Computation:

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f
Locations	0.01594	3	0.00531	13.80
Error	0.00616	16	0.00039	
Total	0.02210	19		

with P-value= 0.0001.

Decision: Reject H_0; the mean ozone levels differ significantly across the locations.

(b) Using Tukey's test, the results are as follows.

$\bar{y}_{4.}$	$\bar{y}_{1.}$	$\bar{y}_{3.}$	$\bar{y}_{2.}$
0.078	0.092	0.096	0.152

Location 2 appears to have much higher ozone measurements than other locations.

Chapter 9

Factorial Experiments (Two or More Factors)

9.1 The hypotheses of the three parts are,

(a) for the main effects temperature,

$$H_0^{'}: \alpha_1 = \alpha_2 = \alpha_3 = 0,$$
$$H_1^{'}: \text{At least one of the } \alpha_i\text{'s is not zero;}$$

(b) for the main effects ovens,

$$H_0^{''}: \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0,$$
$$H_1^{''}: \text{At least one of the } \beta_i\text{'s is not zero;}$$

(c) and for the interactions,

$$H_0^{'''}: (\alpha\beta)_{11} = (\alpha\beta)_{12} = \cdots = (\alpha\beta)_{34} = 0,$$
$$H_1^{'''}: \text{At least one of the } (\alpha\beta)_{ij}\text{'s is not zero.}$$

$\alpha = 0.05$.
Critical regions: (a) $f_1 > 3.00$; (b) $f_2 > 3.89$; and (c) $f_3 > 3.49$.
Computations: From the computer printout we have

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f
Temperatures	5194.08	2	2597.0400	8.13
Ovens	4963.12	3	1654.3733	5.18
Interaction	3126.26	6	521.0433	1.63
Error	3833.50	12	319.4583	
Total	17,116.96	23		

Decision: (a) Reject H_0'; (b) Reject H_0''; (c) Do not reject H_0'''.

9.2 The hypotheses of the three parts are,

(a) for the main effects brands,

$$H_0' : \alpha_1 = \alpha_2 = \alpha_3 = 0,$$
$$H_1' : \text{At least one of the } \alpha_i\text{'s is not zero;}$$

(b) for the main effects times,

$$H_0'' : \beta_1 = \beta_2 = \beta_3 = 0,$$
$$H_1'' : \text{At least one of the } \beta_i\text{'s is not zero;}$$

(c) and for the interactions,

$$H_0''' : (\alpha\beta)_{11} = (\alpha\beta)_{12} = \cdots = (\alpha\beta)_{33} = 0,$$
$$H_1''' : \text{At least one of the } (\alpha\beta)_{ij}\text{'s is not zero.}$$

$\alpha = 0.05$.
Critical regions: (a) $f_1 > 3.35$; (b) $f_2 > 3.35$; and (c) $f_3 > 2.73$.
Computations: From the computer printout we have

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f
Brands	32.7517	2	16.3758	1.74
Times	227.2117	2	113.6058	12.04
Interaction	17.3217	4	4.3304	0.46
Error	254.7025	27	9.4334	
Total	531.9875	35		

Decision: (a) Do not reject H_0'; (b) Reject H_0''; (c) Do not reject H_0'''.

9.3 The hypotheses of the three parts are,

(a) for the main effects environments,

$$H_0' : \alpha_1 = \alpha_2 = 0, \text{ (no differences in the environment)}$$
$$H_1' : \text{At least one of the } \alpha_i\text{'s is not zero;}$$

(b) for the main effects strains,

$$H_0'' : \beta_1 = \beta_2 = \beta_3 = 0, \text{ (no differences in the strains)}$$
$$H_1'' : \text{At least one of the } \beta_i\text{'s is not zero;}$$

(c) and for the interactions,

$H_0''': (\alpha\beta)_{11} = (\alpha\beta)_{12} = \cdots = (\alpha\beta)_{23} = 0,$ (environments and strains do not interact)
H_1''' : At least one of the $(\alpha\beta)_{ij}$'s is not zero.

$\alpha = 0.01$.
Critical regions: (a) $f_1 > 7.29$; (b) $f_2 > 5.16$; and (c) $f_3 > 5.16$.
Computations: From the computer printout we have

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f
Environments	14,875.521	1	14,875.521	14.81
Strains	18,154.167	2	9,077.083	9.04
Interaction	1,235.167	2	617.583	0.61
Error	42,192.625	42	1004.586	
Total	76,457.479	47		

Decision: (a) Reject H_0'; (b) Reject H_0''; (c) Do not reject H_0'''. Interaction is not significant, while both main effects, environment and strain, are all significant.

9.4 (a) The hypotheses of the three parts are,

$$H_0' : \alpha_1 = \alpha_2 = \alpha_3 = 0$$
$$H_1' : \text{At least one of the } \alpha_i\text{'s is not zero;}$$
$$H_0'' : \beta_1 = \beta_2 = \beta_3 = 0,$$
$$H_1'' : \text{At least one of the } \beta_i\text{'s is not zero;}$$
$$H_0''' : (\alpha\beta)_{11} = (\alpha\beta)_{12} = \cdots = (\alpha\beta)_{33} = 0,$$
$$H_1''' : \text{At least one of the } (\alpha\beta)_{ij}\text{'s is not zero.}$$

$\alpha = 0.01$.
Critical regions: for H_0', $f_1 > 3.21$; for H_0'', $f_2 > 3.21$; and for H_0''', $f_3 > 2.59$.
Computations: From the computer printout we have

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f
Coating	1,535,021.37	2	767,510.69	6.87
Humidity	1,020,639.15	2	510,319.57	4.57
Interaction	1,089,989.63	4	272,497.41	2.44
Error	5,028,396.67	45	111,742.15	
Total	76,457.479	47		

Decision: Reject H_0'; Reject H_0''; Do not reject H_0'''. Coating and humidity do not interact, while both main effects are all significant.

(b) The three means for the humidity are $\bar{y}_L = 733.78$, $\bar{y}_M = 406.39$ and $\bar{y}_H = 638.39$. Using Duncan's test, the means can be grouped as

$$\underline{\bar{y}_M \quad \bar{y}_L}\quad \underline{\bar{y}_H}$$
$$\underline{406.39 \quad 638.39}\quad \underline{733.78}$$

Therefore, corrosion damage is different for medium humidity than for low or high humidity.

9.5 The hypotheses of the three parts are,

(a) for the main effects subjects,

$$H_0^{'} : \alpha_1 = \alpha_2 = \alpha_3 = 0,$$
$$H_1^{'} : \text{At least one of the } \alpha_i\text{'s is not zero;}$$

(b) for the main effects muscles,

$$H_0^{''} : \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0,$$
$$H_1^{''} : \text{At least one of the } \beta_i\text{'s is not zero;}$$

(c) and for the interactions,

$$H_0^{'''} : (\alpha\beta)_{11} = (\alpha\beta)_{12} = \cdots = (\alpha\beta)_{35} = 0,$$
$$H_1^{'''} : \text{At least one of the } (\alpha\beta)_{ij}\text{'s is not zero.}$$

$\alpha = 0.01$.
Critical regions: (a) $f_1 > 5.39$; (b) $f_2 > 4.02$; and (c) $f_3 > 3.17$.
Computations: From the computer printout we have

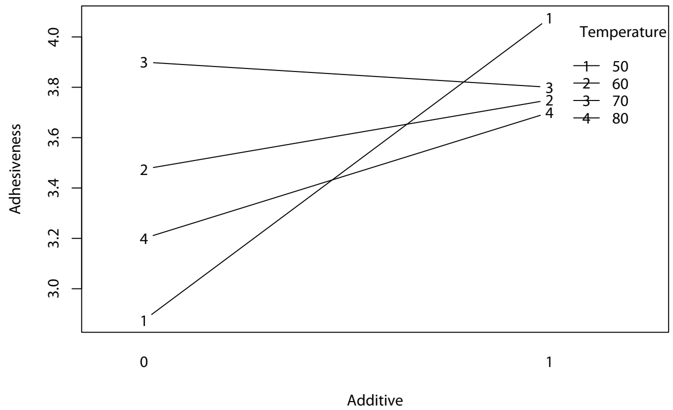
Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f
Subjects	4,814.74	2	2,407.37	34.40
Muscles	7,543.87	4	1,885.97	26.95
Interaction	11,362.20	8	1,420.28	20.30
Error	2,099.17	30	69.97	
Total	25,819.98	44		

Decision: (a) Reject $H_0^{'}$; (b) Reject $H_0^{''}$; (c) Reject $H_0^{'''}$.

9.6 The ANOVA table is shown as

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f	P-value
Additive	1.7578	1	1.7578	22.29	< 0.0001
Temperature	0.8059	3	0.2686	3.41	0.0338
Interaction	1.7934	3	0.5978	7.58	0.0010
Error	1.8925	24	0.0789		
Total	6.2497	31			

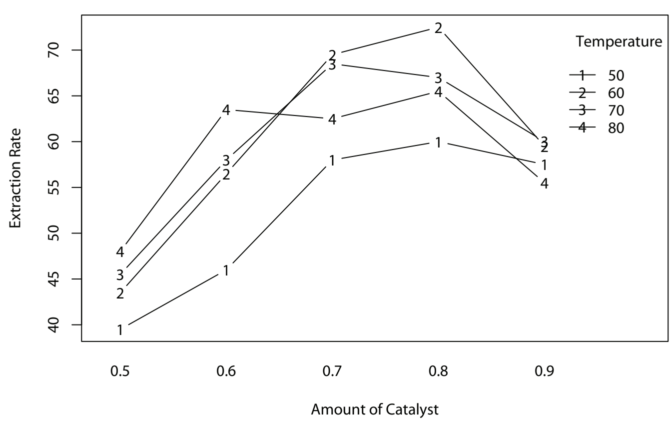
Decision: All main effects and interaction are significant.
An interaction plot is given here.



9.7 The ANOVA table is

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f	P-value
Temperature	430.475	3	143.492	10.85	0.0002
Catalyst	2,466.650	4	616.663	46.63	< 0.0001
Interaction	326.150	12	27.179	2.06	0.0745
Error	264.500	20	13.225		
Total	3,487.775	39			

Decision: All main effects are significant and the interaction is significant at level 0.0745. Hence, if 0.05 significance level is used, interaction is not significant.
An interaction plot is given here.



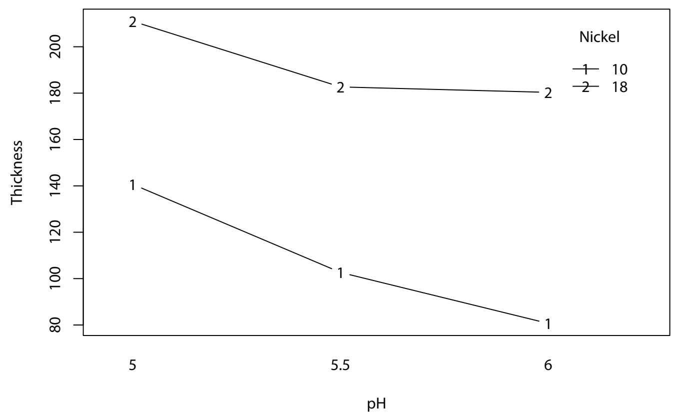
9.8 (a) The ANOVA table is

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f	P-value
Nickel	31,250.00	1	31,250.00	44.52	< 0.0001
pH	6,606.33	2	3,303.17	4.71	0.0310
Nickel*pH	670.33	2	335.17	0.48	0.6316
Error	8,423.33	12	701.94		
Total	46,950.00	17			

Decision: Nickel contents and levels of pH do not interact to each other, while both main effects of nickel contents and levels of pH are all significant, at level higher than 0.0310.

(b) In comparing the means of the six treatment combinations, a nickel content level of 18 and a pH level of 5 resulted in the largest values of thickness.

(c) The interaction plot is given here and it shows no apparent interactions.

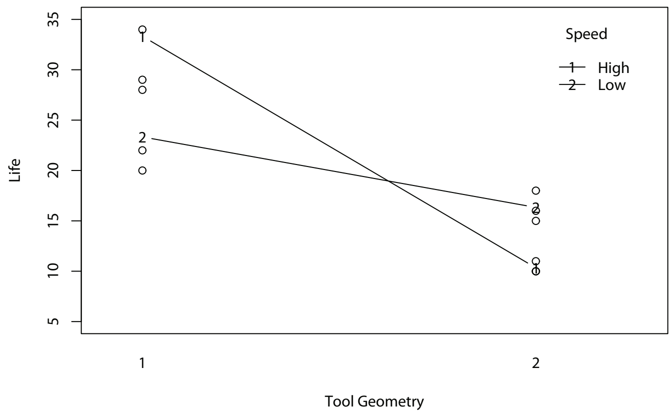


9.9 (a) The ANOVA table is

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f	P-value
Tool	675.00	1	675.00	74.31	< 0.0001
Speed	12.00	1	12.00	1.32	0.2836
Tool*Speed	192.00	1	192.00	21.14	0.0018
Error	72.67	8	9.08		
Total	951.67	11			

Decision: The interaction effects are significant. Although the main effects of speed showed insignificance, we might not make such a conclusion since its effects might be masked by significant interaction.

(b) In the graph shown, we claim that the cutting speed that results in the longest life of the machine tool depends on the tool geometry, although the variability of the life is greater with tool geometry at level 1.



(c) Since interaction effects are significant, we do the analysis of variance for separate tool geometry.

(i) For tool geometry 1, an f-test for the cutting speed resulted in a P-value = 0.0405 with the mean life (standard deviation) of the machine tool at 33.33 (4.04) for high speed and 23.33 (4.16) for low speed. Hence, a high cutting speed has longer life for tool geometry 1.

(ii) For tool geometry 2, an f-test for the cutting speed resulted in a P-value = 0.0031 with the mean life (standard deviation) of the machine tool at 10.33 (0.58) for high speed and 16.33 (1.53) for low speed. Hence, a low cutting speed has longer life for tool geometry 2.

For the above detailed analysis, we note that the standard deviations for the mean life are much higher at tool geometry 1.

(d) See part (b).

9.10 (a) $y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk}$, $i = 1, 2, \ldots, a$; $j = 1, 2, \ldots, b$; $k = 1, 2, \ldots, n$.

(b) The ANOVA table is

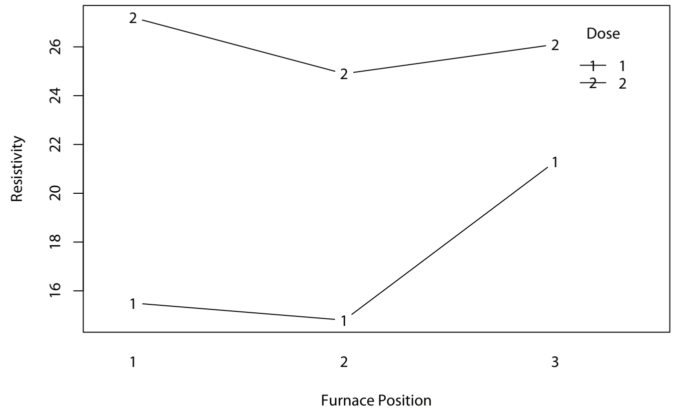
Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f	P-value
Dose	117.9267	1	117.9267	18.08	0.0511
Position	15.0633	2	7.5317	1.15	0.4641
Error	13.0433	2	6.5217		
Total	146.0333	5			

(c) $(n-1) - (a-1) - (b-1) = 5 - 1 - 2 = 2$.

(d) At level 0.05, Tukey's result for the furnace position is shown here:

$\bar{y}_2$	$\bar{y}_1$	$\bar{y}_3$
19.850	21.350	23.700

Although Tukey's multiple comparisons resulted in insignificant differences among the furnace position levels, based on a P-value of 0.0511 for the Dose and on the plot shown we can see that Dose=2 results in higher resistivity.



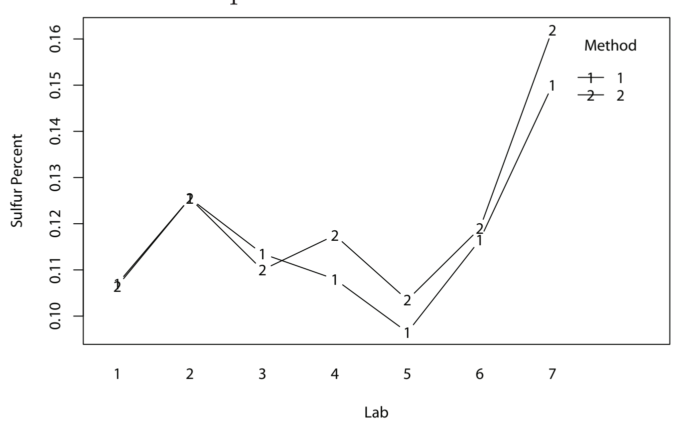
9.11 (a) The ANOVA table is

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f	P-value
Method	0.000104	1	0.000104	6.57	0.0226
Lab	0.008058	6	0.001343	84.70	< 0.0001
Method*Lab	0.000198	6	0.000033	2.08	0.1215
Error	0.000222	14	0.000016		
Total	0.00858243	27			

(b) Since the P-value $= 0.1215$ for the interaction, the interaction is not significant. Hence, the results on the main effects can be considered meaningful to the scientist.

(c) Both main effects, method of analysis and laboratory, are all significant.

(d) The interaction plot is show here.



(e) When the tests are done separately, i.e., we only use the data for Lab 1, or Lab 7 alone, the P-values for testing the differences of the methods at Lab 1 and 7 are 0.8600 and 0.1557, respectively. Using partial data such as we did here as above, the degrees of freedom in errors are often smaller (2 in both cases discussed here). Hence, we do not have much power to detect the difference between treatments.

However, if we compare the treatment differences within the full ANOVA model, the degrees of freedom in error can be quite large, e.g., 18 in this case. So, for this case, we obtain the P-values for testing the difference of the methods at Lab 1 and 7 as 0.9010 and 0.0093, respectively. Hence, methods are no difference in Lab 1 and are significantly different in Lab 7. Similar results may be found in the interaction plot in (d).

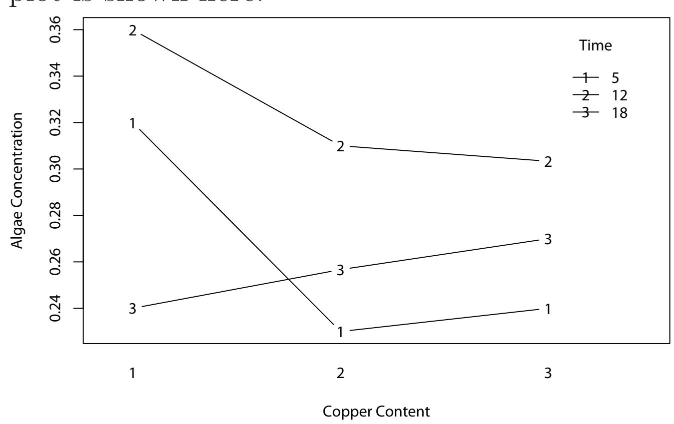
9.12 (a) The ANOVA table is

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f	P-value
Time	0.025622	2	0.012811	38.87	< 0.0001
Copper	0.008956	2	0.004478	13.58	0.0003
Time*Copper	0.012756	4	0.003189	9.67	0.0002
Error	0.005933	18	0.000330		
Total	0.053267	26			

(b) The P-value < 0.0001. There is a significant time effect.

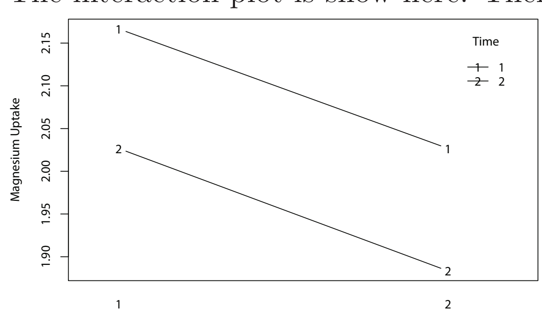
(c) The P-value $= 0.0003$. There is a significant copper effect.

(d) The interaction effect is significant since the P-value $= 0.0002$. The interaction plot is shown here.



The algae concentrations for the various copper contents are all clearly influenced by the time effect shown in the graph.

9.13 (a) The interaction plot is show here. There seems no interaction effect.



(b) The ANOVA table is

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f	P-value
Treatment	0.060208	1	0.060208	157.07	< 0.0001
Time	0.060208	1	0.060208	157.07	< 0.0001
Treatment*Time	0.000008	1	0.000008	0.02	0.8864
Error	0.003067	8	0.000383		
Total	0.123492	11			

(c) The magnesium uptake are lower using treatment 2 than using treatment 1, no matter what the times are. Also, time 2 has lower magnesium uptake than time 1. All the main effects are significant.

(d) Using the regression model and making "Treatment" as categorical, we have the following fitted model:

$$\hat{y} = 2.4433 - 0.13667 \text{ Treatment} - 0.13667 \text{ Time} - 0.00333 \text{ Treatment} \times \text{Time}.$$

(e) The P-value of the interaction for the above regression model is 0.8864 and hence it is insignificant.

9.14 (a) A natural linear model with interaction would be

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2.$$

The fitted model would be

$$\hat{y} = 0.41772 - 0.06631 x_1 - 0.00866 x_2 + 0.00416 x_1 x_2,$$

with the P-values of the t-tests on each of the coefficients as 0.0092, 0.0379 and 0.0318 for x_1, x_2, and x_1x_2, respectively. They are all significant at a level larger than 0.0379. Furthermore, $R^2_{\text{adj}} = 0.1788$.

(b) The new fitted model is

$$\hat{y} = 0.3368 - 0.15965 x_1 + 0.02684 x_2 + 0.00416 x_1 x_2 + 0.02333 x_1^2 - 0.00155 x_2^2,$$

with P-values of the t-tests on each of the coefficients as 0.0004, < 0.0001, 0.0003, 0.0156, and < 0.0001 for x_1, x_2. x_1x_2, x_1^2, and x_2^2, respectively. Furthermore, $R^2_{\text{adj}} = 0.7700$ which is much higher than that of the model in (a). Model in (b) would be more appropriate.

9.15 The ANOVA table is given as follows.

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f	P-value
Launderings	202.13	1	202.13	7.55	0.0087
Bath	715.34	1	715.34	26.73	< 0.0001
Launderings*Bath	166.14	1	166.14	6.21	0.0166
Error	1177.68	44	26.77		
Total	2261.28	47			

(a) The interaction is significant if the level of significance is larger than 0.0166. So, using traditional 0.05 level of significance we would claim that the interaction is significant.

(b) Looking at the ANOVA, it seems that both main effects are significant.

9.16 (a) When only A, B, C, and BC factors are in the model, the P-value for BC interaction is 0.0806. Hence at level of 0.05, the interaction is insignificant.

(b) When the sum of squares of the BC term is pooled with the sum of squares of the error, we increase the degrees of freedom of the error term. The P-values of the main effects of A, B, and C are 0.0275, 0.0224, and 0.0131, respectively. All these are significant.

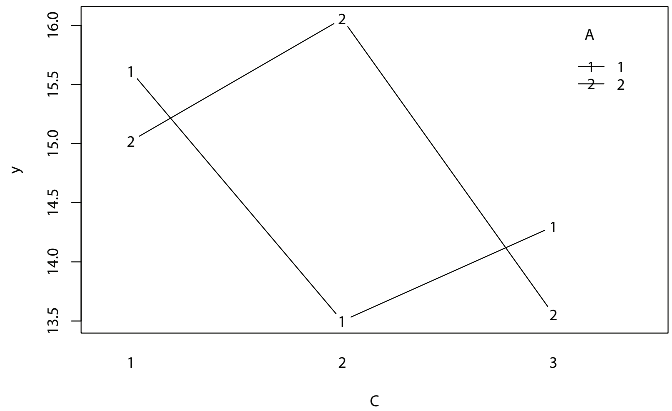
9.17 The ANOVA table is given here.

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f	P-value
Main effect					
A	2.24074	1	2.24074	0.54	0.4652
B	56.31815	2	28.15907	6.85	0.0030
C	17.65148	2	8.82574	3.83	0.1316
Two-factor Interaction					
AB	31.47148	2	15.73574	3.83	0.0311
AC	31.20259	2	15.60130	3.79	0.0320
BC	21.56074	4	5.39019	1.31	0.2845
Three-factor Interaction					
ABC	26.79852	4	6.69963	1.63	0.1881
Error	148.04000	36	4.11221		
Total	335.28370	53			

(a) Based on the P-values, only AB and AC interactions are significant.

(b) The main effect B is significant. However, due to significant interactions mentioned in (a), the insignificance of A and C cannot be counted.

(c) Look at the interaction plot of the mean responses versus C for different cases of A.



Apparently, the mean responses at different levels of C varies in different patterns for the different levels of A. Hence, although the overall test on factor C is insignificant, it is misleading since the significance of the effect C is masked by the significant interaction between A and C.

9.18 The ANOVA table is given here.

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f	P-value
A	1.90591	3	0.63530	8.28	0.0003
B	0.02210	1	0.02212	0.29	0.5951
C	38.93402	1	38.93402	507.57	< 0.0001
AB	0.88632	3	0.29544	3.85	0.0185
AC	0.53594	3	0.17865	2.33	0.0931
BC	0.45435	1	0.45435	5.92	0.0207
ABC	0.42421	3	0.14140	1.84	0.1592
Error	2.45460	32	0.07671		
Total	45.61745	47			

(a) Two-way interactions of AB and BC are all significant and main effect of A and C are all significant. The insignificance of the main effect B may not be valid due to the significant BC interaction.

(b) Based on the P-values, Duncan's tests and the interaction means, the most important factor is C and using $C = 2$ is the most important way to increase percent weight. Also, using factor A at level 1 is the best.

9.19 Letting A, B, and C designate coating, humidity, and stress, respectively, the ANOVA table is given here.

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f	P-value
Main effect					
A	$216,384.1$	1	$216,384.1$	0.05	0.8299
B	$19,876,891.0$	2	$9,938,445.5$	2.13	0.1257
C	$427,993,946.4$	2	$213,996,973.2$	45.96	< 0.0001
Two-factor Interaction					
AB	$31,736,625$	2	$15,868,312.9$	3.41	0.0385
AC	$699,830.1$	2	$349,915.0$	0.08	0.9277
BC	$58,623,693.2$	4	$13,655,923.3$	3.15	0.0192
Three-factor Interaction					
ABC	$36,034,808.9$	4	$9,008,702.2$	1.93	0.1138
Error	$335,213,133.6$	72	$4,655,738.0$		
Total	$910,395,313.1$	89			

(a) The Coating and Humidity interaction, and the Humidity and Stress interaction have the P-values of 0.0385 and 0.0192, respectively. Hence, they are all significant. On the other hand, the Stress main effect is strongly significant as well. However, both other main effects, Coating and Humidity, cannot be claimed as insignificant, since they are all part of the two significant interactions.

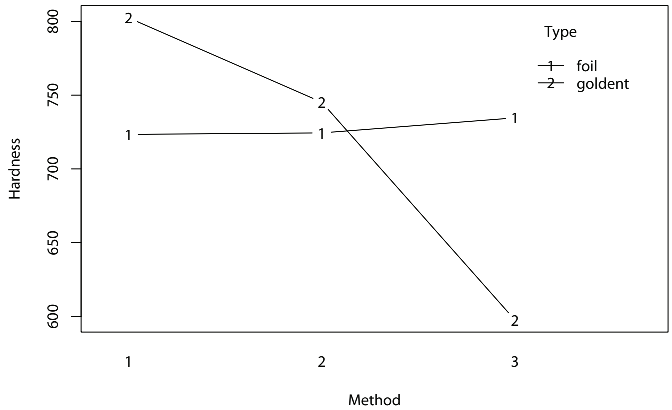
(b) A Stress level of 20 consistently produces low fatigue. It appears to work best with medium humidity and an uncoated surface.

9.20 (a) $y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + (\beta\gamma)_{jk} + \epsilon_{ijk}$; $\sum_j \beta_j = 0$, $\sum_k \gamma_k = 0$, $\sum_j (\beta\gamma)_{jk} = 0$, $\sum_k (\beta\gamma)_{jk} = 0$, and $\epsilon_{ijk} \sim n(x; 0, \sigma^2)$.

(b) The P-value of the Method and Type of Gold interaction is 0.10587. Hence, the interaction is at least marginally significant.

(c) The best method depends on the type of gold used.
The tests of the method effect for different type of gold yields the P-values as 0.9801 and 0.0099 for "Gold Foil" and "Goldent", respectively. Hence, the methods are significantly different for the "Goldent" type.

Here is an interaction plot.

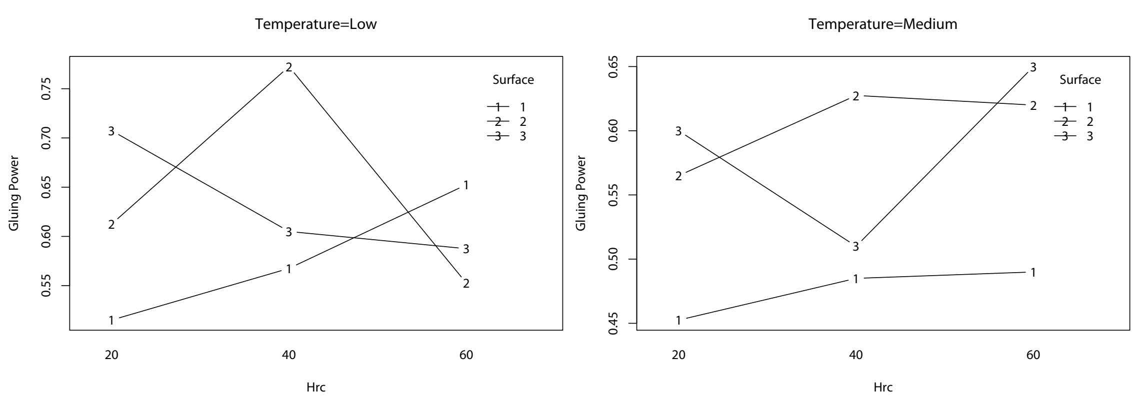


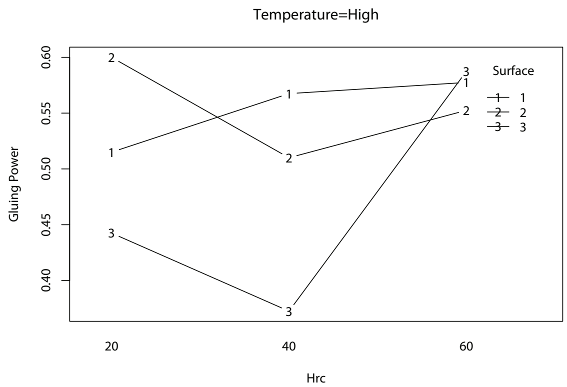
It appears that when Type is "Goldent" and Method is 1, it yields the best hardness.

9.21 The ANOVA table shows:

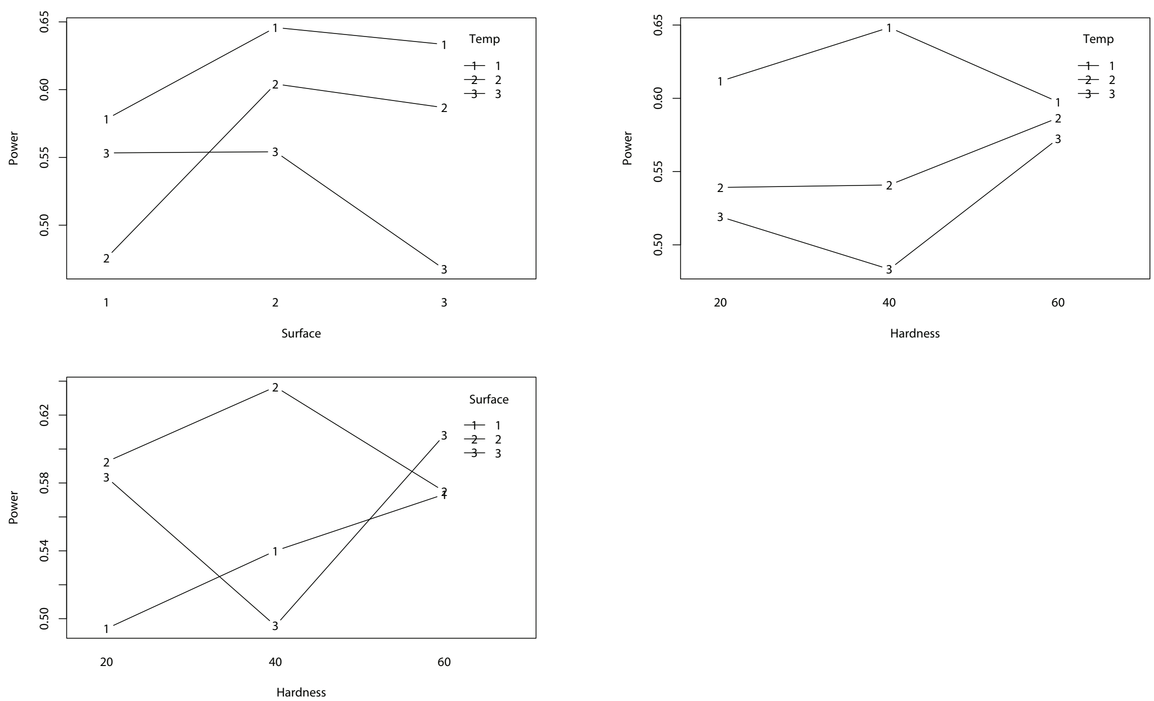
Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f	P-value
A	0.16617	2	0.08308	14.22	< 0.0001
B	0.07825	2	0.03913	6.70	0.0020
C	0.01947	2	0.00973	1.67	0.1954
AB	0.12845	4	0.03211	5.50	0.0006
AC	0.06280	4	0.01570	2.69	0.0369
BC	0.12644	4	0.03161	5.41	0.0007
ABC	0.14224	8	0.01765	3.02	0.0051
Error	0.47323	81	0.00584		
Total	1.19603	107			

There is a significant three-way interaction by Temperature, Surface, and Hrc. A plot of each Temperature is given to illustrate the interaction





9.22 The two-way interaction plots are given here and they all show significant interactions.



9.23 (a) Yes, the P-values for $Brand * Type$ and $Brand * Temp$ are both < 0.0001.

(b) The main effect of Brand has a P-value < 0.0001. So, three brands averaged across the other two factore are significantly different.

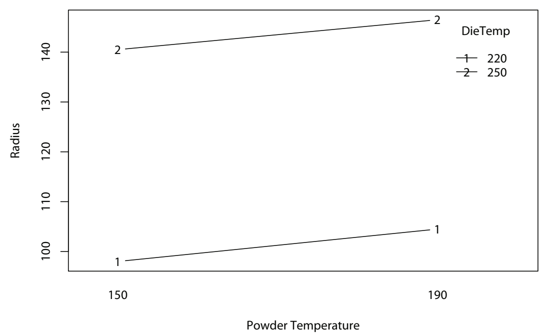
(c) Using brand Y, powdered detergent and hot water yields the highest percent removal of dirt.

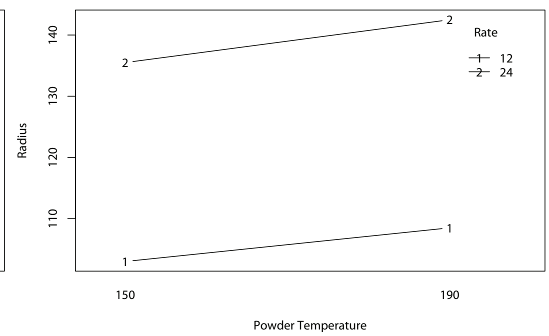
9.24 (a) Define A, B, and C as "Powder Temperature," "Die Temperature," and "Extrusion Rate," respectively. The ANOVA table shows:

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f	P-value
A	78.125	1	78.125	125.00	0.0079
B	3570.125	1	3570.125	5712.20	0.0002
C	2211.125	1	2211.125	3537.80	0.0003
AB	0.125	1	0.125	0.20	0.6985
AC	1.125	1	1.125	1.80	0.3118
Error	1.25	2	0.625		
Total	5861.875	7			

The ANOVA results only show that the main effects are all significant and no two-way interactions are significant.

(b) The interaction plots are shown here.

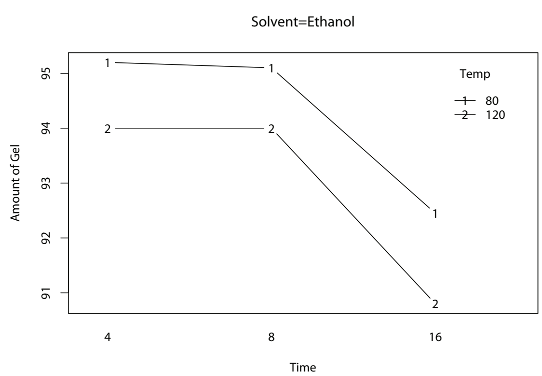


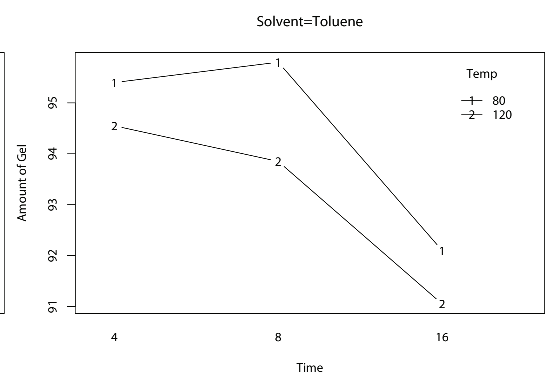


(c) The interaction plots in part (b) are consistent with the findings in part (a) that no two-way interactions present.

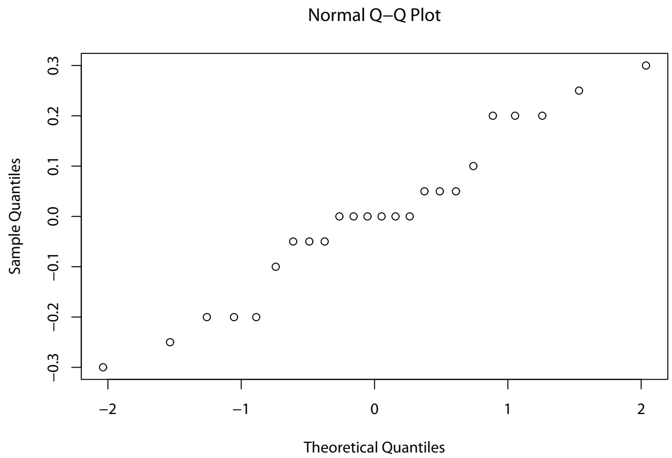
9.25 (a) The P-values of two-way interactions Time$\times$Temperature, Time$\times$Solvent, Temperature $\times$ Solvent, and the P-value of the three-way interaction Time$\times$Temperature$\times$Solvent are 0.1103, 0.1723, 0.8558, and 0.0140, respectively.

(b) The interaction plots for different levels of Solvent are given here.





(c) A normal probability plot of the residuals is given and it appears that normality assumption may not be valid.

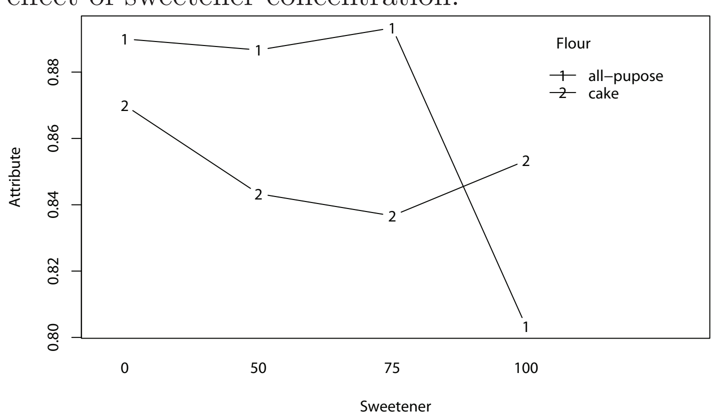


9.26 (a) The ANOVA table is given.

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f	P-value
Sweetener	0.00871	3	0.00290	2.90	0.0670
Flour	0.00184	1	0.00184	1.84	0.1941
Interaction	0.01015	3	0.00338	3.38	0.0442
Error	0.01600	16	0.00100		
Total	0.03670	23			

Sweetener factor is close to be significant, while the P-value of the Flour shows insignificance. However, the interaction effects appear to be significant.

(b) Since the interaction is significant with a P-value $= 0.0442$, testing the effect of sweetener on the specific gravity of the cake samples by flour type we get P-value $= 0.0077$ for "All Purpose" flour and P-value $= 0.6059$ for "Cake" flour. We also have the interaction plot which shows that sweetener at 100% concentration yielded a specific gravity significantly lower than the other concentrations for all-purpose flour. For cake flour, however, there were no big differences in the effect of sweetener concentration.



9.27 (a) The ANOVA table is given.

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f	P-value
Sauce	1,031.3603	1	1,031.3603	7.32	0.0123
Fish	16,505.8640	2	8,252.9320	58.58	< 0.0001
Sauce*Fish	724.6107	2	362.3053	2.57	0.0973
Error	3,381.1480	24	140.8812		
Total	21,642.9830	29			

Interaction effect is not significant.

(b) Both P-values of Sauce and Fish Type are all small enough to call significance.

9.28 (a) The ANOVA table is given here.

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f	P-value
Plastic Type	142.6533	2	71.3267	16.92	0.0003
Humidity	143.7413	3	47.9138	11.36	0.0008
Interaction	133.9400	6	22.3233	5.29	0.0070
Error	50.5950	12	4.2163		
Total	470.9296	23			

The interaction is significant.

(b) The SS for AB with only Plastic Type A and B is 24.8900 with 3 degrees of freedom. Hence $f = \frac{24.8900/3}{4.2163} = 1.97$ with P-value $= 0.1727$. Hence, there is no significant interaction when only A and B are considered.

9.29 (a) The ANOVA table is given here.

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f	P-value
Environment	0.8624	1	0.8624	15.25	0.0021
Stress	40.8140	2	20.4020	360.94	< 0.0001
Interaction	0.0326	2	0.0163	0.29	0.7547
Error	0.6785	12	0.0565		
Total	42.3875	17			

The interaction is insignificant.

(b) The mean fatigue life for the two main effects are all significant.

9.30 The ANOVA table is given here.

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f	P-value
Sweetener	1.26893	3	0.42298	2.65	0.0843
Flour	1.77127	1	1.77127	11.09	0.0042
Interaction	0.14647	3	0.04882	0.31	0.8209
Error	2.55547	16	0.15972		
Total	5.74213	23			

The interaction effect is insignificant. The main effect of Sweetener is somewhat insignificant, since the P-value $= 0.0843$. The main effect of Flour is strongly significant.

9.31 The ANOVA table is given here.

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f	P-value
	1133.5926	2	566.7963	8.30	0.0016
A	26896.2593	2	13448.1296	196.91	< 0.0001
B	40.1482	2	20.0741	0.29	0.7477
C	216.5185	4	54.1296	0.79	0.5403
AB	1.6296	4	0.4074	0.01	0.9999
AC	2.2963	4	0.5741	0.01	0.9999
Error	2.5926	8	0.3241	0.00	1.0000
	1844.0000	27	68.2963		
Total	30137.0370	53			

All the two-way and three-way interactions are insignificant. In the main effects, only A and B are significant.

9.32 (a) Treating Solvent as a class variable and Temperature and Time as continuous variable, only three terms in the ANOVA model show significance. They are (1) Intercept; (2) Coefficient for Temperature and (3) Coefficient for Time.

(b) Due to the fact that none of the interactions are significant, we can claim that the models for ethanol and toluene are equivalent apart from the intercept.

(c) The three-way interaction in Exercise 9.25 was significant. However, the general patterns of the gel generated are pretty similar for the two Solvent levels.

9.33 The ANOVA table is displayed.

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f	P-value
Surface	2.22111	2	1.11056	0.02	0.9825
Pressure	39.10778	2	19.55389	0.31	0.7402
Interaction	112.62222	4	28.15556	0.45	0.7718
Error	565.72000	9	62.85778		
Total	719.67111	17			

All effects are insignificant.

9.34 (a) This is a two-factor fixed-effects model with interaction.

$$y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk},$$
$$\sum_i \alpha_i = 0, \ \sum_j \beta_j = 0, \ \sum_i (\alpha\beta)_{ij} = \sum_j (\alpha\beta)_{ij} = 0, \ \epsilon_{ijk} \sim n(x; 0, \sigma)$$

(b) The ANOVA table is displayed.

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f	P-value
Time	0.16668	3	0.05556	48.67	< 0.0001
Temperature	0.27151	2	0.13575	118.91	< 0.0001
Interaction	0.03209	6	0.00535	4.68	0.0111
Error	0.01370	12	0.00114		
Total	0.48398	23			

The interaction is insignificant, while two main effects are significant.

(c) It appears that using a temperature of $-20°C$ with drying time of 2 hours would speed up the process and still yield a flavorful coffee. It might be useful to try some additional runs at this combination.

9.35 (a) Since it is more reasonable to assume the data come from Poisson distribution, it would be dangerous to use standard analysis of variance because the normality assumption would be violated. It would be better to transform the data to get at least stable variance.

(b) The ANOVA table is displayed.

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f	P-value
Teller	1379.42	3	459.81	19.91	< 0.0001
Time	465.79	2	232.90	10.08	0.0003
Interaction	165.21	6	27.53	1.19	0.3326
Error	831.50	36	23.10		
Total	2841.92	47			

The interaction effect is insignificant. Two main effects, Teller and Time, are all significant.