INSTRUCTOR'S SOLUTIONS MANUAL

INTRODUCTION TO MATHEMATICAL STATISTICS

SEVENTH EDITION

Robert Hogg

University of Iowa

Joseph McKean

Western Michigan University

Allen Craig

PEARSON

Boston Columbus Indianapolis New York San Francisco Upper Saddle River
Amsterdam Cape Town Dubai London Madrid Milan Munich Paris Montreal Toronto
Delhi Mexico City Sao Paulo Sydney Hong Kong Seoul Singapore Taipei Tokyo

Reproduced by Pearson from electronic files supplied by the author.

Publishing as Pearson, 75 Arlington Street, Boston, MA 02116.

ISBN-13: 978-0-321-79565-6
ISBN-10: 0-321-79565-2

www.pearsonhighered.com

PEARSON

Contents

Chapter 1

Probability and Distributions

1.2.1 Part (c): $C_1 \cap C_2 = \{(x, y) : 1 < x < 2, 1 < y < 2\}$.

1.2.3 $C_1 \cap C_2 = \{\text{mary,mray}\}$.

1.2.6 $C_k = \{x : 1/k \le x \le 1 - (1/k)\}$.

1.2.7 $C_k = \{(x, y) : 0 \le x \le 1/k, 0 \le y \le 1/k\}$.

1.2.8 $\lim_{k\to\infty} C_k = \{x : 0 < x < 3\}$. Note: neither the number 0 nor the number 3 is in any of the sets C_k, $k = 1, 2, 3, \ldots$

1.2.9 Part (b): $\lim_{k\to\infty} C_k = \phi$, because no point is in all the sets C_k, $k = 1, 2, 3, \ldots$

1.2.11 Because $f(x) = 0$ when $1 \le x < 10$,

$$Q(C_3) = \int_0^{10} f(x)\, dx = \int_0^1 6x(1 - x)\, dx = 1.$$

1.2.13 Part (c): Draw the region C carefully, noting that $x < 2/3$ because $3x/2 < 1$. Thus

$$Q(C) = \int_0^{2/3} \left[\int_{x/2}^{3x/2} dy \right] dx = \int_0^{2/3} x\, dx = 2/9.$$

1.2.16 Note that

$$25 = Q(\mathcal{C}) = Q(C_1) + Q(C_2) - Q(C_1 \cap C_2) = 19 + 16 - Q(C_1 \cap C_2).$$

Hence, $Q(C_1 \cap C_2) = 10$.

1.2.17 By studying a Venn diagram with 3 intersecting sets, it should be true that

$$11 \ge 8 + 6 + 5 - 3 - 2 - 1 = 13.$$

It is not, and the accuracy of the report should be questioned.

1.3.3
$$P(\mathcal{C}) = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdots = \frac{1/2}{1-(1/2)} = 1.$$

1.3.6
$$P(\mathcal{C}) = \int_{-\infty}^{\infty} e^{-|x|}\,dx = \int_{-\infty}^{0} e^{x}\,dx + \int_{0}^{\infty} e^{-x}\,dx = 2 \neq 1.$$
We must multiply by 1/2.

1.3.8
$$P(C_1^c \cup C_2^c) = P[(C_1 \cap C_2)^c] = P(\mathcal{C}) = 1,$$
because $C_1 \cap C_2 = \phi$ and $\phi^c = \mathcal{C}$.

1.3.11 The probability that he does not win a prize is
$$\binom{990}{5} \Big/ \binom{1000}{5}.$$

1.3.13 Part (a): We must have 3 even or one even, 2 odd to have an even sum. Hence, the answer is
$$\frac{\binom{10}{3}\binom{10}{0}}{\binom{20}{3}} + \frac{\binom{10}{1}\binom{10}{2}}{\binom{20}{3}}.$$

1.3.14 There are 5 mutual exclusive ways this can happen: two "ones", two "twos", two "threes", two "reds", two "blues." The sum of the corresponding probabilities is
$$\frac{\binom{2}{2}\binom{6}{0} + \binom{2}{2}\binom{6}{0} + \binom{2}{2}\binom{6}{0} + \binom{5}{2}\binom{3}{0} + \binom{3}{2}\binom{5}{0}}{\binom{8}{2}}.$$

1.3.15

(a) $$1 - \frac{\binom{48}{5}\binom{2}{0}}{\binom{50}{5}}$$

(b) $$1 - \frac{\binom{48}{n}\binom{2}{0}}{\binom{50}{n}} \geq \frac{1}{2}, \text{ Solve for n.}$$

1.3.20 Choose an integer $n_0 > \max\{a^{-1}, (1-a)^{-1}\}$. Then $\{a\} = \cap_{n=n_0}^{\infty} \left(a - \frac{1}{n}, a + \frac{1}{n}\right)$. Hence by (1.3.10),
$$P(\{a\}) = \lim_{n\to\infty} P\left[\left(a - \frac{1}{n}, a + \frac{1}{n}\right)\right] = \frac{2}{n} = 0.$$

1.4.2
$$P[(C_1 \cap C_2 \cap C_3) \cap C_4] = P[C_4|C_1 \cap C_2 \cap C_3]P(C_1 \cap C_2 \cap C_3),$$
and so forth. That is, write the last factor as
$$P[(C_1 \cap C_2) \cap C_3] = P[C_3|C_1 \cap C_2]P(C_1 \cap C_2).$$

1.4.5

$$\frac{\left[\binom{4}{3}\binom{48}{10}+\binom{4}{4}\binom{48}{9}\right]/\binom{52}{13}}{\left[\binom{4}{2}\binom{48}{11}+\binom{4}{3}\binom{48}{10}+\binom{4}{4}\binom{48}{9}\right]/\binom{52}{13}}.$$

1.4.10

$$P(C_1|C)=\frac{(2/3)(3/10)}{(2/3)(3/10)+(1/3)(8/10)}=\frac{3}{7}<\frac{2}{3}=P(C_1).$$

1.4.12 Part (c):

$$\begin{aligned} P(C_1\cup C_2^c) &= 1-P[(C_1\cup C_2^c)^c]=1-P(C_1^*\cap C_2)\\ &= 1-(0.4)(0.3)=0.88. \end{aligned}$$

1.4.14 Part (d):

$$1-(0.3)^2(0.1)(0.6).$$

1.4.16 $1-P(TT)=1-(1/2)(1/2)=3/4$, assuming independence and that H and T are equilikely.

1.4.19 Let C be the complement of the event; i.e., C equals at most 3 draws to get the first spade.

(a) $P(C)=\frac{1}{4}+\frac{3}{4}\frac{1}{4}+\left(\frac{3}{4}\right)^2\frac{1}{4}$.

(b) $P(C)=\frac{1}{4}+\frac{13}{51}\frac{39}{52}+\frac{13}{50}\frac{38}{51}\frac{39}{52}$.

1.4.22 The probability that A wins is $\sum_{n=0}^{\infty}\left(\frac{5}{6}\frac{4}{6}\right)^n\frac{1}{6}=\frac{3}{8}$.

1.4.27 Let Y denote the bulb is yellow and let T_1 and T_2 denote bags of the first and second types, respectively.

(a)

$$P(Y)=P(Y|T_1)P(T_1)+P(Y|T_2)P(T_2)=\frac{20}{25}.6+\frac{10}{25}.4.$$

(b)

$$P(T_1|Y)=\frac{P(Y|T_1)P(T_1)}{P(Y)}.$$

1.4.30 Suppose without loss of generality that the prize is behind curtain 1. Condition on the event that the contestant switches. If the contestant chooses curtain 2 then she wins, (In this case Monte cannot choose curtain 1, so he must choose curtain 3 and, hence, the contestant switches to curtain 1). Likewise, in the case the contestant chooses curtain 3. If the contestant chooses curtain 1, she loses. Therefore the conditional probability that she wins is $\frac{2}{3}$.

1.4.31 (1) The probability is $1-\left(\frac{5}{6}\right)^4$.

(2) The probability is $1-\left[\left(\frac{5}{6}\right)^2+\frac{10}{36}\right]^{24}$.

1.5.2 Part (a):

$$c[(2/3)+(2/3)^2+(2/3)^3+\cdots] = \frac{c(2/3)}{1-(2/3)} = 2c = 1,$$

so $c = 1/2$.

1.5.5 Part (a):

$$p(x) = \begin{cases} \frac{\binom{13}{x}\binom{39}{5-x}}{\binom{52}{5}} & x = 0, 1, \ldots, 5 \\ 0 & \text{elsewhere.} \end{cases}$$

1.5.9 Part (b):

$$\sum_{x=1}^{50} x/5050 = \frac{50(51)}{2(5050)} = \frac{51}{202}.$$

1.5.10 For Part (c): Let $C_n = \{X \le n\}$. Then $C_n \subset C_{n+1}$ and $\cup_n C_n = R$. Hence, $\lim_{n\to\infty} F(n) = 1$. Let $\epsilon > 0$ be given. Choose n_0 such that $n \ge n_0$ implies $1 - F(n) < \epsilon$. Then if $x \ge n_0$, $1 - F(x) \le 1 - F(n_0) < \epsilon$.

1.6.2 Part (a):

$$p(x) = \frac{\binom{9}{x-1}}{\binom{10}{x-1}} \frac{1}{11-x} = \frac{1}{10}, \quad x = 1, 2, \ldots 10.$$

1.6.3

$$\text{(a)} \quad p(x) = \left(\frac{5}{6}\right)^{x-1}\left(\frac{1}{6}\right), \quad x = 1, 2, 3, \ldots$$

$$\text{(b)} \quad \sum_{x=1}^{\infty}\left(\frac{5}{6}\right)^{x-1}\left(\frac{1}{6}\right) = \frac{1/6}{1-(25/36)} = \frac{6}{11}.$$

1.6.8 $\mathcal{D}_y = \{1, 2^3, 3^3, \ldots\}$. The pmf of Y is

$$p(y) = \left(\frac{1}{2}\right)^{y^{1/3}}, \quad y \in \mathcal{D}_y.$$

1.7.1 If $\sqrt{x} < 10$ then

$$F(x) = P[X(c) = c^2 \le x] = P(c \le \sqrt{x}) = \int_0^{\sqrt{x}} \frac{1}{10}\, dz = \frac{\sqrt{x}}{10}.$$

Thus

$$f(x) = F'(x) = \begin{cases} \frac{1}{20\sqrt{x}} & 0 < x < 100 \\ 0 & \text{elsewhere.} \end{cases}$$

1.7.2

$$C_2 \subset C_1^c \Rightarrow P(C_2) \le P(C_1^c) = 1 - (3/8) = 5/8.$$

1.7.4 Among other characteristics,

$$\int_{-\infty}^{\infty} \frac{1}{\pi(1+x^2)}\, dx = \frac{1}{\pi}\text{arctan}x\bigg|_{-\infty}^{\infty} = \frac{1}{\pi}\left[\frac{\pi}{2} - \left(-\frac{\pi}{2}\right)\right] = 1.$$

1.7.6 Part (b):

$$\begin{aligned} P(X^2 < 9) &= P(-3 < X < 3) = \int_{-2}^{3} \frac{x+2}{19}\, dx \\ &= \frac{1}{18}\left[\frac{x^2}{2} + 2x\right]_{-2}^{3} = \frac{1}{18}\left[\frac{21}{2} - (-2)\right] = \frac{25}{36}. \end{aligned}$$

1.7.8 Part (c):

$$f'(x) = \frac{1}{2} 2xe^{-x} = 0;$$

hence, $x = 2$ is the mode because it maximizes $f(x)$.

1.7.9 Part (b):

$$\int_0^m 3x^2\, dx = \frac{1}{2};$$

hence, $m^3 = 2^{-1}$ and $m = (1/2)^{1/3}$.

1.7.10

$$\int_0^{\xi_{0.2}} 4x^3\, dx = 0.2 :$$

hence, $\xi_{0.2}^4 = 0.2$ and $\xi_{0.2} = 0.2^{1/4}$.

1.7.13 $x = 1$ is the mode because for $0 < x < \infty$ because

$$\begin{aligned} f(x) &= F'(x) = e^{-x} - e^{-x} + xe^{-x} = xe^{-x} \\ f'(x) &= -xe^{-x} + e^{-x} = 0, \end{aligned}$$

and $f'(1) = 0$.

1.7.16 Since $\Delta > 0$

$$X > z \Rightarrow Y = X + \Delta > z.$$

Hence, $P(X > z) \leq P(Y > z)$.

1.7.19 Since $f(x)$ is symmetric about 0, $\xi_{.25} < 0$. So we need to solve,

$$\int_{-2}^{\xi_{.25}} \left(-\frac{x}{4}\right)\, dx = .25.$$

The solution is $\xi_{.25} = -\sqrt{2}$.

1.7.20 For $0 < y < 27$,

$$
\begin{aligned}
x &= y^{1/3}, \quad \frac{dx}{dy} = \frac{1}{3}y^{-2/3} \\
g(y) = &= \frac{1}{3y^{2/3}}\frac{y^{2/3}}{9} = \frac{1}{27}.
\end{aligned}
$$

1.7.22

$$
\begin{aligned}
f(x) &= \frac{1}{\pi}, \quad \frac{-\pi}{2} < x < \frac{\pi}{2}. \\
x &= \arctan y, \quad \frac{dx}{dy} = \frac{1}{1+y^2}, \quad -\infty < y < \infty. \\
g(y) &= \frac{1}{\pi}\frac{1}{1+y^2}, \quad -\infty < y < \infty.
\end{aligned}
$$

1.7.23

$$
\begin{aligned}
G(y) &= P(-2\log X^4 \le y) = P(X \ge e^{-y/8}) = \int_{e^{-y/8}}^{1} 4x^3\,dx = 1 - e^{-y/2}, \quad 0 < y < \infty \\
g(y) &= G'(y) = \begin{cases} e^{-y/2} & 0 < y < \infty \\ 0 & \text{elsewhere.} \end{cases}
\end{aligned}
$$

1.7.24

$$
\begin{aligned}
G(y) &= P(X^2 \le y) = P(-\sqrt{y} \le X \le \sqrt{y}) \\
&= \begin{cases} \int_{-\sqrt{y}}^{\sqrt{y}} \frac{1}{3}\,dx = \frac{2\sqrt{y}}{3} & 0 \le y < 1 \\ \int_{-1}^{\sqrt{y}} \frac{1}{3}\,dx = \frac{\sqrt{y}}{3} + \frac{1}{3} & 1 \le y < 4 \end{cases} \\
g(y) &= \begin{cases} \frac{1}{3\sqrt{y}} & 0 \le y < 1 \\ \frac{1}{6\sqrt{y}} & 1 \le y < 4 \\ 0 & \text{elsewhere.} \end{cases}
\end{aligned}
$$

1.8.4

$$E(1/X) = \sum_{x=51}^{100} \frac{1}{x}\frac{1}{50}.$$

The latter sum is bounded by the two integrals

$$\int_{51}^{101} \tfrac{1}{x}\,dx \text{ and } \int_{50}^{100} \tfrac{1}{x}\,dx.$$

An appropriate approximation might be

$$\frac{1}{50}\int_{50.5}^{101.5} \frac{1}{x}\,dx = \frac{1}{50}(\log 100.5 - \log 50.5).$$

1.8.6
$$E[X(1-X)] = \int_0^1 x(1-x)3x^2\,dx.$$

1.8.8 When $1 < y < \infty$

$$\begin{aligned}
G(y) &= P(1/X \le y) = P(X \ge 1/y) = \int_{1/y}^1 2x\,dx = 1 - \frac{1}{y^2} \\
g(y) &= \frac{2}{y^3} \\
E(Y) &= \int_1^\infty y\frac{2}{y^3}\,dy = 2, \text{ which equals } \textstyle\int_0^1 (1/x)2x\,dx.
\end{aligned}$$

1.8.10 The expectation of X does not exist because

$$E(|X|) = \frac{2}{\pi}\int_0^\infty \frac{x}{1+x^2}\,dx = \frac{1}{\pi}\int_1^\infty \frac{1}{u}\,du = \infty,$$

where the change of variable $u = 1 + x^2$ was used.

1.9.2
$$M(t) = \sum_{x=1}^\infty \left(\frac{e^t}{2}\right)^x = \frac{e^t/2}{1-(e^t/2)}, \quad e^t/2 < 1.$$

Find $E(X) = M'(0)$ and $\mathrm{Var}(X) = M''(0) - [M'(0)]^2$.

1.9.4
$$0 \le \mathrm{var}(X) = E(X^2) - [E(X)]^2.$$

1.9.6
$$E\left[\left(\frac{X-\mu}{\sigma}\right)^2\right] = \frac{1}{\sigma^2}\sigma^2 = 1.$$

1.9.8

$$\begin{aligned}
K(b) &= E[(X-b)^2] = E(X^2) - 2bE(X) + b^2 \\
K'(b) &= -2E(X) + 2b = 0 \Rightarrow b = E(X).
\end{aligned}$$

1.9.11 For a continuous type random variable,

$$\begin{aligned}
K(t) &= \int_{-\infty}^\infty t^x f(x)\,dx. \\
K'(t) &= \int_{-\infty}^\infty xt^{x-1} f(x)\,dx \Rightarrow K'(1) = E(X). \\
K''(t) &= \int_{-\infty}^\infty x(x-1)t^{x-2} f(x)\,dx \Rightarrow K''(1) = E[X(X_1)];
\end{aligned}$$

and so forth.

1.9.12

$$\begin{aligned} 3 &= E(X-7) \Rightarrow E(X) = 10 = \mu. \\ 11 &= E[(X-7)^2] = E(X^2) - 14E(X) + 49 = E(X^2) - 91 \\ &\quad \Rightarrow E(X^2) = 102 \text{ and } \mathrm{var}(X) = 102 - 100 = 2. \\ 15 &= E[(X-7)^3]. \text{ Expand } (X-7)^3 \text{ and continue.} \end{aligned}$$

1.9.16

$$\begin{aligned} E(X) &= 0 \Rightarrow \mathrm{var}(X) = E(X^2) = 2p. \\ E(X^4) &= 2p \Rightarrow \text{kurtosis} = 2p/4p^2 = 1/2p. \end{aligned}$$

1.9.17

$$\begin{aligned} \psi'(t) &= M'(t)/M(t) \Rightarrow \psi'(0) = M'(0)/M(0) = E(X). \\ \psi''(t) &= \frac{M(t)M''(t) - M'(t)M'(t)}{[M(t)^2]} \\ &\quad \Rightarrow \psi''(0) = \frac{M(0)M''(0) - M'(0)M'(0)}{[M(0)^2]} = M''(0) - [M'(0)]^2 = \mathrm{var}(X). \end{aligned}$$

1.9.19

$$M(t) = (1-t)^{-3} = 1 + 3t + 3\cdot 4\frac{t^2}{2!} + 3\cdot 4\cdot 5\frac{t^3}{3!} + \cdots$$

Considering the coefficient of $t^r/r!$, we have

$$E(X^r) = 3\cdot 4\cdot 5\cdots(r+2), \quad r = 1, 2, 3 \ldots.$$

1.9.20 Integrating the parts with $u = 1 - F(x)$, $dv = dx$, we get

$$\{[1-F(x)]x\}_0^b - \int_0^b x[-f(x)]\,dx = \int_0^b xf(x)\,dx = E(X).$$

1.9.23

$$\begin{aligned} E(X) &= \int_0^1 x\frac{1}{4}\,dx + 0\cdot\frac{1}{4} + 1\cdot\frac{1}{2} = \frac{5}{8}. \\ E(X^2) &= \int_0^1 x^2\frac{1}{4}\,dx + 0\cdot\frac{1}{4} + 1\cdot\frac{1}{2} = \frac{7}{12}. \\ \mathrm{var}(X) &= \frac{7}{12} - \left(\frac{5}{8}\right)^2 = \frac{37}{192}. \end{aligned}$$

1.9.24

$$E(X) = \int_{-\infty}^{\infty} x[c_1 f_1(x) + \cdots + c_k f_k(x)]\,dx = \sum_{i=1}^{k} c_i\mu_i = \mu.$$

Because $\int_{-\infty}^{\infty}(x-\mu)^2 f_i(x)\,dx = \sigma_i^2 + (\mu_i-\mu)^2$, we have

$$E[(X-\mu)^2] = \sum_{i=1}^{k} c_i[\sigma_i^2 + (\mu_i-\mu)^2].$$

1.10.2

$$\mu = \int_0^\infty x f(x)\,dx \geq \int_{2\mu}^\infty 2\mu f(x)\,dx = 2\mu P(X > 2\mu).$$

Thus $\frac{1}{2} \geq P(X > 2\mu)$.

1.10.4 If, in Theorem 1.10.2, we take $u(X) = \exp\{tX\}$ and $c = \exp\{ta\}$, we have

$$P(\exp\{tX\} \geq \exp\{ta\}] \leq M(t)\exp\{-ta\}.$$

If $t > 0$, the events $\exp\{tX\} \geq \exp\{ta\}$ and $X \geq a$ are equivalent. If $t < 0$, the events $\exp\{tX\} \geq \exp\{ta\}$ and $X \leq a$ are equivalent.

1.10.5 We have $P(X \geq 1) \leq [1-\exp\{-2t\}]/2t$ for all $0 < t < \infty$, and $P(X \leq -1) \leq [\exp\{2t\}-1]/2t$ for all $-\infty < t < 0$. Each of these bounds has the limit 0 as $t \to \infty$ and $t \to -\infty$, respectively.

Chapter 2

Multivariate Distributions

2.1.2

$$P(A_5) = \frac{7}{8} - \frac{4}{8} - \frac{3}{8} + \frac{2}{8} = \frac{2}{8}.$$

2.1.5

$$\begin{aligned}
\int_0^\infty \int_0^\infty \left[2g(\sqrt{x_1^2 + x_2^2})/\pi\sqrt{x_1^2 + x_2^2} \right] dx_1 dx_2 &= \int_0^\infty \int_0^{\pi/2} [2g(\rho)/\pi\rho]\rho \, d\theta d\rho \\
&= \int_0^\infty g(\rho) \, d\rho = 1.
\end{aligned}$$

2.1.6

$$\begin{aligned}
G(z) &= P(X + Y \le z) = \int_0^z \int_0^{z-x} e^{-x-y} \, dy dx \\
&= \int_0^z [1 - e^{-(z-x)}] e^{-x} \, dx = 1 - e^{-z} - ze^{-z}. \\
g(z) &= G'(z) = \begin{cases} ze^{-z} & 0 < z < \infty \\ 0 & \text{elsewhere.} \end{cases}
\end{aligned}$$

2.1.7

$$\begin{aligned}
G(z) &= P(XY \le z) = 1 - \int_z^1 \int_{z/x}^1 dy dx \\
&= 1 - \int_z^1 \left(1 - \frac{z}{x}\right) dx = z - z \log z \\
g(z) &= G'(z) = \begin{cases} -\log z & 0 < z < 1 \\ 0 & \text{elsewhere.} \end{cases}
\end{aligned}$$

Why is $-\log z > 0$?

2.1.8

$$f(x,y) = \begin{cases} \frac{\binom{13}{x}\binom{13}{y}\binom{26}{13-x-y}}{\binom{52}{13}} & x \geq 0, y \geq 0, x+y \leq 13, x \text{ and } y \text{ integers} \\ 0 & \text{elsewhere.} \end{cases}$$

2.1.10

$$P(X_1 + X_2 \leq 1) = 15 \int_0^{1/2} x_1^2 \left[\int_{x_1}^{1-x_1} x_2 \, dx_2 \right] dx_1.$$

2.1.14

$$\begin{aligned} E\left[e^{t_1 X_1 + t_2 X_2}\right] &= \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} e^{t_1 i + t_2 j} \left(\frac{1}{2}\right)^{i+j} \\ &= \sum_{i=1}^{\infty} \left(e^{t_1} \frac{1}{2}\right)^i \sum_{j=1}^{\infty} \left(e^{t_1} \frac{1}{2}\right)^j \\ &= \left[\frac{1}{1 - 2^{-1} e^{t_1}} - 1\right] \left[\frac{1}{1 - 2^{-1} e^{t_2}} - 1\right], \end{aligned}$$

provided $t_i < \log 2$, $i = 1, 2$.

2.2.1

$$p(y_1, y_2) = \begin{cases} \left(\frac{2}{3}\right)^{y_2} \left(\frac{1}{3}\right)^{2-y_2} & (y_1, y_2) = (0,0), (-1,1), (1,1), (0,2) \\ 0 & \text{elsewhere.} \end{cases}$$

2.2.2

$$p(y_1, y_2) = \begin{cases} y_1/36 & y_1 = y_2, 2y_2, 3y_2; y_2 = 1,2,3 \\ 0 & \text{elsewhere.} \end{cases}$$

y_1	1	2	3	4	6	9
$p(y_1)$	1/36	4/36	6/36	4/36	12/36	9/36

2.2.4 The inverse transformation is given by $x_1 = y_1 y_2$ and $x_2 = y_2$ with Jacobian $J = y_2$. By noting what the boundaries of the space $\mathcal{S}(X_1, X_2)$ map into, it follows that the space $\mathcal{T}(Y_1, Y_2) = \{(y_1, y_2) : 0 < y_i < 1, i = 1, 2\}$. The pdf of (Y_1, Y_2) is $f_{Y_1,Y_2}(y_1, y_2) = 8 y_1 y_2^3$.

2.2.5 The inverse transformation is $x_1 = y_1 - y_2$ and $x_2 = y_2$ with Jacobian $J = 1$. The space of (Y_1, Y_2) is $\mathcal{T} = \{(y_1, y_2) : -\infty < y_i < \infty, i = 1, 2\}$. Thus the joint pdf of (Y_1, Y_2) is

$$f_{Y_1,Y_2}(y_1, y_2) = f_{X_1,X_2}(y_1 - y_2, y_2),$$

which leads to formula (2.2.1).

2.3.2

$$
\begin{array}{ll}
(a) & c_1 \int_0^{x_2} x_1/x_2^2\,dx_1 = \dfrac{c_1}{2} = 1 \Rightarrow c_1 = 2 \text{ and } c_2 = 5. \\
(b) & 10x_1x_2^2, 0 < x_1 < x_2 < 1; \text{zero elsewhere} \\
(c) & \displaystyle\int_{1/4}^{1/2} 2x_1/(5/8)^2\,dx = \frac{64}{25}\left(\frac{1}{4} - \frac{1}{16}\right) = \frac{12}{25}. \\
(d) & \displaystyle\int_{1/4}^{1/2}\int_{x_1}^{1} 10x_1x_2^2\,dx_2dx_1 = \int_{1/4}^{1/2} \frac{10}{3}x_1(1-x_1^3)\,dx_1 = \frac{135}{512}.
\end{array}
$$

2.3.3

$$
\begin{aligned}
f_2(x_2) &= \int_0^{x_2} 21x_1^2x_2^3\,dx_1 = 7x_2^6, \quad 0 < x_2 < 1. \\
f_{1|2}(x_1|x_2) &= 21x_1^2x_2^3/7x_2^6 = 3x_1^2/x_2^3, \quad 0 < x_1 < x_2. \\
E(X_1|x_2) &= \int_0^{x_2} x_1(3x_1^2/x_2^3)\,dx_1 = \frac{3}{4}x_2. \\
G(y) &= P\left(\frac{3}{4}X_2 \le y\right) = \int_0^{4y/3} 7x_2^6\,dx_2 = \left(\frac{4y}{3}\right)^7, \quad 0 < y < \frac{3}{4} \\
g(y) &= \begin{cases} 7\left(\frac{4}{3}\right)^7 y^6 & 0 < y < \frac{3}{4} \\ 0 & \text{elsewhere.} \end{cases} \\
E(Y) &= \frac{7}{8}\frac{3}{4} = \frac{21}{32}. \\
\text{Var}(Y) &= \frac{7}{1024}. \\
E(X_1) &= \frac{21}{32}. \\
\text{Var}(X_1) &= \frac{553}{15360} > \frac{7}{1024}.
\end{aligned}
$$

2.3.8 The marginal pdf of X is

$$f_X(x) = 2\int_x^\infty e^{-x}e^{-y}\,dy = 2e^{-2x}, \quad 0 < x < \infty.$$

Hence, the conditional pdf of Y given $X = x$ is

$$f_{Y|X}(y|x) = \frac{2e^{-x}e^{-y}}{2e^{-2x}} = e^{-(y-x)}, \quad 0 < x < y < \infty,$$

with conditional mean

$$E(Y|X = x) = \int_x^\infty ye^{-(y-x)}\,dy = x + 1, \quad x > 0.$$

2.3.9 For Part (c):

$$\binom{13}{x_2}\binom{13}{x_3}\binom{13}{2-x_2-x_3}\Big/\binom{39}{2}, \quad \text{where integers } x_2, x_3 \geq 0 \text{ and } x_2 + x_3 \leq 2.$$

2.3.11

(a) $$f_1(x_1)f_{2|1}(x_2|x_1) = 1 \cdot \frac{1}{x_1}, \quad 0 < x_2 < x_1 < 1.$$

(b) $$\int_{1/2}^{1}\int_{1-x_1}^{x_1} \frac{1}{x_1}\, dx_2 dx_1 = \int_{1/2}^{1} \frac{2x_1 - 1}{x_1}\, dx_1 = 2(1/2) + \log(1/2) = 1 - \log 2.$$

2.3.12

(b) $$\int_2^\infty e^{-x}\, dx \Big/ \int_1^\infty e^{-x}\, dx = e^{-2}/e^{-1} = e^{-1}.$$

2.4.1 For Part (c):

$$\text{cov} = (0)(0)(1/3) + (1)(1)(1/3) + (2)(0)(1/3) - (1)(1/3) = 0.$$

Thus $\rho = 0$ and yet X and Y are dependent.

2.4.3

$$\rho^2 = (1/2)(1/2) = 1/4 \Rightarrow \rho = 1/2.$$

2.4.7

$$h(v) = \text{var}(X) + 2v\text{cov}(X,Y) + v^2\text{var}(Y) \geq 0,$$

for all v. Hence, the discriminant of this quadratic must satisfy $b^2 - 4ac \leq 0$ which yields

$$[2\text{cov}(X,Y)]^2 - 4\text{var}(X)\text{var}(Y) \leq 0.$$

Equivalently,

$$\rho^2 = [\text{cov}(X,Y)]^2/\text{var}(X)\text{var}(Y) \leq 1.$$

2.4.11 Let $Y = (X_1 - \mu_1) + (X_2 - \mu_2)$. Then the mean of Y is 0 and its variance is

$$\text{Var}(Y) = \text{Var}(X_1 + X_2) = \sigma^2 + \sigma^2 + 2\rho\sigma^2 = 2\sigma^2(1-\rho).$$

Use Chebyshev's inequality to obtain the result.

2.5.2 X_1 and X_2 are dependent because $0 < x_1 < x_2 < \infty$ is not a product space.

2.5.4 Because X_1 and X_2 are independent, the probability equals

$$\left[\int_0^{1/3} 2x_1\, dx_1\right]\left[\int_0^{1/3} 2(1-x_2)\, dx_2\right] = (1/3)^2[1 - (2/3)^2] = 5/81.$$

2.5.7 The marginal pdf of X_1 is given by

$$f_{X_1}(x_1) = \int_{-2-\sqrt{1-(x_1-1)^2}}^{-2+\sqrt{1-(x_1-1)^2}} \frac{1}{\pi}\,dx_2 = \frac{2}{\pi}\sqrt{1-(x_1-1)^2}, \quad 0 < x < 2.$$

The random variables X_1 and X_2 are not independent.

2.5.8 X and Y are dependent because $0 < y < x < 1$ is not a product space.

$$E(X|y) = \int_y^1 x[2x/(1-y^2)]\,dx = \frac{2(1-y^2)}{3(1-y^2)}.$$

2.5.9

$$\begin{aligned} P(X+Y \le 60) &= P(X \le 10) + \int_{10}^{20}\int_{40}^{60-x} \frac{1}{300}\,dy\,dx \\ &= \frac{1}{3} + \int_{10}^{20} (20-x)/300\,dx = \frac{1}{3} + \frac{1}{6} = \frac{1}{2}. \end{aligned}$$

2.5.12

$$\begin{aligned} P(|X_1 - X_2| = 1) &= P(X_1 = 0, X_2 = 1) + P(X_1 = 1, X_2 = 0) \\ &= P(X_1 = 0)P(X_2 = 1) + P(X_1 = 1)P(X_2 = 0) = \frac{1}{3}. \end{aligned}$$

2.6.1 For Part (g):

$$E(X|y,z) = \int_0^1 x\frac{3(x+y+z)/2}{3((1/2)+y+z)/2}\,dx = \frac{(1/3)+(y/2)+(z/2)}{(1/2)+y+z}.$$

2.6.3

$$\begin{aligned} G(y) &= 1 - P(y < X_i, i = 1,2,3,4) = 1 - [(1-y)^3]^4 = 1-(1-y)^{12} \\ g(y) &= G'(y) = \begin{cases} 12(1-y)^{11} & 0 < y < 1 \\ 0 & \text{elsewhere.} \end{cases} \end{aligned}$$

2.6.6 Multiply both members of $E[X_1 - \mu_1|x_2, x_3] = b_2(x_2 - \mu_2) + b_3(x_3 - \mu_3)$ by the joint pdf of X_2 and X_3 and denote the result by (1). Multiply both members of (1) by $(x_2 - \mu_2)$ and integrate (or sum) on x_2 and x_3. This gives (2), $\rho_{12}\sigma_1\sigma_2 = b_2\sigma_2^2 + 3\rho_{23}\sigma_1\sigma_2$. Return to (1) and multiply each member by $(x_3 - \mu_3)$ and integrate (or sum) on x_2 and x_3. This yields (3) $\rho_{13}\sigma_1\sigma_3 = b_2\rho_{23}\sigma_2\sigma_3 + b_3\sigma_3^2$. Solve (2) and (3) for b_2 and b_3.

2.6.9

$$\begin{aligned} (a) \quad & \int_0^\infty\int_{x_1}^\infty e^{-x_1-x_2}\,dx_2\,dx_1 \Big/ \int_0^\infty\int_{x_1/2}^\infty e^{-x_1-x_2}\,dx_2\,dx_1 \\ & + \int_0^\infty e^{-2x_1}\,dx_1 \Big/ \int_0^\infty e^{-3x_1/2}\,dx_1 = \frac{1}{2}\frac{2}{3} = \frac{3}{4}. \end{aligned}$$

2.7.1

$$x_1 = y_1y_2y_3,\ x_2 = y_2y_3 - y_1y_2y_3,\ x_3 = y_3 - y_2y_3.$$

with $J = y_2y_3^2$, and $0 < y_1 < 1, 0 < y_2 < 1, 0 < y_3 < \infty$. This yields

$$g(y_1, y_2, y_3) = y_2y_3^2e^{-y_3} = (1)(2y_2)(y_3^2e^{-y_3}/2) = g_1(y_1)g_2(y_2)g_3(y_3).$$

2.7.2

$$x_1 = \sqrt{y},\ x_2 = -\sqrt{y} \text{ and } J_i = \tfrac{1}{2\sqrt{y}},\ i = 1.2.$$

This yields

$$g(y) = \frac{1}{2}\left(\frac{1}{2\sqrt{y}}\right) + \frac{1}{2}\left(\frac{1}{2\sqrt{y}}\right) = \frac{1}{2\sqrt{y}}, \quad 0 < y < 1.$$

2.7.5 The inverse transformation is $x_1 = \frac{y_1y_3}{1+y_1}$, $x_2 = \frac{y_3}{1+y_1}$, and $x_3 = y_2y_3$, with space $y_i > 0$, $i = 1, 2, 3$. The Jacobian is

$$J = \begin{vmatrix} \frac{y_3}{(1+y_1)^2} & 0 & \frac{y_1}{1+y_1} \\ \frac{-y_3}{(1+y_1)^2} & 0 & \frac{1}{1+y_1} \\ 0 & y_3 & y_2 \end{vmatrix} = \left[\frac{y_3^2}{(1+y_1)^3} + \frac{y_1y_3^2}{(1+y_1)^3}\right] = \frac{y_3^2}{(1+y_1)^2}.$$

2.7.8 Expanding $M(t)$ we get

$$M(t) = \left(\frac{3}{4}\right)^2 e^0 + 2\left(\frac{3}{4}\right)\left(\frac{1}{4}\right)e^t + \left(\frac{1}{4}\right)^2 e^{2t}.$$

From this, we immediately get the probabilities

$$P(X = 0) = \left(\tfrac{3}{4}\right)^2,\ P(X = 1) = 2\left(\tfrac{3}{4}\right)\left(\tfrac{1}{4}\right) \text{ and } P(X = 2) = \left(\tfrac{1}{4}\right)^2.$$

2.8.2 Note that

$$\begin{aligned} \mu_1 &= E(X_i) = \int_0^1 2x^2\,dx = \frac{2}{3}x^3\Big|_0^1 = \frac{2}{3} \\ E(X_i^2) &= \int_0^1 2x^3\,dx = \frac{2}{4}x^4\Big|_0^1 = \frac{1}{2} \end{aligned}$$

So

$$\sigma^2 = \frac{1}{2} - \frac{4}{9} = \frac{1}{18}.$$

Hence,

$$\begin{aligned} E(Y) &= \sum_{i=1}^4 E(X_i) = \frac{8}{3} \\ V(Y) &= \sum_{i=1}^4 V(X_i) = \frac{4}{18}, \end{aligned}$$

where we used the independence of $X_1, \ldots, X_4$ to establish the variance of Y.

2.8.4 By independence

$$\begin{aligned} E(X_1X_2) &= E(X_1)E(X_2) = \mu_1\mu_2 \\ E(X_1^2X_2^2) &= E(X_1^2)E(X_2^2) = (\sigma_1^2+\mu_1^2)(\sigma_2^2+\mu_2^2). \end{aligned}$$

So,

$$V(X_1X_2) = (\sigma_1^2+\mu_1^2)(\sigma_2^2+\mu_2^2) - \mu_1^2\mu_2^2,$$

which simplifies to the answer.

2.8.8 Because in these cases, the correlation coefficient is never influenced by the means, let $\mu_1 = \mu_2 = 0$. Then

$$\begin{aligned} \operatorname{cov}(X,Z) &= E[X(X-Y)] = E(X^2) = \sigma_1^2 \\ \rho &= \sigma_1^2/\sqrt{\sigma_1^2(\sigma_1^2+\sigma_2^2)} = \sigma_1/\sqrt{\sigma_1^2+\sigma_2^2}. \end{aligned}$$

2.8.11

$$\begin{aligned} \operatorname{cov}(W,Z) &= E[(aX+b-a\mu_1-b)(cY+d-c\mu_2-d)] \\ &= acE[(X-\mu_1)(Y-\mu_2)] = ac\rho\sigma_1\sigma_2 \\ \text{correlation coef.} &= \frac{ac\rho\sigma_1\sigma_2}{\sqrt{a^2c^2\sigma_1^2\sigma_2^2}} = \rho. \end{aligned}$$

2.8.13

$$\begin{aligned} \operatorname{cov}(X_1X_2, X_1) &= E[(X_1X_2-\mu_1\mu_2)(X_1-\mu_1)] \\ &= (\mu_1^2+\sigma_1^2)\mu_2 - \mu_1^2\mu_2 - \mu_1^2\mu_2 + \mu_1^2\mu_2 = \sigma_1^2\mu_2. \end{aligned}$$

2.8.15 Without loss of generality, let the means equal zero

$$\begin{aligned} \operatorname{cov}(Y,Z) &= (0.3+0.5+1.0+0.2)\sigma^2 = 2\sigma^2, \\ \text{Answer} &= 2\sigma^2/\sqrt{[1+2(0.3)+1]\sigma^2[1+2(0.2)+1]\sigma^2} = \frac{2}{\sqrt{(2.6)(2.4)}} = 0.801. \end{aligned}$$

2.8.17 Again let $\mu_1 = \mu_2 = 0$.

$$\operatorname{cov}E\{X[Y-\rho(\sigma_2/\sigma_1)X]\} = \rho\sigma_1\sigma_2 - \rho(\sigma_2/\sigma_1)\sigma_1^2 = 0.$$

2.8.18 The function $g(x) = x^2$ is strictly convex. Hence, by Jensen's inequality,

$$(E(S))^2 < E(S^2),$$

which leads to $E(S) < \sigma$.

Chapter 3

Some Special Distributions

3.1.2 Since $n = 9$ and $p = 1/3$, $\mu = 3$ and $\sigma^2 = 2$. Hence, $\mu - 2\sigma = 3 - 2\sqrt{2}$ and $\mu + 2\sigma = 3 + 2\sqrt{2}$ and $P(\mu - 2\sigma < X < \mu + 2\sigma) = P(X = 1, 2, \ldots, 5)$.

3.1.3

$$\begin{aligned} E\left(\frac{X}{n}\right) &= \frac{1}{n}E(X) = \frac{1}{n}(np) = p \\ E\left[\left(\frac{X}{n} - p\right)^2\right] &= \frac{1}{n^2}E[(X - np)^2] = \frac{np(1-p)}{n^2} = \frac{p(1-p)}{n}. \end{aligned}$$

3.1.4 $p = P(X > 1/2) = \int_{1/2}^{1} 3x^2\, dx = \frac{7}{8}$ and $n = 3$. Thus $\binom{3}{2}\left(\frac{7}{8}\right)^2\left(\frac{1}{8}\right) = \frac{147}{512}$.

3.1.6 $P(Y \geq 1) = 1 - P(Y = 0) = 1 - (3/4)^n \geq 0.70$. That is, $0.30 \geq (3/4)^n$ which can be solved by taking logarithms.

3.1.9 Assume X and Y are independent with binomial distributions $b(2, 1/2)$ and $b(3, 1/2)$, respectively. Thus we want

$$\begin{aligned} P(X > Y) &= P(X = 1, 2 \text{ and } Y = 0) + P(X = 2 \text{ and } Y = 1) \\ &= \left[\binom{2}{1}\left(\frac{1}{2}\right)^2 + \binom{2}{2}\left(\frac{1}{2}\right)^2\right]\left[\left(\frac{1}{2}\right)^3\right] + \left[\left(\frac{1}{2}\right)^2\right] + \left[3\left(\frac{1}{2}\right)^3\right]. \end{aligned}$$

3.1.11

$$\begin{aligned} P(X \geq 1) &= 1 - (1-p)^2 = 5/9 \Rightarrow (1-p)^2 = 4/9 \\ P(Y \geq 1) &= 1 - (1-p)^4 = 1 - (4/9)^2 = 65/81. \end{aligned}$$

3.1.12 Let $f(x)$ denote the pmf which is $b(n, p)$. Show, for $x \geq 1$, that $f(x)/f(x-1) = 1 + [(n+1)p - x]/x(1-p)$. Then $f(x) > f(x-1)$ if $(n+1)p > x$ and $f(x) < f(x-1)$ if $(n+1)p < x$. Thus the mode is the greatest integer less than $(n+1)p$. If $(n+1)p$ is an integer, there is no unique mode but $f[(n+1)p] = f[(n+1)p - 1]$ is the maximum of $f(x)$.

3.1.14

$$\begin{aligned} P(X \geq 3) &= (1/3)(2/3)^3 + (1/3)(2/3)^4 + \cdots = \frac{(1/3)(2/3)^3}{1-(2/3)} = (2/3)^3. \\ p(x|X \geq 3) &= \frac{(1/3)(2/3)^x}{(2/3)^3} = (1/3)(2/3)^{x-3}, \quad x = 3, 4, 5, \ldots \end{aligned}$$

3.15

$$\frac{5!}{2!1!2!}\left(\frac{3}{6}\right)^2\left(\frac{2}{6}\right)\left(\frac{1}{6}\right)^2.$$

3.1.16 $M(t) = \sum_{y=0}^{\infty}\binom{y+r-1}{r-1}p^r[(1-p)e^t]^y = p^r[1-(1-p)e^t]^{-r}$, because the summation equals $p^r(1-w)^{-r}$, where $w = (1-p)e^t$.

3.1.18

$$\binom{5}{5}\left(\frac{1}{2}\right)^5 \Big/ \left[\binom{5}{4}\left(\frac{1}{2}\right)^5 + \binom{5}{5}\left(\frac{1}{2}\right)^5\right] = \frac{1}{6},$$

which is much different than 1/2 that some might have arrived at by letting 4 coins be heads and tossing the fifth coin.

3.1.19

$$\left[\frac{7!}{2!1!\cdots1!}\left(\frac{1}{6}\right)^7\right] \Big/ \left[\frac{7!}{2!5!}\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^5\right] = \frac{5!}{1!\cdots1!}\left(\frac{1}{5}\right)^5.$$

3.1.21

$$\begin{aligned} &(a) \qquad E(X_2) = \sum_{x_1=1}^{5}\sum_{x_2=0}^{x_1}\left[x_2\binom{x_1}{x_2}\left(\frac{1}{2}\right)^{x_1}\right]\frac{x_1}{15} = \sum_{x_1=1}^{5}\frac{x_1}{2}\frac{x_1}{15} = \frac{11}{6}. \\ &(b) \qquad f_{2|1}(x_2|x_1) \text{ is } b(x_1, 1/2) \Rightarrow E(X_2|x_1) = x_1/2. \\ &(c) \qquad E(x_1/2) = 11/6. \end{aligned}$$

3.1.22

$$p_1 = 6(1/6)^3 = 1/36,\ p_2 = 6 \cdot 5 \cdot 3 \cdot (1/6)^3 = 15/36.$$

Thus X and Y are trinomial ($n = 10, p_1 = 1/36, p_2 = 15/36$). $\mathrm{cov}(X, Y) = -np_1p_2$. Thus $E(XY) = -np_1p_2 + np_1np_2 = 25/24$.

3.1.25 Use the mgf technique and independence to get

$$\begin{aligned} E[e^{t(X_1-X_2+n_2)}] &= E[e^{tX_1}]E[e^{-tX_2}]e^{tn_2} \\ &= \left(\frac{1}{2}+\frac{1}{2}e^t\right)^{n_1}\left(\frac{1}{2}+\frac{1}{2}e^{-t}\right)^{n_2}e^{tn_2} \\ &= \left(\frac{1}{2}+\frac{1}{2}e^t\right)^{n_1+n_2}. \end{aligned}$$

3.1.27 Part (b): $D = 100$, so

$$\begin{aligned} P[X \geq 2] &= 1 - P[X \leq 1] \\ &= 1 - \frac{\binom{100}{0}\binom{900}{10}}{\binom{1000}{10}} - \frac{\binom{100}{1}\binom{900}{9}}{\binom{1000}{10}} = 0.2637. \end{aligned}$$

Part (c): For the binomial approximation for Part (b), $p = 0.10$ and $n = 10$; hence,

$$\begin{aligned} P[X \geq 2] &= 1 - P[X \leq 1] \\ &\approx 1 - 0.9^{10} - \binom{10}{1}.1^1.9^9 = 0.2639. \end{aligned}$$

3.2.1

$$\frac{e^{-\mu}\mu}{1!} = \frac{e^{-\mu}\mu^2}{2!} \Rightarrow \mu = 2 \text{ and } P(X = 4) = \tfrac{e^{-2}2^4}{4!}.$$

3.2.4 Given $p(x) = 4p(x-1)/x$, $x = 1, 2, 3, \ldots$. Thus $p(1) = 4p(0)$, $p(2) = 4^2p(0)/2!$, $p(3) = 4^3p(0)/3!$. Use induction to show that $p(x) = 4^xp(0)/x!$. Then

$$1 = \sum_{x=0}^{\infty} p(x) = p(0)\sum_{x=0}^{\infty} 4^x/x! = p(0)e^4 \text{ and } p(x) = 4^xe^{-4}/x!, x = 0, 1, 2, \ldots.$$

3.2.6 For $x = 1$, $D_w[g(1, w)] + \lambda g(1, w) = \lambda e^{-\lambda w}$. The general solution to $D_w[g(1, w)] + \lambda g(1, w) = 0$ is $g(1, w) = ce^{-\lambda w}$. A particular solution to the full differential equation is $\lambda we^{-\lambda w}$. Thus the most general solution is

$$g(1, w) = \lambda we^{-\lambda w} + ce^{-\lambda w}.$$

However, the boundary condition $g(1, 0)$requires that $c = 0$. Thus $g(1, w) = \lambda we^{-\lambda w}$. Now assume that the answer is correct for $x = -1$, and show that it is correct for x by exactly the same type of argument used for $x = 1$.

3.2.8

$$P(X \geq 2) = 1 - P(X = 0 \text{ or } X = 1) = 1 - [e^{-\mu} + e^{-\mu}\mu] \geq 0.99.$$

Thus $0.01 \geq (1+\mu)e^{-\mu}$. Solve by trying several values of μ using a calculator.

3.2.10

$$\sum_{x=0}^{k} \frac{e^{-3}3^x}{x!} \geq 0.99.$$

From tables, $k = 8$.

3.2.11 $\frac{e^{-\mu}\mu}{1!} = \frac{e^{-\mu}\mu^3}{3!}$ requires $\mu^2 = 6$ and $\mu = \sqrt{6}$. Since $\frac{e^{-\sqrt{6}}(\sqrt{6})^2}{2!} = 3e^{-\sqrt{6}} > \frac{e^{-\sqrt{6}}\sqrt{6}}{1!}$, $x = 2$ is the mode.

3.2.12

$$E(X!) = \sum_{x=0}^{\infty} x! \frac{e^{-1}}{x!} = \sum_{x=0}^{\infty} e^{-1} \text{ does not exist.}$$

3.2.13 For Part (a),

$$\begin{aligned} M(t_1, t_2) &= \sum_{y=0}^{\infty} \sum_{x=0}^{y} e^{t_1 x + t_2 y} \frac{e^{-2}}{x!(y-x)!} \\ &= \sum_{y=0}^{\infty} \left[\sum_{x=0}^{y} \frac{y!}{x!(y-x)!} \left(e^{t_1}\right)^x \right] \frac{e^{-2} e^{t_2 y}}{y!} \\ &= \sum_{y=0}^{\infty} \frac{e^{-2}[(1+e^{t_1})e^{t_2}]^y}{y!} \\ &= e^{-2} \exp[(1+e^{t_1})e^{t_2}]. \end{aligned}$$

3.3.1

$$(1-2t)^{-6} = (1-2t)^{-12/2} \Rightarrow X \text{ is } \chi^2(12).$$

From tables, $P(X < 5.23) = 0.05$.

3.3.4

$$\begin{aligned} M(t) &= 1 + \frac{2!2t}{1!} + \frac{3!2^2t^2}{3!} + \frac{4!2^3t^3}{3!} + \cdots \\ &= 1 + 2(2t) + 3(2t)^2 + 4(2t)^3 + \cdots \\ &= (1-2t)^{-2} = (1-2t)^{-4/2}, \end{aligned}$$

so X is $\chi^2(4)$.

3.3.6 Part (a):

$$\begin{aligned} Part(a) : P(Y \le y) &= 1 - [P(X > y)]^3 = 1 - (e^{-y})^3 = 1 - e^{-3y} = G(y). \\ g(y) &= G'(y) = 3e^{-3y}, \quad 0 < y < \infty. \end{aligned}$$

3.3.7 $f'(x) = \frac{1}{\beta^2} e^{-x/\beta} + \frac{1}{\beta^2} x e^{-x/\beta}(-1/\beta) = 0$; hence, $x = \beta$ which is given as 2. Thus X is $\chi^2(4)$.

3.3.9

$$P(X \ge 2\alpha\beta) \le e^{-2\alpha\beta t}(1-\beta t)^{-\alpha},$$

for all $t < 1/\beta$. The minimum of the right side, say $K(t)$, can be found by

$$K'(t) = e^{-2\alpha\beta t}(\alpha\beta)(1-\beta t)^{-\alpha-1} + e^{-2\alpha\beta t}(-2\alpha\beta)(1-\beta t)^{-\alpha} = 0$$

which implies that

$$(1-\beta t)^{-1} - 2 = 0 \text{ and } t = 1/2\beta.$$

That minimum is

$$K(1/2\beta) = e^{-\alpha}(1-(1/2))^{-\alpha} = (2/e)^{\alpha}.$$

3.3.10 If $r = 0$, $M(t) = 1 = e^{(0)t}$, which is the mgf of a degenerate distribution at $x = 0$.

3.3.14 The differential equation requires

$$\log g(0, w) = -kw^r + c.$$

The boundary condition $g(0, 0) = 1$ imlies that $c = 0$. Thus $g(0, w) = \exp\{-kw^r\}$ and $G(w) = 1 - \exp\{-kw^r\}$ and

$$G'(w) = krw^{r-1}\exp\{-kw^r\}, \quad 0 < w < \infty.$$

3.3.15 The joint pdf of X and the parameter is

$$\begin{aligned}
f(x|m)g(m) &= \frac{e^{-m}m^x}{x!}me^{-m}, \quad x = 0, 1, 2, \ldots, \; 0 < m < \infty \\
P(X = 0, 1, 2) &= \sum_{x=0}^{2}\int_0^\infty \frac{m^{x+1}e^{-2m}}{x!}\,dm = \sum_{x=0}^{2}\frac{\Gamma(x+2)(1/2)^{x+2}}{x!} \\
&= \sum_{x=0}^{2}(x+1)(1/2)^{x+2} = \frac{1}{4} + \frac{2}{8} + \frac{3}{16} = \frac{11}{16}.
\end{aligned}$$

3.3.16

$$\begin{aligned}
G(y) &= P(Y \le y) = P(-2\log X \le y) = P(X \ge \exp\{-y/2\}) \\
&= \int_{\exp\{-y/2\}}^{1}(1)\,dx = 1 - \exp\{-y/2\}, \quad 0 < y < \infty \\
g(y) &= G'(y) = (1/2)\exp\{-y/2\}, \quad 0 < y < \infty;
\end{aligned}$$

so Y is $\chi^2(2)$.

3.3.17 $f(x) = 1/(b-a)$, $a < x < b$, has respective mean and variance of

$$\tfrac{a+b}{2} = 8 \text{ and } \tfrac{(b-a)^2}{12} = 16.$$

Solve for a and b.

3.3.18

$$\begin{aligned}
E(X) &= \frac{\Gamma(\alpha+\beta)\Gamma(\alpha+1)\Gamma(\beta)}{\Gamma(\alpha)\Gamma(\beta)\Gamma(\alpha+\beta+1)} = \frac{\alpha}{\alpha+\beta}. \\
E(X^2) &= \frac{\Gamma(\alpha+\beta)\Gamma(\alpha+2)\Gamma(\beta)}{\Gamma(\alpha)\Gamma(\beta)\Gamma(\alpha+\beta+2)} = \frac{(\alpha+1)\alpha}{(\alpha+\beta+1)(\alpha+\beta)}. \\
\sigma^2 &= E(X^2) - [E(X)]^2 = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}.
\end{aligned}$$

3.3.20

$$\begin{aligned} 1 &= c\int_o^3 x(3-x)^4\,dx. \text{ Let } x = 3y,\ \tfrac{dx}{dy} = 3. \\ 1 &= c\int_0^1 (3y)(3-3y)^4 3\,dy = 3^6 c\frac{\Gamma(2)\Gamma(5)}{\Gamma(7)};\ \text{so } c = 6\cdot 5/3^6 = 10/3^5. \end{aligned}$$

3.3.21 If $\alpha = \beta$, show that $f\left(\frac{1}{2}+z\right) = f\left(\frac{1}{2}-z\right)$.

3.3.22 Note that

$$D_z\left[-\sum_{w=0}^{k-1}\binom{n}{w} z^w(1-z)^{n-w}\right] = \frac{n!}{(k-1)!(n-k)!}z^{k-1}(1-z)^{n-k}.$$

3.3.26 (a). Using the mean value theorem, we have

$$r(x) = \lim_{\Delta\to 0}\frac{P(x\le X < x+\Delta)}{\Delta P(X\ge x)} = \lim_{\Delta\to 0}\frac{f(\xi)\Delta}{\Delta(1-F(x))},$$

where $\xi \to x$ as $\Delta \to 0$. Letting $\Delta\to 0$, we then get the desired result.

(d) The pdf of X is

$$f_X(x) = \exp\left\{\frac{c}{b}(1-e^{bx})\right\} ce^{bx},$$

from which the desired result follows.

3.4.1 In the integral for $\Phi(-z)$, let $w = -v$ and it follows that $\Phi(-z) = 1-\Phi(z)$.

3.4.4

$$\begin{aligned} P\left(\frac{X-\mu}{\sigma} < \frac{89-\mu}{\sigma}\right) &= 0.90 \\ P\left(\frac{X-\mu}{\sigma} < \frac{94-\mu}{\sigma}\right) &= 0.95. \end{aligned}$$

Thus $\frac{89-\mu}{\sigma} = 1.282$ and $\frac{94-\mu}{\sigma} = 1.645$. Solve for μ and σ.

3.4.5

$$c2^{-x^2} = ce^{-x^2\log 2} = c\exp\left\{-\frac{(2\log 2)x^2}{2}\right\}.$$

Thus if $c = 1/[\sqrt{2\pi}\sqrt{1/(2\log 2)}]$, we would have a $N(0, 1/(2\log 2))$ distribution.

3.4.6

$$\begin{aligned} E(|X-\mu|) &= 2\int_\mu^\infty (x-\mu)\frac{1}{\sqrt{2\pi}\sigma}e^{-(x-\mu)^2/2\sigma^2}\,dx \\ &= 2\left[\frac{-\sigma}{\sqrt{2\pi}}\exp\left\{-(x-\mu)^2/2\sigma^2\right\}\right]_\mu^\infty = \sigma\sqrt{\frac{2}{\pi}}. \end{aligned}$$

3.4.8

$$\begin{aligned}\int_2^3 \exp\{-(x-3)^2/2(1/4)\}\,dx &= \sqrt{2\pi}\sqrt{1/4}\int_2^3 \frac{1}{\sqrt{2\pi}\sqrt{1/4}}\exp\{-(x-3)^2/2(1/4)\}\,dx \\ &= \sqrt{\frac{\pi}{2}}\left[\Phi\left(\frac{3-3}{1/2}\right)-\Phi\left(\frac{2-3}{1/2}\right)\right] = \sqrt{\frac{\pi}{2}}\left[\frac{1}{2}-\Phi(-2)\right].\end{aligned}$$

3.4.10 Of course, X is $N(3,16)$.

3.4.12

$$P\left[0.0004 < \tfrac{(X-5)^2}{10} < 3.84\right] \text{ and } \tfrac{(X-5)^2}{10} \text{ is } \chi^2(1),$$

so, the answer is $0.95 - 0.05 = 0.90$.

3.4.13

$$\begin{aligned}P(1 < X^2 < 9) &= p(-3 < X < -1) + P(1 < X < 3) \\ &= \left[\Phi\left(\frac{-1-1}{2}\right)-\Phi\left(\frac{-3-1}{2}\right)\right]+\left[\Phi\left(\frac{3-1}{2}\right)-\Phi\left(0\right)\right].\end{aligned}$$

3.4.15

$$\begin{aligned}M(t) &= 1+0+\frac{2!/(2)1!}{2!}t^2+0+\frac{4!/(2^2)2!}{4!}t^4+\cdots \\ &= 1+\frac{t^2/2}{1!}+\frac{(t^2/2)^2}{2!}+\cdots = \exp\{t^2/2\};\end{aligned}$$

so X is $N(0,1)$.

3.4.20

$$\lim_{\sigma^2\to 0}\left[\exp\left\{\mu t+\frac{\sigma^2 t^2}{2}\right\}\right] = \exp\{\mu t\},$$

which is the mgf of a degenerate distribution at $x=\mu$.

3.4.22

$$\int_{-\infty}^{b} yf(y)/F(b)\,dy = -f(b)/F(b).$$

Multiply both sides by $F(b)$ then differentiate both sides with respect to b. This yields,

$$bf(b) = f'(b) \text{ and } -(b^2/2)+c = \log f(b).$$

Thus

$$f(b) = c_1 e^{-b^2/2},$$

which is the pdf of a $N(0,1)$ distribution.

3.4.25 Using $W = ZI_{1-\epsilon} + \sigma_c z(1 - I_{1-\epsilon})$, the independence of Z and $I_{1-\epsilon}$, and $I_{1-\epsilon}^2 = I_{1-\epsilon}$, we get

$$\begin{aligned}
E(W) &= 0(1-\epsilon) + \sigma_c 0[1-(1-\epsilon)] = 0 \\
\mathrm{Var}(W) &= E(W^2) \\
&= e[Z^2 I_{1-\epsilon}^2 + 2\sigma_c Z^2 I_{1-\epsilon}(1-I_{1-\epsilon}) + \sigma_c^2 Z^2 (1-I_{1-\epsilon})^2] \\
&= (1-\epsilon) + \sigma_c^2[1-(1-\epsilon)],
\end{aligned}$$

which is the desired.

3.4.26 If R or SPLUS is available, the code on page 168, i.e.,

```
(1-eps)*pnorm(w)+eps*pnorm(w/sigc)
```

evaluates the contaminated normal cdf with parameters `eps` and `sigc`. Using R, the probability asked for in Part (d) is

```
> eps = .25
> sigc = 20
> w = -2
> w2 = 2
> (1-eps)*pnorm(w)+eps*pnorm(w/sigc)
    + (1-(1-eps)*pnorm(w2)-eps*pnorm(w2/sigc))
[1] 0.2642113
```

3.4.28 Note $X_1 - X_2$ is $N(-1, 2)$. Thus

$$P(X_1 - X_2 > 0) = 1 - \Phi(1/\sqrt{2}) = 1 - \Phi(0.707).$$

3.4.30 The distribution of the sum Y is $N(43, 9)$, so

$$P(Y < 40) = \Phi\left(\frac{40-43}{3}\right) = \Phi(-1).$$

3.5.1 For Part (b),

$$\begin{aligned}
E(Y|x = 3.2) &= 110 + (0.6)\frac{10}{0.4}(3.2 - 2.8) = 116. \\
\mathrm{Var}(Y|x = 3.2) &= 100(1 - 0.36) = 64 \\
\text{Answer} &= \Phi\left(\frac{124-116}{8}\right) - \Phi\left(\frac{106-116}{8}\right) = \Phi(1) - \Phi(-1.25).
\end{aligned}$$

3.5.3

$$\begin{aligned}
\frac{\partial \psi}{\partial t_2} &= \frac{\partial M(t_1,t_2)}{\partial t_2} / M(t_1,t_2) \\
\frac{\partial^2 \psi}{\partial t_1 \partial t_2} &= \left[M(t_1,t_2)\frac{\partial^2 M(t_1,t_2)}{\partial t_1 \partial t_2} - \frac{\partial M(t_1,t_2)}{\partial t_2}\frac{\partial M(t_1,t_2)}{\partial t_1}\right] / M(t_1,t_2)^2 \\
\left.\frac{\partial^2 \psi}{\partial t_1 \partial t_2}\right|_{t_1=t_2=0} &= \frac{\partial^2 M(0,0)}{\partial t_1 \partial t_2} - \frac{\partial M(0,0)}{\partial t_1}\frac{\partial M(0,0)}{\partial t_2},
\end{aligned}$$

because $M(0,0) = 1$. This is the covariance.

3.5.5 Because $E(Y|x=5) = 10 + \rho(5/1)(5-5) = 10$, this probability requires that

$$\frac{16-10}{5\sqrt{1-\rho^2}} = 2,\ \frac{9}{25} = 1-\rho^2, \text{ and } \rho = \frac{4}{5}.$$

3.5.8 $f_1(x) = \int_{-\infty}^{\infty} f(x,y)\,dy = (1/\sqrt{2\pi})\exp\{-x^2/2\}$, because the first term of the integral is obviously equal to the latter expression and the second term integrates to zero as it is an odd function of y. Likewise

$$f_2(y) = \frac{1}{\sqrt{2\pi}}\exp\{-y^2/2\}.$$

Of course, each of these marginal standard normal densities integrates to one.

3.5.9 Similar to 3.5.8 as the second term of

$$\int_{-\infty}^{\infty} f(x,y,z)\,dx$$

equals zero because it is an integral of an odd function of x.

3.5.10 Write

$$\mathbf{Z} = \begin{bmatrix} a & b \end{bmatrix}\begin{bmatrix} X \\ Y \end{bmatrix}.$$

Then apply Theorem 3.5.1.

3.5.14

$$\begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} = \begin{bmatrix} 3 & 1 & -2 \\ 1 & -5 & 1 \end{bmatrix} = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} = \mathbf{BX}.$$

Evaluate $\mathbf{B}\boldsymbol{\mu}$ and $\mathbf{BVB}'$.

3.5.16 Write

$$(X_1 + X_2, X_1 - X_2) = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} X_1 \\ X_2 \end{bmatrix}.$$

Then apply Theorem 3.5.1.

3.5.21 This problem requires statistical software which at least returns the spectral decomposition of a matrix. The following is from an R output where the variable `amat` contains the matrix $\mathbf{\Sigma}$.

```
> sum(diag(amat))
[1] 1026                                         Total Variation

> eigen(amat)
$values
[1] 925.36363  60.51933  25.00226  15.11478   The first eigen value
```

```
                                                    is the variance of the
                                                    first principal component.

$vectors                                            First column is the
                                                    first principal component.

           [,1]       [,2]       [,3]       [,4]
[1,] -0.5357934  0.1912818  0.7050231 -0.4234138
[2,] -0.4320336  0.7687151 -0.3416228  0.3251431
[3,] -0.5834990 -0.4125759 -0.5727115 -0.4016360
[4,] -0.4310468 -0.4497438  0.2413252  0.7441044

> 925.36363/1026
[1] 0.9019139                                        Over 90%
```

3.6.8 Since $F = \frac{U/r_1}{V/r_2}$, then $\frac{1}{F} = \frac{V/r_2}{U/r_1}$, which has an F-distribution with r_2 and r_1 degrees of freedom.

3.6.10 Note

$$T^2 = W^2/(V/r) = (W^2/1)/(V/r).$$

Since W is $N(0,1)$, then W^2 is $\chi^2(1)$, Thus T^2 is F with one and r degrees of freedom.

3.6.12 The change-of-variable technique can be used. An alternative method is to observe that

$$Y = \frac{1}{1+(U/V)} = \frac{V}{V+U},$$

where V and U are independent gamma variables with respective parameters $(r_2/2, 2)$ and $(r_1/2, 2)$. Hence, Y is beta with $\alpha = r_2/2$ and $\beta = r_1/2$.

3.6.13 Note that the distribution of X_i is $\Gamma(1,1)$. It follows that the mgf of $Y_i = 2X_i$ is

$$M_{Y_i}(t) = (1-2t)^{-2/2}, \quad t < 1/2.$$

Hence $2X_i$ is distributed as $\chi^2(2)$. Since X_1 and X_2 are independent, we have that

$$\frac{X_1}{X_2} = \frac{2X_1/2}{2X_2/2}$$

has an F-distribution with $\nu_1 = 2$ and $\nu_2 = 2$ degrees of freedom.

3.6.14 For Part (a), the inverse transformation is $x_1 = (y_1y_2)/(1+y_1)$ and $x_2 = y_2/(1+y_1)$. The space is $y_i > 0$, $i = 1, 2$. The Jacobian is $J = y_2/(1+y_1)^2$. It is easy to show that the joint density factors into two positive functions, one of which is a function of y_1 alone while the other is a function y_2 alone. Hence, Y_1 and Y_2 are independent.

3.7.3 Recall from Section 3.4, that we can write the random variable of interest as

$$X = IZ + 3(1-I)Z,$$

where Z has a $N(0,1)$ distribution, I is 0 or 1 with probabilities 0.1 and 0.9, respectively, and I and Z are independent. Note that $E(X) = 0$ and the variance of X is given by expression (3.4.13); hence, for the kurtosis we only need the fourth moment. Because I is 0 or 1, $I^k = I$ for all positive integers k. Also $I(I-1) = 0$. Using these facts, we see that

$$E(X^4) = .9E(Z^4) + 3^4(.1)E(Z^4) = E(Z^4)(.9 + (.i)3^4).$$

Use expression (1.9.1) to get $E(Z^4)$.

3.7.4 The joint pdf is

$$f_{X,\theta}(x,\theta) = \theta(1-\theta)^{x-1}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1}.$$

Integrating out θ, we have

$$\begin{aligned} f_X(x) &= \int_0^1 \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\theta^{\alpha+1-1}(1-\theta)^{\beta+x-1-1}\,d\theta \\ &= \frac{\Gamma(\alpha+\beta)\Gamma(\alpha+1)\Gamma(\beta+x-1)}{\Gamma(\alpha)\Gamma(\beta)\Gamma(\alpha+\beta+x)}. \end{aligned}$$

3.7.7 Both the pdf and cdf of the Pareto distribution are given on page 193 of the text. Their ratio $(h(x)/(1-H(x)))$, quickly gives the result.

3.7.10 Part (b). The joint pdf of X and α is given by

$$f_{X,\alpha}(x,\alpha) = \frac{\alpha\tau x^{\tau-1}}{(1+\beta x^\tau)^{\alpha+1}}e^{-\alpha/\beta}.$$

Integrating out α, we have

$$\begin{aligned} f_X(x) &= \frac{\tau x^{\tau-1}}{1+\beta x^\tau}\int_0^\infty e^{-\alpha[\log(1+\beta x^\tau)+(1/\beta)]}\,d\alpha \\ &= \frac{\tau x^{\tau-1}}{1+\beta x^\tau}[\log(1+\beta x^\tau)+(1/\beta)]^{-1}, \quad x > 0. \end{aligned}$$

Chapter 4

Some Elementary Statistical Inferences

4.1.1 Parts (b), (c), and (d).

(b) The likelihood function is

$$L(\theta) = \prod_{i=1}^{n} \left(\frac{1}{\theta}\right) e^{-x_i/\theta} = \theta^{-n} e^{-\sum_{i=1}^{n} x_i/\theta}.$$

Hence,

$$l(\theta) = -n \log \theta - \frac{1}{\theta} \sum_{i=1}^{n} x_i.$$

So,

$$\frac{\partial l(\theta)}{\partial \theta} = \frac{-n}{\theta} + \frac{1}{\theta^2} \sum_{i=1}^{n} x_i,$$

resulting in the mle $\widehat{\theta} = \overline{X}$. For the data in this problem, the estimate of θ is 101.15.

(c) Since the cdf $F(x) = 1 - e^{-x/\theta}$, the population median is ξ where ξ solves the equation $e^{-x/\theta} = 1/2$; hence, $\xi = \theta \log 2$. The sample median is an estimator of ξ. For the data set of this problem, the sample median is 55.5.

(d) Because the mle of θ is $\overline{X}$, the mle of the population median is $\overline{X} \log 2$. For the data of this problem, this estimate is $101.15 \log 2 = 70.11$.

4.1.2 Parts (c) and (d). The parameter of interest is

Part (c) Using the binomial model, the estimate of $P(X > 215)$ is

$$\widehat{p}_b = \frac{\#\{x_i > 215\}}{26} = \frac{7}{26} = 0.2692.$$

Part (d) Under the normal probability model, the parameter of interest is

$$\begin{aligned} p &= P(X > 215) = P\left(Z > \frac{215-\mu}{\sigma}\right) \\ p &= 1 - \Phi\left(\frac{215-\mu}{\sigma}\right). \end{aligned}$$

Because $\overline{X}$ and $\widehat{\sigma}^2 = n^{-1}\sum_{i=1}^n (X_i - \overline{X})^2$ are the mles of μ and σ^2, respectively, the mle of p is

$$\widehat{p}_N = 1 - \Phi\left(\frac{215 - \overline{X}}{\hat{\sigma}}\right).$$

For the data in this problem, $\overline{x} = 201$ and $\widehat{\sigma} = 17.144$. Hence, a calculation using a computer package or using the normal tables results in $\widehat{p}_N = 0.2701$ as the mle estimate of p.

4.1.5 Parts (a) and (b).

Part (a). Using conditional expectation we have

$$\begin{aligned} P(X_1 \le X_i, i = 2, 3, \ldots, j) &= E[P(X_1 \le X_i, i = 2, 3, \ldots, j | X_1)] \\ &= E[(1 - F(X_1))^{j-1}] \\ &= \int_0^1 u^{j-1}\, du = j^{-1}, \end{aligned}$$

where we used the fact that the random variable $F(X_1)$ has a uniform$(0,1)$ distribution.

Part (b). In the same way, for $j = 2, 3, \ldots$

$$\begin{aligned} P(Y = j-1) &= P(X_1 \le X_2, \ldots, X_1 \le X_{j-1}, X_j > X_1) \\ &= E[(1 - F(X_1))^{j-2} F(X_1)] = \int_0^1 u^{j-2}(1-u)\, du \\ &= \frac{1}{j(j-1)}. \end{aligned}$$

4.1.6 It follows that

$$\begin{aligned} E[\widehat{p(a_j)}] &= \frac{1}{n}\sum_{i=1}^n E[I_j(X_i)] = \frac{1}{n}\sum_{i=1}^n P[X_i = a_j] \\ &= \frac{1}{n}\sum_{i=1}^n p(a_j) = p(a_j). \end{aligned}$$

Hence, the estimator is unbiased. Using independence, its variance is

$$\begin{aligned} V[\widehat{p(a_j)}] &= \frac{1}{n^2}\sum_{i=1}^n V[I_j(X_i)] = \frac{1}{n^2}\sum_{i=1}^n \{p(a_j)[1 - p(a_j)]\} \\ &= \frac{p(a_j)[1 - p(a_j)]}{n}. \end{aligned}$$

4.1.8 If $X_1, \ldots, X_n$ are iid with a Poisson distribution having mean λ, then the likelihood function is

$$L(\lambda) = \prod_{i=1}^{n} e^{-\lambda} \frac{\lambda^{x_i}}{x_i!} = e^{-n\lambda} \frac{\lambda^{\sum_{i=1}^{n} x_i}}{\prod_{i=1}^{n} x_i!}.$$

Taking the partial of the log of this likelihood function leads to $\overline{x}$ as the mle of λ. Hence, the mle of the pmf at k is

$$\widehat{p(k)} = e^{-\overline{x}} \frac{\overline{x}^k}{k!}$$

and the mle of $P(X \geq 6)$ is

$$\widehat{P(X \geq 6)} = e^{-\overline{x}} \sum_{k=6}^{\infty} \frac{\overline{x}^k}{k!}.$$

For the data set of this problem, we obtain $\overline{x} = 2.1333$. Using R, the mle of $P(X \geq 6)$ is `1 - ppois(5, 2.1333) = 0.0219`. Note, for comparison, from the tabled data, that the nonparametric estimate of this probability is 0.033.

4.1.11 Note in this part of the example that x is fixed and by the Mean Value Theorem that ξ is such that $x-h < \xi < x+h$ and $F(x+h)-F(x-h) = 2hf(\xi)$.

Part(a) The mean of the estimator is

$$\begin{aligned} E[\widehat{f(x)}] &= \frac{1}{2hn} \sum_{i=1}^{n} E[I_i(x)] = \frac{1}{2hn} \sum_{i=1}^{n} [F(x+h) - F(x-h)] \\ &= \frac{n2hf(\xi)}{2hn} = f(\xi). \end{aligned}$$

Hence, the bias of the estimate is $f(\xi) - f(x)$ which goes to 0 as $h \to 0$.

Part (b) Since $I_i(x)$ is a Bernoulli indicator, the variance of the estimator is

$$\begin{aligned} V[\widehat{f(x)}] &= \frac{1}{4h^2n^2} \sum_{i=1}^{n} [F(x+h) - F(x-h)][1 - [F(x+h) - F(x-h)]] \\ &= \frac{f(\xi)[1 - 2hf(\xi)]}{2hn}. \end{aligned}$$

Note for this variance to go to 0 as $h \to 0$ and $n \to \infty$, h must be of order n^δ for $\delta > -1$.

4.2.7 $1 = (1.645)(3/\sqrt{n})$; $\sqrt{n} = 4.935$; $n = 24.35$; so take $n = 25$.

4.2.10 (a). $\overline{X} \pm 1.96\sigma/\sqrt{9}$, length $= (2)(1.96)\sigma/3 = 1.31\sigma$.

(b). $\overline{X} \pm 2.306S/\sqrt{8}$, length $= (2)(2.306)S/\sqrt{8}$. Since

$$\begin{aligned} E(S) &= (\sigma/\sqrt{n})\int_0^\infty w^{1/2}\frac{w^{4-1}e^{-w/2}}{\Gamma(4)2^4}\,dw \\ &= (\sigma/\sqrt{9})\frac{\Gamma(9/2)2^{9/2}}{\Gamma(4)2^4} = \frac{\sigma(7/2)(5/2)(3/2)(1/2)\Gamma(1/2)\sqrt{2}}{3\cdot 3\cdot 2\cdot 1} \\ &= \frac{35\sqrt{2\pi}\sigma}{(6)(16)} = (0.914)\sigma, \\ &= E(\text{length}) = \left[(2)(2.306)(0.914)/\sqrt{8}\right]\sigma = 1.49\sigma. \end{aligned}$$

4.2.11 $\frac{(\overline{X}-X_{n+1})/\sqrt{\sigma^2/n+\sigma^2}}{\sqrt{(nS^2/\sigma^2)/(n-1)}} = \sqrt{\frac{n-1}{n+1}}\frac{\overline{X}-X_{n+1}}{S}$ is $T(n-1)$.
$P(-1.415 < \sqrt{\frac{7}{9}}\left(\frac{\overline{X}-X_{n+1}}{S}\right) < 1.415) = 0.80$, or equivalently,
$P(\overline{X} - 1.415\sqrt{9/7}S < X_{n+1} < \overline{X} + 1.415\sqrt{9/7}S) = 0.80$

4.2.13 $c_1(\mu) = \mu - 2\sigma/\sqrt{n} < \overline{X} < \mu + 2\sigma/\sqrt{n} = c_2(\mu)$
is equivalent to
$d_1(\overline{X}) = \overline{X} - 2\sigma/\sqrt{n} < \mu < \overline{X} + 2\sigma/\sqrt{n} = d_2(\overline{X})$.

4.2.14

$$\begin{aligned} -2 &< 5\overline{X}/2\beta - 10 < 2, \\ 8 &< 5\overline{X}/2\beta < 12, \\ \tfrac{5\overline{X}}{24} &< \beta < \tfrac{5\overline{X}}{16}. \end{aligned}$$

4.2.16 $\frac{y}{n} \pm 1.645\sqrt{\frac{(y/n)(1-y/n)}{n}}$
$2(1.645)\sqrt{\frac{(y/n)(1-y/n)}{n}} \le 2(1.645)\sqrt{\frac{(1/2)(1/2)}{n}} = 0.02$
$\frac{1.645}{2(0.01)} = \sqrt{n}; n \approx 7675.$

4.2.18 (c). Use the fact that $\sum(X_i-\mu)^2/\sigma^2$ is $\chi^2(n)$.

4.2.19 $E[\exp\{t(2X/\beta)\}] = [1-\beta(2t/\beta)]^{-3} = (1-2t)^{-6/2}$.
Since $2X/\beta$ is $\chi^2(6), 2\sum X_i/\beta$ is $\chi^2(6n)$. Using tables for $\chi^2(6n)$, find a and b such that

$$P\left(a < 2\sum X_i/\beta < b\right) = 0.95$$

or, equivalently,

$$P\left(\frac{2\sum X_i}{b} < \beta < \frac{2\sum X_i}{a}\right) = 0.95.$$

4.2.23 Use the fact that

$$P\left[-1.96 < \frac{\overline{X}-\overline{Y}-(\mu_1-\mu_2)}{\sqrt{\sigma_1^2/n+\sigma_2^2/m}} < 1.96\right] = 0.95$$

and solve the inequalities so that $\mu_1 - \mu_2$ is in the middle.

4.2.24 Say Z is the $N(0,1)$ random variable used in 6.32. Thus

$$\frac{Z}{\sqrt{\frac{nS_1^2/\sigma_1^2+mS_2^2/\sigma_2^2}{n+m-2}}} \text{ is } T(n+m-2).$$

However, the unknown variances cannot be eliminated from the expression as can be when $\sigma_1^2 = \sigma_2^2$ but unknown. But if $\sigma_1^2 = k\sigma_2^2$, k known, then that ratio can be written (replacing σ_1^2 by $k\sigma_2^2$) without involving the unknown σ_2^2. It still has a t-distribution with $n+m-2$ degrees of freedom.

4.2.26 The distribution of $\overline{X}$ is $N(\mu_1, \sigma^2/n)$ and the distribution of $\overline{Y}$ is $N(\mu_2, \sigma^2/n)$. Because the samples are independent the distribution of $\overline{X} - \overline{Y}$ is $N(\mu_1 - \mu_2, 2\sigma^2/n)$. After some algebra, the equation to solve for n can be written as

$$P\left[\left|\frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sigma/\sqrt{n}}\right| < \frac{\sqrt{n}}{5}\right] = 0.90,$$

which is equivalent to

$$P\left[|Z| < \frac{\sqrt{n}}{5}\right] = 0.90,$$

where Z has a $N(0,1)$ distribution. Hence, $\sqrt{n}/5 = 1.645$ or $n = 67.65$, i.e., $n = 68$.

4.3.1 Note that

$$\begin{aligned}\int_0^p \frac{n!}{(k-1)!(n-k)!} z^{k-1}(1-z)^{n-k}\,dz \quad &+ \quad \int_p^1 \frac{n!}{(k-1)!(n-k)!} z^{k-1}(1-z)^{n-k}\,dz \\ &= \quad \sum_{w=0}^{n} \binom{n}{w} p^w (1-p)^{n-w}.\end{aligned}$$

Then using Exercise 3.3.22 we have the result, i.e.,

$$\int_0^p \frac{n!}{(k-1)!(n-k)!} z^{k-1}(1-z)^{n-k}\,dz + = \sum_{w=k}^{n} \binom{n}{w} p^w (1-p)^{n-w}.$$

4.3.4 For Part (a), use Exercise 3.3.5 or reason as follows. Let W_n be the waiting time until the nth event. Then $W_n > 1$ if and only if at most $n-1$ events occurred in the the interval $(0,1]$. Since W_n has a $\Gamma(n, 1/\lambda)$ distribution, we have

$$\frac{\lambda^n}{\Gamma(n)} \int_1^\infty x^{n-1} e^{-x\lambda}\,dx = \sum_{j=0}^{n-1} e^{-\lambda} \frac{\lambda^j}{j!}.$$

In the integral, make the substitution $z = x\lambda$. This results in the identity

$$\frac{1}{\Gamma(n)} \int_\lambda^\infty z^{n-1} e^{-z}\,dx = \sum_{j=0}^{n-1} e^{-\lambda} \frac{\lambda^j}{j!}.$$

For Part(b), replace n by $n\overline{x}+1$ and replace λ by $n\theta$ which yields the result.

4.4.2 Part (b). The cdf and its inverse are

$$\begin{aligned} F(x) &= \frac{1}{1+e^{-x}}, \quad -\infty < x < \infty \\ F^{-1}(u) &= \log\left[\frac{1-u}{u}\right], \quad 0 < u < 1. \end{aligned}$$

Hence, $\xi_{.25} = \log(.25/.75) = -1.099$. Because the pdf is symmetric about 0, $\xi_{.75} = 1.099$. Thus $h = 1.5(\xi_{.75} - \xi_{.25}) = 3.296$. Thus, the upper inner fence is $\xi_{.75} + h = 4.395$ and the probability of a potential outlier is

$$2[1 - F(4.395)] = 0.0244.$$

4.4.5 The cdf of the Y_4 is

$$P(Y_4 \le t) = (1 - e^{-t})^4, \quad t > 0.$$

Hence, $P(Y_4 \ge 3) = 1 - (1 - e^{-3})^4 = 0.1848$.

4.4.7 Since the distribution is of the discrete type, we cannot use the formulas in the book. However,

$$\begin{aligned} P(Y_1 = y_1) &= P(\text{all} \ge y_1) - P(\text{all} \ge y_1 + 1) \\ &= \left(\frac{7-y_1}{6}\right)^5 - \left(\frac{6-y_1}{6}\right)^5. \end{aligned}$$

4.4.9 Here $F(x) = x,\ 0 \le x \le 1$. Thus, using the Remark,

$$g_k(y_k) = \frac{n!}{(k-1)!(n-k)!} y_k^{k-1}(1-y_k)^{n-k}(1), \quad 0 < y_k < 1,$$

which is beta $(\alpha = k, \beta = n-k+1)$.

4.4.11 The distribution of the range $Y_4 - Y_1$ could be found. An alternative method is

$$P(Y_4 - Y_1 < 1/2) = 1 - \int_0^{1/2} \int_{y_1+1/2}^{1} 12(y_4 - y_1)^2 \, dy_4 \, dy_1.$$

4.4.12 $y_1 = z_1 z_2 z_3$, $y_2 = z_2 z_3$, $y_3 = z_3$, with $J = z_2 z_3^2$, $0 < z_1 < 1$, $0 < z_2 < 1$, $0 < z_3 < 1$. Accordingly,

$$\begin{aligned} g(z_1, z_2, z_3) &= 3!\, 2(z_1 z_2 z_3)\, 2(z_2 z_3)\, 2(z_3) z_2 z_3^2 \\ &= (2z_1)(4z_2^3)(6z_3^5), \quad 0 < z_i < 1, \quad i = 1,2,3. \end{aligned}$$

4.4.13 $P(2Y_1 < Y_2) = \int_0^{1/2} \int_{2y_1}^{1} 8(1-y_1)(1-y_2)\, dy_2\, dy_1$.

4.4.15 (a).

$$\begin{aligned} E(Y_1) &= \int_{-\infty}^{\infty}\int_{-\infty}^{y_2} y_1\left(\frac{2}{2\pi\sigma^2}\right)\exp\left\{-\frac{y_1^2+y_2^2}{2\sigma^2}\right\} dy_1\,dy_2 \\ &= \left(\frac{-2}{2\pi\sigma^2}\right)\int_{-\infty}^{\infty}\sigma^2\exp\left\{-\frac{2y_2^2}{2\sigma^2}\right\} dy_2 = \left(\frac{-1}{\pi}\right)\sqrt{2\pi\sigma^2/2} = -\sigma/\sqrt{\pi}. \end{aligned}$$

4.4.17

$$\begin{aligned} F(x) &= x^2,\ \ 0 \le x < 1. \\ g_{34}(y_3, y_4) &= \frac{4!}{2!}(y_3^2)^2(2y_3)(2y_4),\ \ 0 < y_3 < y_4 < 1. \\ g_4(y_4) &= 4(y_4^2)^3(2y_4) = 8y_4^7,\ \ 0 < y_4 < 1. \\ g_{3|4}(y_3|y_4) &= 6y_3^5/y_4^6,\ \ 0 < y_3 < y_4. \\ E(Y_3|y_4) &= (6/7)y_4. \end{aligned}$$

4.4.18 To form a triangle, we must have $y_1 < 1-y_1$, $y_2-y_1 < y_1+(1-y_2)$, $1-y_2 < y_2$. That is, $y_1 < 1/2$, $y_2 > 1/2$, $y_2-y_1 < 1/2$; so answer $= \int_0^{1/2}\int_{1/2}^{1/2+y_1} 2\,dy_2\,dy_1 = \frac{1}{4}$.

4.4.19

$$\begin{aligned} g(u,v) &= (2u)(3v^2)(1) + (2v)(3u^2)(1) \\ &= 6uv(v+u),\ \ 0 < u < v < 1, \end{aligned}$$

since each Jacobian is equal to one.

4.4.21 (a). Since Y_{10} is greater than 9 other Y values, $a_{10} = 9$. Since Y_9 is greater than 8 others but smaller than one, $a_9 = 7$. And so on. Thus,

$$G = (9Y_{10} + 7Y_9 + 5Y_8 + 3Y_7 + Y_6 - Y_5 - 3Y_4 - 5Y_3 - 7Y_2 - 9Y_1)/45.$$

(b). It follows from Exercise 3.4.6 that $E(|X_i - X_j|) = 2\sigma/\sqrt{\pi}$, and G is the mean of $\binom{n}{2}$ such absolute differences. Thus $E(G) = 2\sigma/\sqrt{\pi}$.

4.4.22 (a). $y_1 = z_1/n$, $y_2 = z_2/(n-1) + z_1/n$, $y_3 = z_3/(n-2) + z_2/(n-1) + z_1/n$, etc., which has $J = 1/n!$. Moreover, $0 < y_1 < y_2 < \cdots < y_n < \infty$ maps onto $0 < z_i < \infty$, $i = 1, 2, \ldots, n$. Thus

$$g(z_1, z_2, \ldots, z_n) = \left(\frac{1}{n!}\right)(n!e^{-z_1-z_2-\cdots-z_n}) = e^{-z_1-z_2-\cdots-z_n}.$$

That $Z_1, Z_2, \ldots, Z_n$ are independent, each with an exponential distribution.

4.4.24 Let $F(x)$ denote the common cdf of the sample. Then $\xi_{0.9} = F^{-1}(0.9)$. The solution to the desired inequality is

$$\begin{aligned}
1-(F(\xi_{0.9}))^n &\geq 0.75 \\
1-F(F^{-1}(0.9)))^n &\geq \frac{3}{4} \\
1-0.9^n &\geq \frac{3}{4} \\
n\log(0.9) &\leq \frac{1}{4} \\
n &\geq -\frac{\log(4)}{\log(0.9)} = 13.14.
\end{aligned}$$

Hence, take $n = 14$.

4.4.27 $\sum_{w=1}^{n-1}\binom{n}{w}\left(\frac{1}{2}\right)^n = 1-2\left(\frac{1}{2}\right)^n \geq 0.99;\;\; 0.01 \geq \frac{1}{2^{n-1}};\;\; n=8.$

4.4.28 (a). It follows from Exercise 3.4.6 that $E(Y_1) = \mu - \sigma/\sqrt{\pi}$. So $E(Y_2 - Y_1) = 2\sigma/\sqrt{\pi} \approx 1.13\sigma$.

(b). $\overline{X} \pm (0.65)\sigma/\sqrt{2}$ is a 50 percent confidence interval for μ with length $(0.65)\sqrt{2}\sigma = 0.92\sigma$.

4.5.3 For a general θ the probability of rejecting H_0 is

$$\gamma(\theta) = \int_{3/4}^{1}\int_{3/4x_1}^{1} \theta^2(x_1x_2)^{\theta-1}\,dx_2\,dx_1 = 1-\left(\frac{3}{4}\right)^\theta + \theta\left(\frac{3}{4}\right)^\theta \log\left(\frac{3}{4}\right)$$

$\gamma(1)$ is the significance level and $\gamma(2)$ is the power when $\theta = 2$.

4.5.5 $\frac{(1/2)^2\exp\{-(x_1+x_2)/2\}}{\exp\{-(x_1+x_2)\}} \leq \frac{1}{2};\; \frac{1}{4}\exp\{(x_1+x_2)/2\} \leq \frac{1}{2}$. So $x_1 + x_2 \leq 2\log 2$ describes the critical region.

4.5.8

$$\begin{aligned}
\gamma(\theta) &= P(\overline{X} \geq c;\theta) = P\left(\frac{\overline{X}-\theta}{5000/\sqrt{n}} \geq \frac{c-\theta}{5000/\sqrt{n}};\theta\right) \\
&= 1-\Phi\left(\frac{c-\theta}{5000/\sqrt{n}}\right).
\end{aligned}$$

Thus, solve for n and c knowing that

$$\frac{c-30000}{5000/\sqrt{n}} = 2.325 \;\text{ and }\; \frac{c-35000}{5000/\sqrt{n}} = -2.05.$$

4.5.10

$$\begin{aligned}
\gamma(p) &= P(Y \geq c;p) = P\left(\frac{Y-np}{\sqrt{np(1-p)}} \geq \frac{c-np}{\sqrt{np(1-p)}};p\right) \\
&\approx 1-\Phi\left(\frac{c-np}{\sqrt{np(1-p)}}\right).
\end{aligned}$$

So solve for n and c knowing that approximately

$$\frac{c - n(1/2)}{\sqrt{n(1/2)(1/2)}} = 1.282, \quad \frac{c - n(2/3)}{\sqrt{n(2/3)(1/3)}} = -1.645.$$

4.5.12 Let $Y = \sum_{i=1}^{8} X_i$. Then Y has a Poisson(8μ) distribution.

Part (a). The significance level of the test is

$$\alpha = P_{H_0}[Y \geq 8] = P[\text{Poisson}(4) \geq 8] = 0.051.$$

Part (b). The power function is

$$\gamma(\mu) = P_\mu[Y \geq 8] = P[\text{Poisson}(8\mu) \geq 8].$$

Part (c). $\gamma(0.75) = 0.256$.

4.6.2 Suppose $\mu > \mu_0$. Then

$$\left| \frac{\sqrt{n}(\mu_0 - \mu)}{\sigma} + z_{\alpha/2} \right| < \left| \frac{\sqrt{n}(\mu_0 - \mu)}{\sigma} - z_{\alpha/2} \right|.$$

Hence,

$$\phi\left(\left| \frac{\sqrt{n}(\mu_0 - \mu)}{\sigma} + z_{\alpha/2} \right| \right) > \phi\left(\left| \frac{\sqrt{n}(\mu_0 - \mu)}{\sigma} - z_{\alpha/2} \right| \right).$$

Because $\phi(t)$ is symmetric about 0, $\phi(t) = \phi(|t|)$. This observation plus the last inequality shows that $\gamma'(\mu)$ is increasing, (for $\mu > \mu_0$). Likewise for $\mu < \mu_0$, $\gamma'(\mu)$ is decreasing.

4.6.3 Under H_0, the statistic $t = (\overline{X} - \mu_0)/(S/\sqrt{n})$ has a t-distribution with $n-1$ degrees of freedom. Hence,

$$P_{H_0}[|t| > t_{\alpha/2,n-1}] = \alpha.$$

4.6.5 (a). The critical region is

$$t = \frac{\overline{x} - 10.1}{s/\sqrt{15}} \geq 1.753.$$

The observed value of t,

$$t = \frac{10.4 - 10.1}{0.4/\sqrt{15}} = 2.90,$$

is greater than 1.753 so we reject H_0.

(b). Since $t_{0.005}(15) = 2.947$ (from other tables), the approximate p-value of this test is 0.005.

4.6.7 Assume that Y and Y are normally distributed. Then the t-statistic

$$t = \frac{\overline{X} - \overline{Y}}{S_p\sqrt{(1/n_1) + (1/n_2)}}$$

has under H_0 a t-distribution with $n_1 + n_2 - 2$ degrees of freedom. A level α test for the alternative $H_A: \ \mu_1 < \mu_2$ is

Reject H_0 in favor of H_A, if $t < -t_{\alpha, n_1+n_2-2}$.

For Part (b), based on the data we have,

$$\begin{aligned} s_p^2 &= \frac{(13-1)25.6^2 + (16-1)28.3^2}{27} \\ s_p &= \sqrt{s_p^2} = 27.133 \\ t &= \frac{72.9 - 81.7}{27.133\sqrt{(1/13) + (1/16)}} = -0.8685. \end{aligned}$$

Since $t = -0.8685 \not< -t_{.05,27} = -1.703$, we fail to reject H_0 at level 0.05. The p-value is $P[t(27) < -0.8685] = 0.1964$.

4.6.8 For Parts (a) - (c):

Part (a) $H_0 : p = 0.14$; $H_1 : p > 0.14$;

Part (b) $C = \{z : z \geq 2.326\}$ where $z = \frac{y/n - 0.14}{\sqrt{(0.14)(0.86)/n}}$;

Part (c) $z = \frac{104/590 - 0.14}{\sqrt{(0.14)(0.86)/590}} = 2.539 > 2.326$
so H_0 is rejected and conclude that the campaign was successful.

4.7.1 $p_{10} = \int_0^{1/2} \frac{2-x}{2}\, dx = \frac{1}{2} - \frac{1}{16} = \frac{7}{16}$.
Likewise $p_{20} = 5/16, p_{30} = 3/16, p_{40} = 1/16$.
$Q_3 = \frac{(30-35)^2}{35} + \frac{(30-25)^2}{25} + \frac{(10-15)^2}{15} + \frac{(10-15)^2}{5} = 8.38$.
However, $8.38 > 7.81$ so we reject H_0 at $\alpha = 0.05$.

4.7.3 $Q_5 = \frac{(b-20)^2}{20} + \frac{(40-b-20)^2}{20} = \frac{(b-20)^2}{10} = 12.8$,
which is the 97.5 percentile of a $\chi^2(5)$ distribution. Thus $(b-20)^2 = 128$ and $b = 20 \pm 11.3$. Hence $b < 8.7$ or $b > 31.3$ would lead to rejection.

4.7.7 The maximum likelihood statistic for p is defined by that value of p which maximizes

$$\frac{n!}{x_1!x_2!x_3!}[p^2]^{x_1}[2p(1-p)]^{x_2}[(1-p)^2]^{x_3};$$

it is $\hat{p} = (2X_1 + X_2)/(2X_1 + 2X_2 + 2X_3)$. Thus if $\hat{p}_1 = \hat{p}^2$, $\hat{p}_2 = 2\hat{p}(1-\hat{p})$, and $\hat{p}_3 = (1-\hat{p})^2$, the random variable $\sum_1^3 (X_i - n\hat{p}_i)^2/n\hat{p}_i$ has an approximate chi-square distribution with $3 - 1 - 1 = 1$ degree of freedom.

4.7.8 The expected value of each cell is 15; thus the chi-square statistic equals

$$\frac{4(3k)^2}{15} + \frac{4(k)^2}{15} = \frac{40k^2}{15} \geq 12.6,$$

which is the 95th percentile of a $\chi^2(6)$ distribution. Thus $k > \sqrt{(3/8)(12.6)} = 2.17$. So $k = 3$.

4.8.1 Suppose $0 < z < 1$. Then

$$P(Z \leq z) = P[F(X) \leq z] = P[X \leq F^{-1}(z)] = F[F^{-1}(z)] = z.$$

Hence, Z has a uniform $(0, 1)$ distribution.

4.8.3 Note that

$$1.96 \int_0^{1.96} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}u^2\right\} \frac{1}{1.96}\, du = 1.96E\left[\frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}U^2\right\}\right],$$

where U has a uniform distribution on $(0, 1.96)$. The following R-code draws 10,000 variates $Z_i = 1.96\frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}U_i^2\right\}$ where U_i are iid with a common uniform distribution on $(0, 1.96)$. A 95% confidence interval for mean of Z_i is obtained. Notice that it does trap the true mean $\mu = 0.475$.

```
> u = runif(10000,0,1.96)
> z = 1.96*(1/sqrt(2*pi))*exp(-u^2/2)
> mean(z)
[1] 0.4750519                    *** Estimate of mu
> se = var(z)^.5/sqrt(10000)
> se
[1] 0.002225439                  *** standard error of estimation
> cil = mean(z) - 1.96*se
> ciu = mean(z) + 1.96*se
> cil
[1] 0.4706901                    *** Lower limit of CI
> ciu
[1] 0.4794138                    *** Upper limit of CI
```

4.8.5 The cdf of the logistic distribution is

$$F(x) = \frac{1}{1 + e^{-x}}, \quad -\infty < x < \infty.$$

To determine the inverse of this function, set $u = 1/(1 + e^{-x})$ and then solve for x. After some algebra, we get

$$F^{-1}(u) = \log \frac{u}{1 - u}.$$

Hence, if U is uniform $(0, 1)$ then $\log[U/(1-U)]$ has a logistic distribution with cdf $F(x)$. The following R function returns a random sample of n observations from this logistic distribution:

```
rlogist = function(n){
  u = runif(n)
  rlogist = log(u/(1-u))
  rlogist
}
```

4.8.7 First show that the cdf of the Laplace distribution is given by

$$F(t) = \begin{cases} \frac{1}{2}e^{t} & -\infty < t < 0 \\ 1 - \frac{1}{2}e^{-t} & 0 < t < \infty. \end{cases}$$

Then show that the inverse of the cdf is

$$F^{-1}(u) = \begin{cases} \log(2u) & 0 < u < \frac{1}{2} \\ -\log(2 - 2u) & \frac{1}{2} < u < 1. \end{cases}$$

Hence, if U is uniform$(0,1)$ then $F^{-1}(U)$ has the Laplace pdf (5.2.9). The following R-code generates n observations from this Laplace distribution.

```
> uni = runif(n)
> x=rep(0,n)
> x[uni<.5]=log(2*uni[uni<.5])
> x[uni>=.5]=-log(2-2uni[uni>=.5])
```

4.8.10 By a simple change of variable ($z = x^3/\theta^3$) in its integrand (pdf), the cdf is

$$F(t) = 1 - \exp\left\{\frac{t^3}{\theta^3}\right\}, \quad t > 0.$$

Its inverse is given by

$$F^{-1}(u) = -\theta[\log(1-u)]^{1/3}, \quad 0 < u < 1.$$

Hence, if U has a uniform $(0,1)$ distribution then $F^{-1}(U)$ has the Weibull distribution.

4.8.12 The logistic cdf corresponding to the pdf given in expression (4.4.9) is $F(x) = 1/(1+e^{-x})$, $-\infty < x < \infty$. Its inverse function is $F^{-1}(u) = \log[u/(1-u)]$, $0 < u < 1$. Hence, if $U_1, U_2, \ldots, U_{20}$ is a random sample of size 20 from the uniform $(0,1)$ distribution then $X_1, X_2, \ldots, X_{20}$, where $X_i = F^{-1}(U_i)$, is a random sample of size 20 from this logistic distribution. Use this and the algorithm given on page 267.

4.8.17 By simple differentiation the derivative of the ratio is

$$D_x = -x\exp\left\{-\frac{x^2}{2}\right\}(x^2 - 1).$$

hence, ± 1 are critical values. The second derivative is

$$D_{xx} = \exp\left\{-\frac{x^2}{2}\right\}(x^4 - 4x^2 + 1).$$

Notice that it is negative at ± 1; hence, ± 1 are minimums.

4.8.18 Parts (a) and (b).

Part(a) Note that $F(x) = x^{\beta}$, which has the inverse function $F^{-1}(u) = u^{1/\beta}$.

Part(b) There are many accept-reject algorithms to generate observations from this distribution. One such algorithm is to take Y to have a uniform $(0,1)$ distribution and $M = \beta$. Then it follows that $f(x) \leq Mg(x)$, because $0 < x < 1$ and $\beta > 1$. The following R function returns n observations from this distribution based on this accept-reject algorithm.

```
rpareto = function(n,beta){
   ic = 0
   x = rep(0,n)
   while(ic <= n){
     u1 = runif(1)
     u2 = runif(1)
     chk = u1^(beta-1)
     if(u2 <= chk){
         ic = ic + 1
         x[ic] = u1
     }
   }
   x
}
```

.

4.8.21 If $W = U^2 + V^2 > 1$ the algorithm begins anew. So suppose $W < 1$. Note that X_1 and X_2 are functions of U and V. So first we get the conditional distribution of U and V given $U^2 + V^2 < 1$. But this is easily seen to be a uniform distribution over the unit circle. Hence, the conditional pdf of (U, V) is

$$f_{U,V|W<1}(u, v|w < 1) = \frac{1}{\pi}, \quad u^2 + v^2 < 1.$$

Now transform to polar coordinates. Let

$$\begin{aligned} u &= r\sin\theta, \quad 0 < \theta < 2\pi, \\ v &= r\cos\theta, \quad 0 < r < 1. \end{aligned}$$

The partials are

$$\begin{array}{ll} \frac{\partial u}{\partial r} = \sin\theta & \frac{\partial u}{\partial \theta} = r\cos\theta \\ \frac{\partial v}{\partial r} = \cos\theta & \frac{\partial v}{\partial \theta} = -r\sin\theta. \end{array}$$

It follows that the Jacobian is r. Hence, the conditional pdf of R, Θ given $W < 1$ is

$$f_{R,\Theta|W<1}(r, \theta|w < 1) = \frac{1}{\pi}r, \quad 0 < \theta < 2\pi, \ 0 < r < 1. \tag{4.0.1}$$

Now transform to X_1 and X_2, which gives

$$\begin{aligned} x_1 &= r\sin\theta \left[-4\frac{\log r}{r^2}\right]^{1/2}, \quad -\infty < x_1 < \infty, \\ x_2 &= r\cos\theta \left[-4\frac{\log r}{r^2}\right]^{1/2}, \quad -\infty < x_2 < \infty. \end{aligned}$$

For the inverse transform, note that

$$\begin{aligned} x_1^2 &= r^2\sin^2\theta \left[-4\frac{\log r}{r^2}\right] \\ x_2^2 &= r^2\cos^2\theta \left[-4\frac{\log r}{r^2}\right], \end{aligned}$$

which leads to $x_1^2 + x_2^2 = -4\log r$ or

$$r = \exp\left\{-\frac{1}{4}(x_1^2 + x_2^2)\right\}. \tag{4.0.2}$$

For θ, note that $x_1/x_2 = \tan\theta$, or

$$\theta = \tan^{-1}\left(\frac{x_1}{x_2}\right). \tag{4.0.3}$$

Taking partials, we get the Jacobian

$$J = \begin{vmatrix} \frac{-2x_1}{4}r & \frac{-2x_2}{4}r \\ \frac{x_2}{r^2} & \frac{-x_1}{r^2} \end{vmatrix} = \frac{r}{2}. \tag{4.0.4}$$

Putting (4.0.1), (4.0.2) and (4.0.4) together, we have the pdf of X_1 and X_2, (note, by the algorithm, this is the unconditional pdf of X_1 and X_2),

$$f_{X_1,X_2}(x_1,x_2) = \frac{1}{\pi}\exp\left\{-\frac{1}{2}(x_1^2 + x_2^2)\right\}, \quad -\infty < x_1 < \infty,\ -\infty < x_2 < \infty.$$

Thus, X_1 and X_2 are iid $N(0,1)$ random variables.

4.9.1 (a). This follows immediately because the sampling is with replacement.

(b).

$$E(x_i^*) = \sum_{j=1}^{n} x_j \frac{1}{n} = \overline{x}.$$

(c). Although a discrete distribution we do have

$$P[x_i^* < x_{((n+1)/2)}] = \frac{n-1}{2n} = P[x_i^* > x_{((n+1)/2)}].$$

4.9.3 (a). The median $\xi_{0.5}$ solves

$$1 - e^{-\xi_{0.5}/\beta} = \frac{1}{2},$$

or $\xi_{0.5} = \beta \log 2$.

(b). The following R-function produces the bootstrap percentile confidence interval:

```
percentcimed<-function(x,b,alpha){
#
theta<-median(x)
thetastar<-rep(0,b)
n<-length(x)
for(i in 1:b){xstar<-sample(x,n,replace=T)
             thetastar[i]<-median(xstar)
             }
thetastar<-sort(thetastar)
pick<-round((alpha/2)*(b+1))
lower<-thetastar[pick]
upper<-thetastar[b-pick+1]
list(theta=theta,lower=lower,upper=upper,thetasta=thetastar)
#list(theta=theta,lower=lower,upper=upper)
}
```

Below is the output of a 90% confidence interval based on 1000 bootstraps. Note the the confidence interval did trap the true value in this case.

```
$theta
[1] 67.6

$lower
[1] 32.25

$upper
[1] 126.8
 truemed = 100*log2
> 100*log(2)
[1] 69.31472
```

4.9.5 The following R-code gives a function which returns the confidence interval defined in expression (4.9.13).

```
prob595bs<-function(x,b,alpha){
#
```

```
theta<-mean(x)
stan <- var(x)^.5
n = length(x)
teestar<-rep(0,b)
n<-length(x)
for(i in 1:b){xstar<-sample(x,n,replace=T)
              teestar[i] = (mean(xstar) - theta)/(var(xstar)^.5/sqrt(n))
             }
teestar<-sort(teestar)
pick<-round((alpha/2)*(b+1))
lower0<-teestar[pick]
upper0<-teestar[b-pick+1]
lower = theta - upper0*(stan/sqrt(n))
upper = theta - lower0*(stan/sqrt(n))
list(theta=theta,lower=lower,upper=upper,teestar=teestar)
#list(theta=theta,lower=lower,upper=upper)
}
```

The results for data in Example 4.9.3 based on 1000 bootstraps are:

```
> temp=prob595bs(x,1000,.10)
> temp$theta
[1] 90.59
> temp$lower
[1] 63.67547
> temp$upper
[1] 129.4924
```

4.9.7 Here are the results from a Minitab run on the data of Example 4.9.2:

```
TWOSAMPLE T FOR C2 VS C1
      N      MEAN     STDEV   SE MEAN
 C2  15     117.7      18.6       4.8
 C1  15     111.1      20.4       5.3

 95 PCT CI FOR MU C2 - MU C1: ( -8.0,  21.2)

 TTEST MU C2 = MU C1 (VS GT): T= 0.93  P=0.18  DF=  28

 POOLED STDEV =       19.5
```

where the data are in $C1$ and $C2$.

4.9.10

$$\begin{aligned} E[z^*] &= \sum_{i=1}^{n}(x_i - \overline{x} + \mu_0)\frac{1}{n} = \mu_0. \\ \text{Var}[z^*] &= \sum_{i=1}^{n}(x_i - \overline{x} + \mu_0 - \mu_0)\frac{1}{n} = \mu_0. \\ &= \sum_{i=1}^{n}(x_i - \overline{x})^2\frac{1}{n}. \end{aligned}$$

4.9.13 The paired test is a one sample test based on the paired differences. So the bootstrap test discussed on page 280 can be used. In this case a bootstrap sample consists of a sample drawn with replacement from the observations $d_i = (x_i - y_i) - (\overline{x} - \overline{y})$, $i = 1, 2, \ldots, n$. The following R function performs this bootstrap:

```
pairsbs2=function(x,y,nb){
d = x-y - mean(x)+mean(y)
n=length(d)
ts = mean(x) - mean(y)
tsstar = rep(0,nb)
pval = 0
for(i in 1:nb){dstar = sample(d,n,replace=T)
               tsstar[i] = mean(dstar)
               if(tsstar[i]>= ts) pval = pval + 1}
pval = pval/nb
list(teststat=ts,pval=pval,tsstar=tsstar)
}
```

Here are results of a run based on 10,000 bootstraps:

```
>  temp=pairsbs2(x,y,10000)
>  temp$teststat
[1] 2.62
>  temp$pval
[1] 0.0062
```

4.10.1 $F(Y_n) - F(Y_1)$ is distributed like $V = F(Y_{n-1})$. So

$$\begin{aligned} P(V \geq 0.5) &= \int_{0.5}^{1} n(n-1)v^{n-2}(1-v)\,dv \\ &= 1 - n(0.5)^{n-1} + (n-1)(0.5)^n \geq 0.95. \end{aligned}$$

That is, $0.05 \geq n(0.5)^{n-1} - (n-1)(0.5)^n = (0.5)^n(n+1)$ means that (by trial) $n = 9$ is that smallest value.

4.10.3 $F(Y_{45}) - F(Y_4)$ is distributed as $V = F(Y_{41})$. So

$$\begin{aligned}\gamma &= \int_{0.75}^{1} \frac{48!}{40!7!} v^{40}(1-v)^7\, dv = \sum_{w=41}^{48} \binom{48}{w} (0.75)^w (0.25)^{48-w} \\ &\approx \Phi\left(\frac{48.5-36}{3}\right) - \Phi\left(\frac{40.5-36}{3}\right).\end{aligned}$$

4.10.4 (a). $1 - F(Y_j)$ is distributed as $V = F(Y_{n+1-j})$ which is beta ($\alpha = n+1-j$, $\beta = j$).

(b). There are $n-j+i-1$ coverages so it is distributed as $V = F(Y_{n-j+i-1})$ which is beta ($\alpha = n-j+i-1$, $\beta = j-i+2$),

4.10.5 These variables are distributed like $U_1 = F(Y_2) = W_2$, $U_2 = F(Y_6) - F(Y_2) = W_6 - W_2$, where

$$g(w_2, w_6) = \frac{10!}{1!\,3!\,4!} w_2 (w_6 - w_2)^3 (1-w_6)^4, \;\; 0 < w_2 < w_6 < 1.$$

Here $w_2 = u_1$, $w_6 = u_1 + u_2$ with $J = 1$; so the joint pdf of U_1 and U_2 is

$$h(u_1, u_2) = \frac{10!}{1!\,3!\,4!} u_1 u_2^3 (1 - u_1 - u_2)^4, \;\; 0 < u_1,\, 0 < u_2 \text{ and } u_1 + u_2 < 1,$$

which is Dirichlet ($\alpha_1 = 2$, $\alpha_2 = 4$, $\alpha_3 = 5$).

Chapter 5

Consistency and Limiting Distributions

5.1.3 For all $\epsilon > 0$,

$$P(|W_n - \mu| \geq \epsilon) \leq \frac{b}{n^p \epsilon^2} \to 0,$$

as $n \to \infty$.

5.1.5 Note that $Y_n \geq t \Leftrightarrow X_i \geq t,$ for all $i = 1, 2, \ldots, n$. Hence, for $t > \theta$, the fact that $X_1, X_2, \ldots, X_n$ are iid implies

$$\begin{array}{rcl} P(|Y_n - \theta| \leq \epsilon) & = & P(Y \leq \epsilon + \theta) = 1 - e^{-n(\epsilon + \theta - \theta)} \\ 1 - e^{-n\epsilon} \to 1, & & \end{array}$$

as $n \to \infty$.

5.1.7 The density of Y_n is $f(y) = n\exp\{-n(y-\theta)\}$ for $y > \theta$. Hence,

$$\begin{aligned} E[Y_n] &= n \int_\theta^\infty y e^{-n(y-\theta)}\, dy \\ &= \int_0^\infty \left(\frac{z}{n} + \theta\right) e^{-z}\, dz \\ &= \frac{1}{n} \int_0^\infty z^{2-1} e^{-z}\, dz + \theta \int_0^\infty e^{-z}\, dz = \frac{1}{n} + \theta, \end{aligned}$$

where the integral on the second line results from the substitution $z = n(y - \theta)$. Based on this result $Y_n - \frac{1}{n}$ is an unbiased estimate of θ.

5.2.2

$$\begin{aligned} g_1(y_1) &= ne^{-n(y_1-\theta)}, \quad 0<y_1<\infty \\ z &= n(y_1-\theta) \ \text{ and } \ \frac{dy_1}{dz}=\frac{1}{n}, \\ h_n(z) &= e^{-z} \ \text{ and } \ H_n(z)=1-e^{-z}, 0<z<\infty \\ \lim_{n\to\infty} H_n(z) &= \begin{cases} 1-e^{-z} & 0<z<\infty \\ 0 & \text{elsewhere.} \end{cases} \end{aligned}$$

5.2.4

$$\begin{aligned} g_2(y_2) &= n(n-1)F(y_2)[1-F(y_2)]^{n-2}f(y_2), -\infty<y_2<\infty \\ w &= nF(y_2) \Rightarrow \frac{dy_2}{dw}=\frac{1}{nf(y_2)}. \\ h(w) &= \frac{n-1}{n}w(1-w/n)^{n-2}, 0<w<n \\ \lim_{n\to\infty} H_n(w) &= \lim_{n\to\infty}\int_0^w \frac{n-1}{n}z(1-z/n)^{n-2}\,dz \\ &= \int_0^w ze^{-z}\,dz, \end{aligned}$$

which is a $\Gamma(2,1)$ cdf.

5.2.5

$$\begin{aligned} F_n(y) &= \begin{cases} 0 & y<n \\ 1 & n\le y. \end{cases} \\ \lim_{n\to\infty} F_n(y) &= 0, \quad -\infty<y<\infty. \end{aligned}$$

There is no cdf which equals this limit at every point of continuity.

5.2.7 $\lim_{n\to\infty} E(e^{tX_n/n}) = \lim_{n\to\infty}(1-\beta t/n)^{-n} = e^{\beta t}$, which is the mgf of a degenerate distribution at β.

5.2.9

$$P\left(\frac{40-50}{10}<\frac{X-50}{10}<\frac{60-50}{10}\right)\approx\Phi(1)-\Phi(-1).$$

5.2.10

(a) $$\sum_{x=56}^{60}\binom{60}{x}(0.95)^x(0.05)^{60-x}.$$

(b) $Y=60-X$ is $b(n=60,p=0.05)$.
$np=3$ and $P(Y=0,1,2,3,4)\approx 0.815$, from the Poisson Tables.

5.2.11

$$\begin{aligned}
\lim_{n\to\infty} E[e^{t(Z_n-n)/\sqrt{n}}] &= \lim_{n\to\infty}\{e^{-tsqrtn}\exp[n(e^{t/\sqrt{n}}-1)]\} \\
&= \lim_{n\to\infty}\left\{\exp\left[-t/\sqrt{n}+n\left(t/\sqrt{n}+\frac{t^2}{2n}+\frac{t^3}{6n^{3/2}}-\cdots\right)\right]\right\}. \\
&= \lim_{n\to\infty}\left[\exp\left(\frac{t^2}{2}+\frac{t^3}{6n^{1/2}}\cdots\right)\right] = \exp(t^2/2),
\end{aligned}$$

which is the mgf of $N(0,1)$.

5.2.18 Note that $Y_n \le t \Leftrightarrow X_i \le t,$ for all $i = 1, 2, \ldots, n$. Hence, for $0 < t$, the fact that $X_1, X_2, \ldots, X_n$ are iid implies

$$\begin{aligned}
P(Y_n \le t + \log n) &= (P(X_1 \le t + \log n))^n \\
&= \left[1 - e^{-(t+\log n)}\right]^n \\
&= \left[1 - e^{-t}\frac{1}{n}\right]^n \to \exp\{-e^{-t}\},
\end{aligned}$$

as $n \to \infty$.

5.2.20 Using Stirling's formula,

$$\begin{aligned}
\frac{\Gamma((n+1)/2)}{\sqrt{n/2}\Gamma(n/2)} &\approx \frac{\left(\frac{n}{2}-\frac{1}{2}\right)^{n/2} e^{1/2}}{\left(\frac{n}{2}\right)^{1/2}\left(\frac{n}{2}-1\right)^{(n/2)-(1/2)} e} \\
&= \left\{\left(\frac{n-1}{n-2}\right)^{n/2}\right\} e^{-1/2}\left\{\left(\frac{n-2}{n}\right)^{1/2}\right\}.
\end{aligned}$$

The last factor in braces goes to 1, as $n \to \infty$. The first factor in braces can be expressed as

$$\left\{\left[1+\frac{1}{n-2}\right]^{n-2+2}\right\}^{1/2},$$

which converges to $e^{1/2}$, as $n \to \infty$.

5.3.2

$$\operatorname{var}(\overline{X}) = (2)(4^2)/128 = 1/4 \text{ and } E(\overline{X}) = (2)(4) = 8;$$

$$P\left(\frac{7-8}{1/2} < \frac{\overline{X}-8}{1/2} < \frac{9-8}{1/2}\right) \approx \Phi(2) - \Phi(-2).$$

5.3.3

$$P(21.5 < Y < 28.5) \approx \Phi\left(\frac{28.5-24}{4}\right) - \Phi\left(\frac{21.5-24}{4}\right),$$

because $E(Y) = 24$ and $\operatorname{var}(Y) = 16$.

5.3.5
$$E(X) = 3.5 \text{ and } \mathrm{var}(X) = 35/12 \;\Rightarrow\; E(Y) = 42 \text{ and } \mathrm{var}(Y) = 35.$$

Hence,

$$P(35.5 < Y < 48.5) \approx \Phi\left(\frac{48.5-42}{\sqrt{35}}\right) - \Phi\left(\frac{35.5-42}{\sqrt{35}}\right).$$

5.3.7 $\Phi\left(\frac{50.5-50}{5}\right) - \Phi\left(\frac{49.5-50}{5}\right)$.

5.3.9 Here Y is $b(72, p)$, where $p = \int_1^3 (1/x^2)\,dx = 2/3$. So,

$$P(Y > 50) \approx 1 - \Phi\left(\frac{50.5-48}{4}\right).$$

5.3.12

$$\begin{aligned} u(\overline{X}) &\approx v(\overline{X}) = u(\mu) + u'(\mu)(\overline{X}), \\ \mathrm{var}[v(\overline{X})] &= [u'(\mu)]^2(\mu/n) = c, \\ u'(\mu) &= c_1/\sqrt{\mu}, \text{ a solution is } u(\mu) = c_2\sqrt{\mu}. \end{aligned}$$

Taking $c_2 = 1$, we have $u(\overline{X}) = \sqrt{\overline{X}}$.

5.4.1 Assume that $\mathbf{X}_n \xrightarrow{D} N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Consider the sequence of random variables $\{\mathbf{a}'X_n\}$. Let $t \in R$. Then by the assumption,

$$E\left[e^{t(\mathbf{a}'X_n)}\right] = E\left[e^{(t\mathbf{a})'X_n}\right] \to \exp\{t\mathbf{a}'\boldsymbol{\mu} + \frac{1}{2}t^2(\mathbf{a}'\boldsymbol{\Sigma}\mathbf{a})\}.$$

Hence, the sequence of random variables $\{\mathbf{a}'X_n\}$ converges in distribution to the $N(\mathbf{a}'\boldsymbol{\mu}, \mathbf{a}'\boldsymbol{\Sigma}\mathbf{a})$ distribution. The converse is similar.

5.4.2 Immediate by Theorem 4.5.1.

5.4.5 Use mgfs. Then the result follows because the function $\exp\{\mathbf{t}'\boldsymbol{\mu} + \frac{1}{2}\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t}\}$ is continuous in $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$.

Chapter 6

Maximum Likelihood Methods

6.1.2 (b). $L = e^{-\sum(x_i-\theta)}$, provided $\theta \leq x_i$; otherwise $L = 0$.
$\log L = -\sum(x_i - \theta)$ and $D_\theta(\log L) = n > 0$. That is, $\log L$ is an increasing function of θ provided $\theta \leq x_i$, $i = 1, 2, \ldots, n$. Thus $\hat{\theta} = \min(X_i)$.

6.1.4 The likelihood simplifies to

$$L(\theta) = \frac{2^n}{\theta^{2n}} \prod_{i=1}^{n} x_i I(0 < x_i \leq \theta).$$

But $x_i \leq \theta$ for all $i = 1, \ldots, n$ if and only if $\max_{1\leq i\leq n} x_i \leq \theta$. hence, the likelihood can be written as

$$L(\theta) = \frac{2^n}{\theta^{2n}} I(0 < \max_{1\leq i\leq n} x_i \leq \theta) \prod_{i=1}^{n} x_i.$$

Part(a) It is clear from the form of the likelihood that the maximum of $L(\theta)$ occurs at the smallest value in the range of θ; hence, the mle of θ is $Y = \max_{1\leq i\leq n} X_i$.

Part(b) The cdf of X_i is $F_X(x) = x^2/\theta^2$. Hence, the cdf and pdf of Y are, respectively,

$$\begin{aligned} F_Y(y) &= \frac{y^{2n}}{\theta^{2n}}, \quad 0 < y \leq \theta \\ f_Y(y) &= \frac{2ny^{2n-1}}{\theta^{2n}}, \quad 0 < y \leq \theta. \end{aligned}$$

So

$$E(Y) = \int_0^\theta \frac{2ny^{2n}}{\theta^{2n}}\, dy = \frac{2n}{2n+1}\theta.$$

So $c = (2n+1)/(2n)$.

Part(c) The median is the value of x which solves $x^2/\theta^2 = (1/2)$, which is $\theta/\sqrt{2}$. The mle of the median is therefore $Y/\sqrt{2}$. Note that an unbiased estimate of the median is $[(2n+1)Y]/[2n\sqrt{2}]$.

6.1.5 Since $\hat{\theta} = \overline{X}$ and $P(X \le 2) = 1 - e^{-2/\theta}$, then

$$P(\widehat{X \le 2}) = 1 - e^{-2/\overline{X}}.$$

6.1.6

$$\begin{aligned}
\hat{p} &= \frac{(6)(0)+(10)(1)+(14)(2)+(13)(3)+(6)(4)+(1)(5)}{250} \\
&= \frac{106}{250} = \frac{53}{125} \\
P(\widehat{X \ge 3}) &= \sum_{x=3}^{5} \binom{5}{x} (\hat{p})^x (1-\hat{p})^{5-x}, \quad \text{where } \hat{p} = 53/125.
\end{aligned}$$

6.1.8 The mle is $\overline{X}$. In terms of the summary data

$$\overline{x} = \frac{7(0)+14(1)+12(2)+13(3)+6(4)+3(5)}{7+14+12+13+6+3} = 2.109.$$

6.1.10 The log of the likelihood function is

$$l(\theta) = K_1 - K_2 Q(\theta),$$

where K_1 and $K_2 > 0$ are constants and $Q(\theta) = \sum_{i=1}^{n}(x_i - \theta)^2$. To maximize $l(\theta)$, we must minimize $Q(\theta)$. In the unrestricted case $Q(\theta)$ is minimized at $\overline{x}$. In the restricted case, $\theta > 0$. Hence, if $\overline{x} > 0$ then the minimum occurs at $\overline{x}$. If $\overline{x} \le 0$ then, because $Q(\theta)$ is a quadratic whose leading coefficient is positive, the minimum occurs at 0.

6.2.2 $\frac{\partial \log f(x;\theta)}{\partial \theta} = \frac{-1}{\theta}$; $nE\left[\left(\frac{-1}{\theta}\right)^2\right] = \frac{n}{\theta^2}$.
Also

$$\begin{aligned}
E(Y_n) &= \int_0^{\theta} (ny^n/\theta^n)\, dy = \frac{n}{n+1}\theta, \\
\text{Var}(Y_n) &= \frac{n}{n+2}\theta^2 - \left(\frac{n}{n+1}\right)^2 \theta^2 = \frac{n\theta^2}{(n+2)(n+1)^2}.
\end{aligned}$$

6.2.3

$$\begin{aligned}
\log f(x;\theta) &= -\log \pi - \log[1 + (x-\theta)^2], \\
\frac{\partial \log f(x;\theta)}{\partial \theta} &= \frac{2(x-\theta)}{1+(x-\theta)^2}, \\
I(\theta) &= \int_{-\infty}^{\infty} \frac{4(x-\theta)^2}{\pi[1+(x-\theta)^2]^3}\, dx = \int_{-\pi/2}^{\pi/2} \frac{4\tan^2 z \sec^2 z}{\pi[1+\tan^2 z]^3}\, dz
\end{aligned}$$

by letting $x-\theta=\tan z$. Thus

$$\begin{aligned} I(\theta) &= \left(\frac{4}{\pi}\right)\int_{-\pi/2}^{\pi/2}\cos^2 z\sin^2 z\,dz = \frac{4}{\pi}\int_{-\pi/2}^{\pi/2}\left(\frac{1+\cos 2z}{2}\right)\left(\frac{1-\cos 2z}{2}\right)dz \\ &= \frac{1}{\pi}\int_{-\pi/2}^{\pi/2}\left[1-\left(\frac{1+\cos 4z}{2}\right)\right]dz = \frac{1}{2}-\int_{-\pi/2}^{\pi/2}\frac{\cos 4z}{2}\,dz = \frac{1}{2}. \end{aligned}$$

Accordingly, $1/nI(\theta)=2/n$.

6.2.6 The variance of $\overline{X}$ is σ^2/n, where σ^2 is the variance of a contaminated normal distribution; see expression (3.4.13) on page 167. The asymptotic variance of the sample median is $1/4f^2(0)n$. Here,

$$f(0)=\phi(0)(1-\epsilon)+\phi(0)\frac{\epsilon}{\sigma_c},$$

from which the result follows.

6.2.8 (a).

$$\begin{aligned} \log f(x;\theta) &= -\frac{1}{2}\log(2\pi)-\frac{1}{2}\log\theta-\frac{x^2}{2\theta}, \\ \frac{\partial\log f(x;\theta)}{\partial\theta} &= -\frac{1}{2\theta}+\frac{x^2}{2\theta^2}, \\ \frac{\partial^2\log f(x;\theta)}{\partial\theta^2} &= \frac{1}{2\theta^2}-\frac{x^2}{\theta^3}, \\ -nE\left[\frac{\partial^2\log f(x;\theta)}{\partial\theta^2}\right] &= \frac{-n}{2\theta^2}+\frac{n}{\theta^2}=\frac{n}{2\theta^2}=nI(\theta). \end{aligned}$$

(b). Here $\hat{\theta}=\sum X_i^2/n$. Since $\sum X_i^2/\theta$ is $\chi^2(n)$, we have

$$\mathrm{Var}(\hat{\theta})=\frac{\theta^2}{n^2}\,\mathrm{Var}\left(\frac{\sum X_i^2}{\theta}\right)=\frac{2\theta^2}{n}=\frac{1}{nI(\theta)}.$$

6.2.10 Note that

$$\begin{aligned} E[|X_1|] &= 2\int_0^\infty \frac{x}{\sqrt{2\pi}}\exp\left\{-\frac{1}{2}\frac{x^2}{\theta}\right\}dx \\ &= 2\sqrt{\theta}\int_0^\infty\frac{1}{\sqrt{2\pi}}\exp\{-z\}\,dz=\sqrt{\frac{2}{\pi}}\sqrt{\theta}. \end{aligned}$$

So $c=\sqrt{\pi/2}/n$. Hence, $Y=n^{-1}\sum_{i=1}^n\sqrt{\frac{2}{\pi}}|X_i|$. Note that,

$$\begin{aligned} V\left[\sqrt{\frac{\pi}{2}}|X_1|\right] &= \frac{\pi}{2}\{E(X_1^2)-[E(|X_1|)]^2\} \\ &= \frac{\pi}{2}\left[\theta\left(1-\frac{2}{\pi}\right)\right]=\theta\left[\frac{\pi}{2}-1\right]. \end{aligned}$$

By independence,

$$V(Y) = \theta\left[\frac{\pi}{2} - 1\right]\frac{1}{n}. \tag{6.0.1}$$

To finish, we need the efficiency of the parameter $\sqrt{\theta}$. For convenience, let $\beta = \sqrt{\theta}$. Then

$$\log f(x;\beta) = -\log\sqrt{2\pi} - \log\beta - \frac{1}{2}\frac{x^2}{\beta^2}.$$

The second partial of this expression is,

$$\frac{\partial^2 \log f(x;\beta)}{\partial\beta^2} = \frac{1}{\beta^2} - 3\frac{x^2}{\beta^4}.$$

Hence, using $\sqrt{\theta} = \beta$,

$$I(\sqrt{\theta}) = -E\left[\frac{1}{\theta} - 3\frac{X^2}{\theta^2}\right] = \frac{2}{\theta}. \tag{6.0.2}$$

Thus by (6.0.1) and (6.0.2) we have

$$e(Y) = \frac{\theta/2n}{\theta[(\pi/2) - 1]/n} = \frac{1}{\pi - 2}.$$

6.2.14 For Part (a), recall that $(n-1)S^2/\theta$ has $\chi^2(n-1)$ distribution. Hence,

$$V\left[\frac{(n-1)S^2}{\theta}\right] = 2(n-1).$$

So $V(S^2) = 2\theta^2/(n-1)$. Also, by Problem (6.28), $I(\theta) = (2\theta^2)^{-1}$. Thus, the efficiency of S^2 is $(n-1)/n$.

6.3.1 Note that under θ, the random variable $(\theta_0/\theta)(2/\theta_0)\sum_{i=1}^n X_i$ has a $\chi^2(2n)$ distribution. Therefore, the power function is

$$\gamma(\theta) = P\left[T \le \frac{\theta_0}{\theta}\chi^2_{1-\alpha/2}(2n)\right] + P\left[T \ge \frac{\theta_0}{\theta}\chi^2_{\alpha/2}(2n)\right],$$

where T has a $\chi^2(2n)$ distribution.

6.3.3 The decision rule (6.3.6) is equivalent to the decision rule

$$\text{Reject } H_0 \text{ if } |z| \ge z_{\alpha/2},$$

where $z = (\overline{x} - \theta_0)/(\sigma/\sqrt{n})$. The power function is

$$\begin{aligned}
\gamma(\theta) &= P_\theta\left[\left|\frac{\overline{X} - \theta_0}{\sigma/\sqrt{n}}\right| \ge z_{\alpha/2}\right] \\
&= P_\theta\left[\left|\frac{\overline{X} - \theta}{\sigma/\sqrt{n}}\right| \le -z_{\alpha/2} + \frac{\sqrt{n}(\theta_0 - \theta)}{\sigma}\right] \\
&\quad + P_\theta\left[\left|\frac{\overline{X} - \theta}{\sigma/\sqrt{n}}\right| \ge z_{\alpha/2} + \frac{\sqrt{n}(\theta_0 - \theta)}{\sigma}\right] \\
&= \Phi\left[-z_{\alpha/2} + \frac{\sqrt{n}(\theta_0 - \theta)}{\sigma}\right] + 1 - \Phi\left[z_{\alpha/2} + \frac{\sqrt{n}(\theta_0 - \theta)}{\sigma}\right].
\end{aligned}$$

6.3.6 Let $\chi^2_{1-\alpha/2}$ and $\chi^2_{\alpha/2}$ be the lower and upper $\alpha/2$ critical values of a χ^2-distribution with n degrees of freedom. Then the power curve for a level α test is given by

$$\begin{aligned}\gamma(\theta) &= P_\theta\left[W \le \chi^2_{1-\alpha/2}\right] + P_\theta\left[W \ge \chi^2_{\alpha/2}\right] \\ &= P\left[\chi^2(n) \le \frac{\theta_0}{\theta}\chi^2_{1-\alpha/2}\right] + P_\theta\left[\chi^2(n) \ge \frac{\theta_0}{\theta}\chi^2_{\alpha/2}\right],\end{aligned}$$

where $\chi^2(n)$ represents a random variable with a χ^2-distribution with n degrees of freedom. The following R function computes this power function at the specified value `theta`. The default values of the other arguments are set at values given in the exercise. Using this, it is easy to obtain a plot of the power curve.

```
pcchitst = function(n=10,alpha=.05,theta0=1,theta){
     alp2 = alpha/2
     l = (theta0/theta)*qchisq(alp2,n)
     u = (theta0/theta)*qchisq(1-alp2,n)
     pcchitst = pchisq(l,n) + 1 - pchisq(u,n)
     pcchitst
}
```

6.3.8 Part (a). Under Ω, the mle is $\overline{x}$. After simplification, the likelihood ratio test is

$$\Lambda = e^{-\theta_0}e^{\overline{x}-n\overline{x}\log(\overline{x}/\theta_0)}.$$

Treating Λ as a function of $\overline{x}$, upon differentiating it twice we see that the function has a positive real critical value which is a maximum. Hence, the likelihood ratio test is equivalent to rejecting H_0, if $Y \le c_1$ or $Y \ge c_2$ where $Y = n\overline{X}$. Under H_0, Y has a Poisson distribution with mean $n\theta_0$. The significance level of the test is 0.056 for the situation described in Part (b).

6.3.11 Note that under $\theta = 2$, the distribution is $N(0, 2^{-1})$. Under $\theta = 1$, the distribution is the standard Laplace. Some simplification leads to

$$\Lambda = K\exp\left\{\sum_{i=1}^n (x_i^2 - |x_i|)\right\},$$

where K is a constant.

6.3.15 The likelihood function can be expressed as

$$L(\theta) = \theta^{n\overline{x}}(1-\theta)^{n-n\overline{x}}.$$

To get the information, note that

$$\log p(x;\theta) = x\log\theta + (1-x)\log(1-\theta).$$

Upon taking the first two partial derivatives with respect to θ, we obtain the information

$$I(\theta) = E\left[\frac{X}{\theta^2}\right] - E\left[\frac{1-X}{1-\theta^2}\right] = \frac{1}{\theta(1-\theta)}.$$

(a). Under Ω, the mle is $\overline{x}$. Hence, the likelihood ratio test statistic is

$$\Lambda = \left(\frac{1}{3\overline{x}}\right)^{n\overline{x}}\left(\frac{2/3}{1-\overline{x}}\right)^{n-n\overline{x}}.$$

(b). Wald's test statistic is

$$\chi_W^2 = \left[\frac{\overline{x}-(1/3)}{\sqrt{\overline{x}(1-\overline{x})/n}}\right]^2.$$

(c). For the scores test,

$$l'(\theta_0) = \sum_{i=1}^{n}\left[\frac{x_i}{\theta} - \frac{1-x_i}{1-\theta}\right] = \frac{n(\overline{x}-\theta)}{\theta(1-\theta)}.$$

Hence, the scores test statistic is

$$\chi_R^2 = \left\{\frac{n(\overline{x}-\theta_0)}{\theta_0(1-\theta_0)}\Big/\sqrt{\frac{n}{\theta_0(1-\theta_0)}}\right\}^2 = \left\{\frac{\sqrt{n}(\overline{x}-\theta_0)}{\sqrt{\theta_0(1-\theta_0)}}\right\}^2.$$

6.3.18 Recall the the pdf of the Y_n is

$$f_{Y_n}(y;\theta) = \begin{cases} \frac{n}{\theta}\left(\frac{y}{\theta}\right)^{n-1} & 0 < y < \theta \\ 0 & \text{elsewhere.} \end{cases} \tag{6.0.3}$$

(a). The numerator of the likelihood ratio test is $(1/\theta_0)^n$, if $0 < y_n \le \theta_0$ and is 0 if $y_n > \theta_0$. The mle under Ω is y_n. So, the denominator of the likelihood ratio test is $(1/y_n)^n$. Hence, the result for Λ.

(b). Let $T_n = -2\log\Lambda = -2n\log(Y_n/\theta_0)$. Then the inverse transformation is $y_n = \theta_0\exp\{-t_n/2n\}$ with Jacobian $(-\theta_0/2n)\exp\{-t/2n\}$. Based on (6.0.3) the pdf of T_n is

$$\begin{aligned} f_{T_n}(t) &= \frac{n}{\theta_0}\left\{\frac{\theta_0\exp\{-t/2n\}}{\theta_0}\right\}^{n-1}\frac{\theta_0}{2n}\exp\{-t/2n\} \\ &= \frac{1}{2}\exp\{-t/2\}, \end{aligned}$$

which is the pdf of $\chi^2(2)$ distribution.

6.4.2 Note in general that the log of the likelihood is

$$l(\theta_1,\theta_2,\theta_3,\theta_4) = K - \frac{n}{2}\log\theta_3 - \frac{m}{2}\log\theta_4 - \frac{1}{2\theta_3}\sum_{i=1}^{n}(x_i-\theta_1)^2 - \frac{1}{2\theta_4}\sum_{i=1}^{m}(y_i-\theta_2)^2, \tag{6.0.4}$$

where K is a constant.

(a). For this part, expression (6.0.4) becomes

$$l(\theta_1,\theta_2,\theta_3) = K - \frac{n+m}{2}\log\theta_3 - \frac{1}{2\theta_3}\sum_{i=1}^{n}(x_i-\theta_1)^2 - \frac{1}{2\theta_3}\sum_{i=1}^{m}(y_i-\theta_2)^2.$$

If we take the partial with respect to θ_1 and set the resulting expression to 0, then we see immediately that the mle of θ_1 id $\overline{x}$. Likewise, the mle of θ_2 is $\overline{y}$. Substituting these mles into the above expression and differentiating with respect to θ_3, yields the mle of θ_3:

$$\widehat{\theta_3} = \frac{1}{n+m}\left[\sum_{i=1}^{n}(x_i-\overline{x})^2 + \sum_{i=1}^{m}(y_i-\overline{y})^2\right].$$

(b). Under the assumptions of this part, we have one (combined) sample from a $N(\theta_1,\theta_3)$ distribution. Hence, based on Example 6.4.1 the mles are

$$\begin{aligned}\widehat{\theta_1} &= \frac{1}{n+m}\left[\sum_{i=1}^{n}x_i + \sum_{i=1}^{m}y_i\right]\\ \widehat{\theta_3} &= \frac{1}{n+m}\left[\sum_{i=1}^{n}(x_i-\widehat{\theta_1})^2 + \sum_{i=1}^{m}(y_i-\widehat{\theta_1})^2\right].\end{aligned}$$

6.4.5 $L = \left(\frac{1}{2\rho}\right)^n$, provided $\theta-\rho \le y_1 \le y_n \le \theta+\rho$. To maximize L make ρ as small as possible which is accomplished by setting

$$\hat{\theta}-\hat{\rho} = Y_1 \quad\text{and}\quad \hat{\theta}+\hat{\rho} = Y_n.$$

So

$$\hat{\theta} = \frac{Y_1+Y_n}{2} \quad\text{and}\quad \hat{\rho} = \frac{Y_n-Y_1}{2}.$$

Thus

$$E\left[\frac{(n+1)Y_n}{n}\right] = \theta, \quad \operatorname{Var}\left[\frac{(n+1)Y_n}{n}\right] = \frac{\theta^2}{n(n+2)}.$$

However, we have that

$$\frac{\theta^2}{n(n+2)} < \frac{\theta^2}{n} = \frac{1}{n\,E\left\{\left[\frac{\partial\log f(X;\theta)}{\partial\theta}\right]^2\right\}},$$

which seems like a contradiction to the Rao-Cramér inequality until we recognize that this is not a regular case.

6.4.8 Because $b > 0$,

$$P(X_i \le t) = P\left(e_i \le \frac{t-a}{b}\right),$$

from which the result follows.

6.4.10 Write I_{12} as

$$I_{12} = \frac{1}{b^2}\int_{-\infty}^{\infty}\{z\}\left\{\frac{f'(z)}{f(z)}\right\}^2\{f(z)\}\,dz.$$

Note that the function in the first set of braces is odd while the last two functions are even (the third because of the assumed symmetry). Thus their product is an odd function and hence the integral of it from $-\infty$ to ∞ is 0.

6.5.4 $\lambda = \dfrac{\left\{\frac{1}{(2\pi)[(\sum x_i^2+\sum y_i^2)/(n+m)]}\right\}^{(n+m)/2}}{\left[\frac{1}{(2\pi)(\sum x_i^2/n)}\right]^{n/2}\left[\frac{1}{(2\pi)(\sum y_i^2/m)}\right]^{m/2}} \leq k,$

which is equivalent to $F \leq c_1$ or $F \geq c_2$, where $F = \frac{\sum X_i^2/n}{\sum Y_i^2/m}$ has an $F(r_1 = n, r_2 = m)$ distribution when $\theta_1 = \theta_2$.

6.5.6 Note $\hat{\theta}_i = \max\{-\text{1st order statistic}, n\text{th order statistic}\}$, where $n = n_1 = n_2$. Hence, in a notation that seems clear, we have

$$\lambda = \frac{1/[2\max(\hat{\theta}_X,\hat{\theta}_Y)]^{2n}}{[1/(2\hat{\theta}_X)^n][1/(2\hat{\theta}_Y)^n]} = \left[\frac{\min(\hat{\theta}_X,\hat{\theta}_Y)}{\max(\hat{\theta}_X,\hat{\theta}_Y)}\right]^n.$$

If $U = \min(\hat{\theta}_X,\hat{\theta}_Y)$and $V = \max(\hat{\theta}_X,\hat{\theta}_Y)$, the joint pdf is

$$g(u,v) = 2n^2u^{n-1}v^{n-1}/\theta^{2n},\ \ 0 < u < v < \theta.$$

So the distribution function of λ is

$$\begin{aligned} H(z) &= P(U \leq z^{1/n}V),\ \ 0 \leq z \leq 1, \\ &= \int_0^{\theta}\int_0^{z^{1/n}v} g(u,v)\,du\,dv \\ &= \int_0^{\theta} 2nzv^{2n-1}/\theta^{2n}\,dv \\ &= z,\ \ 0 \leq z \leq 1, \end{aligned}$$

which is uniform (0, 1). Thus $-2\log\lambda$ is $\chi^2(2)$, where the degrees of freedom $= 2 = 2$(dimension of Ω – dimension of ω). Note that this is a nonregular case.

6.5.9 The likelihood ratio test statistic is

$$\Lambda = \frac{\hat{p}^{n_1\overline{x}+n_2\overline{y}}(1-\hat{p})^{n_1+n_2-(n_1\overline{x}+n_2\overline{y})}}{\overline{x}^{n_1\overline{x}}(1-\overline{x})^{n_1-n_1\overline{x}}\overline{y}^{n_2\overline{y}}(1-\overline{y})^{n_2-n_2\overline{y}}}.$$

After simplification, we have

$$\begin{aligned} -2\log\Lambda &= -2\left\{n_1\overline{x}\log\left(\frac{\hat{p}}{\overline{x}}\right) + n_2\overline{y}\log\left(\frac{\hat{p}}{\overline{y}}\right) + (n_1 - n_1\overline{x})\log\left(\frac{1-\hat{p}}{1-\overline{x}}\right)\right. \\ &\quad \left. +(n_2 - n_2\overline{y})\log\left(\frac{1-\hat{p}}{1-\overline{y}}\right)\right\}. \end{aligned}$$

6.5.11 Under H_0, $p_1 = p_2 = p$. Thus both $\overline{X}$ and $\overline{Y}$ are consistent estimators of p. Hence

$$\begin{aligned} \hat{p} &= \frac{n_1}{n}\overline{X} + \frac{n_2}{n}\overline{Y} \\ &\xrightarrow{P} \lambda_1 p + \lambda_2 p = p. \end{aligned}$$

6.5.13 Using the R code below, we obtained the values of the test statistics and their associated p-values:

(a) Wald test, (6.5.25), and p-value: -1.727113, 0.08414743.

(b) LRT test, (Exercise 6.5.9), and p-value: 2.993653, 0.0835914.

(c) Test of Exercise 6.5.11, and p-value: -1.725826, 0.0843787

R code:

```
p6513=function(x,y,n1,n2){
p1=x/n1
p2=y/n2
pc = (n1*p1+n2*p2)/(n1+n2)
zw = (p1-p2)/sqrt((p1*(1-p1)/n1) + (p2*(1-p2)/n2))
pzw = 2*min(pnorm(zw),1-pnorm(zw))
lrt = -2*(n1*p1*log(pc/p1) + n2*p2*log(pc/p2) + (n1-n1*p1)*log((1-pc)/(1-p1)
        + (n2-n2*p2)*log((1-pc)/(1-p2)))
plrt = 1-pchisq(lrt,1)
zpc = (p1-p2)/sqrt(pc*(1-pc)*((1/n1) + (1/n2)))
pzpc = 2*min(pnorm(zpc),1-pnorm(zpc))
list(zw=zw,pzw=pzw,lrt=lrt,plrt=plrt,zpc=zpc,pzpc=pzpc)
}
```

6.6.1 For Part (c), taking the partial of the log likelihood and setting the result to 0 yields

$$\frac{x_1}{2+\theta} - \frac{x_2+x_3}{1-\theta} + \frac{x_4}{\theta} = 0.$$

Upon simplification, we obtain the quadratic equation

$$n\theta^2 - (x_1 - 2x_2 - 2x_3 - x_4)\theta - 2x_4 = 0,$$

which has one positive and one negative root.

6.6.2

(a). The complete likelihood is given by

$$L^c = \frac{n!}{z_{11}!z_{12}!x_2!x_3!x_4!}\left(\frac{1}{2}\right)^{z_{11}}\left(\frac{\theta}{4}\right)^{z_{12}}\left(\frac{1-\theta}{4}\right)^{x_2+x_3}\left(\frac{\theta}{4}\right)^{x_4}.$$

(b). The conditional pmf $k(\mathbf{z}|\theta, \mathbf{x})$ is the ratio of L^c to L, which after simplification is

$$k(\mathbf{z}|\theta, \mathbf{x}) = \frac{x_1!}{z_{12}!(x_1 - z_{12})!}\left(\frac{\theta}{2+\theta}\right)^{z_{12}}\left(1 - \frac{\theta}{2+\theta}\right)^{x_1 - z_{12}};$$

i.e., a binomial distribution with parameters x_1 and $\theta/(2+\theta)$.

(c). Let $\hat{\theta}^{(0)}$ be an initial estimate of θ. For the E step, the conditional expectation of the log of L^c (ignoring constants) simplifies to

$$\begin{aligned} E\left[\log L^c(\theta|\mathbf{x}, \mathbf{z})|\hat{\theta}^{(0)}, \mathbf{x}\right] &= \log\left(\frac{\theta}{4}\right) E\left[Z_{12}|\hat{\theta}^{(0)}, \mathbf{x}\right] + (x_2 + x_3)\log\left(\frac{1-\theta}{4}\right) \\ &\quad + x_4 \log\left(\frac{\theta}{4}\right) \\ &= \log\left(\frac{\theta}{4}\right)\left[x_1 \frac{\hat{\theta}^{(0)}}{2 + \hat{\theta}^{(0)}}\right] + (x_2 + x_3)\log\left(\frac{1-\theta}{4}\right) \\ &\quad + x_4 \log\left(\frac{\theta}{4}\right) \end{aligned}$$

For the M step: Taking the partial of this last expression with respect to θ and setting the result to 0 yields the solution given in Part (d) of the text.

6.6.5 The observable likelihood is

$$L(\theta|\mathbf{x}) \propto \exp\left\{-\frac{1}{2}\sum_{i=1}^{n_1}(x_i - \theta)^2\right\},$$

while the complete likelihood is

$$L^c(\theta|\mathbf{x}, \mathbf{z}) \propto \exp\left\{-\frac{1}{2}\left[\sum_{i=1}^{n_1}(x_i - \theta)^2 + \sum_{i=1}^{n_2}(z_i - \theta)^2\right]\right\}.$$

The conditional pmf $k(\mathbf{z}|\theta, \mathbf{x})$ is the ratio of L^c to L which is easily seen to be the product of n_2 iid $N(\theta, 1)$ pdfs. Let $\hat{\theta}^{(0)}$ be an initial estimate of θ. For the E step, the conditional expectation of the log of L^c (ignoring constants) simplifies to

$$\begin{aligned} E[\log L^c||\hat{\theta}^{(0)}, \mathbf{x}] &= -\frac{1}{2}\left[\sum_{i=1}^{n_1}(x_i - \theta)^2 + \sum_{i=1}^{n_2} E_{\hat{\theta}^{(0)}}(z_i - \theta)^2\right] \\ &= -\frac{1}{2}\left[\sum_{i=1}^{n_1}(x_i - \theta)^2 + \sum_{i=1}^{n_2} E_{\hat{\theta}^{(0)}}[(z_i - \hat{\theta}^{(0)}) + (\hat{\theta}^{(0)} - \theta)]^2\right] \\ &= -\frac{1}{2}\left[\sum_{i=1}^{n_1}(x_i - \theta)^2 + \sum_{i=1}^{n_2}[1 + (\hat{\theta}^{(0)} - \theta)^2]\right] \\ &= -\frac{1}{2}\left[\sum_{i=1}^{n_1}(x_i - \theta)^2 + [n_2 + n_2(\hat{\theta}^{(0)} - \theta)^2]\right]. \end{aligned}$$

For the M step: Taking the partial of this last expression with respect to θ and setting the result to 0 yields the solution given in the text.

Chapter 7

Sufficiency

7.1.2 $\frac{\sum X_i^2}{\sigma^2}$ is $\chi^2(n)$. Hence $E\left[\frac{\sum X_i^2}{\sigma^2}\right] = n$, $\operatorname{var}\left(\frac{\sum X_i^2}{\sigma^2}\right) = 2n$. Accordingly,

$$E\left[\frac{\sum X_i^2}{n}\right] = \sigma^2 \text{ and } \operatorname{var}\left(\frac{\sum X_i^2}{n}\right) = \left(\frac{\sigma^2}{n}\right)^2 (2n) = \frac{2\sigma^4}{n}$$

7.1.3 This is a rather lengthy exercise. One observation that might help is illustrated with the second part. The pdf of Y_2 is

$$\begin{aligned} g(y) &= \frac{3!}{1!1!1!}\left(\frac{y}{\theta}\right)\left(1-\frac{y}{\theta}\right)\frac{1}{\theta} = 6y(\theta-y)/\theta^3,\ 0<y<\theta. \\ E(Y_2) &= \int_0^\theta 6y^2(\theta-y)/\theta^3\,dy. \end{aligned}$$

Let $y = \theta w$ (this is the observation and this substitution can be used in each part); so

$$E(Y_2) = 6\theta\int_0^1 w^2(1-w)\,dw = \frac{6\theta\Gamma(3)\Gamma(2)}{\Gamma(5)} = \frac{\theta}{2}.$$

Likewise

$$E(Y_2^2) = 6\theta^2\frac{\Gamma(4)\Gamma(2)}{\Gamma(6)} = \frac{3\theta^2}{10}; \qquad \operatorname{var}(Y_2) = \frac{\theta^2}{20}.$$

So

$$E(2Y_2) = \theta \text{ and } \operatorname{var}(2Y_2) = (4)(\theta^2/20) = \theta^2/5.$$

7.1.6 We have that $E(Y) = n\theta$, $\operatorname{var}(Y) = n\theta$. Thus

$$E[(\theta - b - Y/n)^2] = (y - b - \theta)^2 + (n\theta)(1/n)^2 = b^2 + \theta^2/n.$$

Thus take $b = 0$ and us $\delta(y) = y/n$. Clearly $\max(\theta^2/n)$ does not exist.

7.1.7

$$\begin{aligned} E(bS^2) &= (b\theta/n)E(nS^2/\theta) = b(n-1)\theta/n, \\ \mathrm{var}(bS^2) &= (b\theta/n)^2(2)(n-1) = 2b^2(n-1)\theta^2/n. \end{aligned}$$

Therefore

$$E[(\theta - bS^2)] = [\theta - b(n-1)\theta/n]^2 + 2b(n-1)\theta^2/n^2.$$

Differentiating this expression with respect to b, we have

$$2[\theta - b(n-1)\theta/n][-(n-1)\theta/n] + 4b(n-1)\theta^2/n^2 = 0.$$

Solving this expression for b yields $b = n/(n+1)$. Thus the estimator that provides minimum mean square error is

$$\frac{n}{n+1}S^2 = \frac{1}{n+1}\sum(X_i - \bar{X})^2.$$

It is interesting to note that this principle implies the use of $n+1$, rather than n or $n-1$ suggested by most books on statistics.

7.2.2

$$\frac{e^{-n\theta}\theta^{\sum x_i}}{x_1!x_2!\cdots x_n!} = \left[e^{-n\theta}\theta^{\sum x_i}\right]\left[\frac{1}{x_1!x_2!\cdots x_n!}\right];$$

so, by the factorization theorem, $Y = \sum X_i$ is a sufficient statistics for θ.

7.2.3 $f(x;\theta) = Q(\theta)M(x)I_{(0,\theta)}(x)$. Therefore

$$\prod_{i=1}^n Q(\theta)M(x_i)I_{(0,\theta)}(x_i) = \left\{[Q(\theta)]^n I_{(0,\theta)}[\max(x_i)]\right\}\left\{\prod_{i=1}^n M(x_i)\right\},$$

because $\prod I_{(0,\theta)}(x_i) = \prod I_{(0,\theta)}[\max(x_i)]$. According to the factorization theorem, $Y = \max(X_i)$ is a sufficient statistic for θ.

7.2.7

$$\prod_{i=1}^n \frac{x_i^{\theta-1}e^{-x_i/6}}{\Gamma(\theta)6^\theta} = \left\{\frac{(\prod x_i)^{\theta-1}}{[\Gamma(\theta)]^n 6^{n\theta}}\right\} e^{-\sum x_i/6}$$

so $Y = \prod X_i$ is a sufficient statistic for θ.

7.2.8

$$\prod_{i=1}^n \frac{\Gamma(2\theta)}{[\Gamma(\theta)]^2}[x_i(1-x_i)]^{\theta-1} = \left\{\frac{[\Gamma(2\theta)]^n}{[\Gamma(\theta)]^{2n}}\left[\prod x_i(1-x_i)\right]^{\theta-1}\right\}(1).$$

Thus $Y = \prod[X_i(1-X_i)]$ is a sufficient statistic for θ.

7.3.2

$$
\begin{aligned}
g(y_3|y_5) &= \frac{(5!/2!)(y_3/\theta)^2[(y_5-y_3)/\theta](1/\theta)^2}{5(y_5/\theta)^4(1/\theta)},\ 0<y_3<y_5<\theta, \\
&= 12y_3^2(y_5-y_3)/y_5^4,\ 0<y_3<y_5<\theta. \\
E(2Y_3|y_5) &= 24\int_0^{y_5} y_3^3(y_5-y_3)/y_5^4\,dy_3.
\end{aligned}
$$

Let $y_3 = y_5 w$ to obtain

$$
\begin{aligned}
E(2Y_3|y_5) &= 24y_5\int_0^1 w^3(1-w)\,dw \\
&= \frac{24y_5\Gamma(4)\Gamma(2)}{\Gamma(6)} = \left(\frac{6}{5}\right) y_5 = \varphi(y_5).
\end{aligned}
$$

7.3.5 For illustration, in Exercise 7.2.1, $Y = \sum X_i^2$ is a sufficient statistic, and

$$
E(Y) = \sum E(X_i^2) = \sum \theta = n\theta
$$

Thus $Y/n = \sum X_i^2/n$ is an unbiased estimator.

7.3.6 It suffices to find the conditional distribution of X_1 given $\sum_{i=1}^n X_i = x$. Assuming $x \geq x_1$ (otherwise the following probability is 0) we have

$$
\begin{aligned}
P[X_1 = x_1|\sum_{i=1}^n X_i = x] &= \frac{P[X_1 = x_1, \sum_{i=1}^n X_i = x]}{P[\sum_{i=1}^n X_i = x]} \\
&= \frac{P[X_1 = x_1, \sum_{i=2}^n X_i = x - x_1]}{P[\sum_{i=1}^n X_i = x]} \\
&= \frac{e^{-\theta}\frac{\theta^{x_1}}{x_1!}e^{-(n-1)\theta}\frac{[(n-1)\theta]^{x-x_1}}{(x-x_1)!}}{e^{-n\theta}\frac{[n\theta]^x}{x!}} \\
&= \binom{x}{x_1}\left(\frac{1}{n}\right)^{x_1}\left(1-\frac{1}{n}\right)^{x-x_1}.
\end{aligned}
$$

Thus, the conditional distribution is binomial and $E[X_1|\sum_{i=1}^n X_i = x] = x/n$. By linearity of conditional expectation it follows that

$$
E[X_1 + 2X_2 + 3X_3|\sum_{i=1}^n X_i = x] = (6x)/n = 6\overline{x}.
$$

7.4.1

$$
\begin{aligned}
\sum_{x=0}^{2} u(x)\binom{2}{x}\theta^x(1-\theta)^{2-x} &= u(0)(1-\theta)^2 + 2u(1)\theta(1-\theta) + u(2)\theta^2 \\
&= [u(0) - 2u(1) + u(2)]\theta^2 + [-2u(0) + 2u(1)]\theta + [u(0)] \\
&\equiv 0
\end{aligned}
$$

Thus each expression in brackets must be zero, which implies that $u(0) = u(1) = u(2) = 0$.

7.4.2 In each case $E(X) = 0$ for all $\theta > 0$.

7.4.3 A generalization of 7.4.1. Since $E[\sum X_i] = n\theta, \sum X_i/n$ is the unbiased minimum variance estimator.

7.4.4

(a) $\int_0^\theta u(x)(1/\theta)\,dx = 0$ implies $\int_0^\theta u(x)\,dx = 0,\ 0 < \theta$. Taking the derivative of the last expression w.r.t. θ, we obtain $u(\theta) = 0,\ 0 < \theta$.

(b) Take $u(x) = x - 1/2, 0 < x < 1$, and zero elsewhere.

$$E[u(x)] = \int_0^1 \left(x - \frac{1}{2}\right) dx + \int_1^\theta 0 \cdot dx = 0, \text{ provided } 1 < \theta.$$

7.4.6

(a) The pdf of Y is

$$\begin{aligned} g(y;\theta) &= P(Y \le y) - P(Y \le y-1) \\ &= [y/\theta]^n - [(y-1)/\theta]^n,\ y = 1, 2, \ldots, \theta. \end{aligned}$$

The quotient of $\prod f(x_i;\theta) = (1/\theta)^n,\ 1 \le x_i \le \theta$, and $g(y;\theta)$ is free of θ. It is easy to show $\sum u(y)g(y;\theta) \equiv 0$ for all $\theta = 1, 2, 3, \ldots$ implies that $u(1) = u(2) = u(3) = \ldots = 0$.

(b) The expected value of that expression in the book, say $v(Y)$, is

$$\sum_{y=1}^{\theta} v(y)g(y;\theta) = \left(\frac{1}{\theta^n}\right) \sum_{y=1}^{\theta} [y^{n+1} - (y-1)^{n+1}].$$

Clearly, by substituting $y = 1, 2, \ldots, \theta$, the summation equals θ^{n+1}; so

$$E[v(y)] = \left(\frac{1}{\theta^n}\right) \theta^{n+1} = \theta.$$

7.4.8 Note that there is a typographical error in the definition of the pmf. The binomial coefficient should be $\binom{n}{|x|}$ not $\binom{n}{x}$.

(a). Just consider the function $u(X) = X$. Then $E(X) = 0$ for all θ, but X is not identically 0.

(b). Y is sufficient because the distribution of X conditioned on $Y = y$ has space $\{-y, y\}$ with probabilities 1/2 for each point, if $y \neq 0$. If $y = 0$ then conditionally $X = 0$ with probability 1. The conditional distribution does not depend on θ. For completeness, suppose, for any function u, $E(u(y)) = 0$. Then we have

$$\begin{aligned} 0 &= \sum_{j=0}^{n} u(j)\binom{n}{j}\theta^j(1-\theta)^{n-j} \\ &= (1-\theta)^n\left\{\sum_{j=0}^{n} u(j)\binom{n}{j}\left(\frac{\theta}{1-\theta}\right)^j\right\}. \end{aligned}$$

Because $1-\theta > 0$ the expression in braces is 0. This, though, is a polynomial in $\frac{\theta}{1-\theta}$ which is equal to the 0 polynomial. Therefore each of the coefficients $u(j)\binom{n}{j} = 0$. Because $\binom{n}{j} > 0$ for all j, we must have $u(j) \equiv 0$ for all j. Thus Y is complete.

7.5.2 By Theorem 7.5.2, Y is a complete sufficient statistic for θ. In addition

$$\begin{aligned} E(1/Y) &= \int_0^\infty (1/y)\frac{\theta^n y^{n-1}e^{-\theta y}}{\Gamma(n)}\,dy \\ &= \frac{\theta^n\Gamma(n-1)}{\theta^{n-1}\Gamma(n)} = \frac{\theta}{n-1}; \end{aligned}$$

so $(n-1)/Y$ is that MVUE estimator.

7.5.4 We know that $E[\psi(\bar{X})] = \theta$ since $E(X_1) = \theta$. Also $E(\bar{X}) = \theta$. Thus

$$E[\psi(\bar{X}) - \bar{X}] \equiv 0 \text{ for all } \theta > 0.$$

Since $\bar{X}$ is a complete and sufficient statistic for θ, $\psi(\bar{X}) - \bar{X} = 0$; that is, $\psi(\bar{X}) = \bar{X}$.

7.5.6

$$\begin{aligned} E\left[e^{tK(X)}\right] &= \int_a^b \exp\{(t+\theta)K(x) + S(x) + q(\theta)\}\,dx \\ &= \exp\{q(\theta) - q(\theta-t)\}\int_a^b \exp\{(t+\theta)K(x) + S(x) + q(\theta+t)\}\,dx. \end{aligned}$$

However the integral equals one since the integrand can be treated as a pdf, provided $\gamma < \theta + t < \delta$.

7.5.10 Since $f(x;\theta) = \exp\{(-\theta)x + \log x + 2\log\theta\}$, $0 < x < \infty$, $Y = \sum X_i$ is a complete and sufficient statistic for θ. Also

$$\begin{aligned} E(1/Y) &= \int_0^\infty (1/y)\frac{\theta^{2n} y^{2n-1}e^{-\theta y}}{\Gamma(2n)}\,dy \\ &= \frac{\theta^{2n}\Gamma(2n-1)}{\theta^{2n-1}\Gamma(2n)} = \theta/(2n-1). \end{aligned}$$

So $(2n-1)/Y$ is the MVUE estimator.

7.5.11 Similar to 7.5.4.

7.6.2 The distribution of Y/θ is $\chi^2(n)$. Thus

$$\begin{aligned} E(Y/\theta) &= n \text{ and } \operatorname{var}(Y/\theta) = 2n. \\ E(Y^2) &= (n\theta)^2 + \theta^2(2n) = (n^2+2n)\theta^2; \end{aligned}$$

thus $Y^2/(n^2+2n)$ is the MVUE of θ^2.

7.6.5 For part (a), since $Y = \sum_{i=1}^n X_i$, we have

$$\begin{aligned} P[X_1 \le 1 | Y = y] &= P[X_1 = 0|Y = y] + P[X_1 = 1|Y = y] \\ &= \frac{P[\{X_1 = 0\} \cap \{\sum_{i=2}^n X_i = y\}]}{P(Y = y)} \\ &\quad + \frac{P[\{X_1 = 1\} \cap \{\sum_{i=2}^n X_i = y-1\}]}{P(Y = y)} \\ &= \frac{e^{-\theta}e^{-(n-1)\theta}[(n-1)\theta]^y/y!}{e^{-n\theta}(n\theta)^y/y!} \\ &\quad + \frac{e^{-\theta}\theta e^{-(n-1)\theta}[(n-1)\theta]^{y-1}/(y-1)!}{e^{-n\theta}(n\theta)^y/y!} \\ &= \left(\frac{n-1}{n}\right)^y + \frac{y}{n-1}\left(\frac{n-1}{n}\right)^y \\ &= \left(\frac{n-1}{n}\right)^y \left(1 + \frac{y}{n-1}\right). \end{aligned}$$

Hence, the statistic $\left(\frac{n-1}{n}\right)^Y \left(1 + \frac{Y}{n-1}\right)$ is the MVUE of $(1+\theta)e^{-\theta}$.

7.6.8 $P(X \le 2) = \int_0^2 (1/\theta)e^{-x/\theta}\,dx = 1 - e^{-2/\theta}$. Since $\bar{X} = Y/n$, where $Y = \sum X_i$, is the mle of θ, then the mle of that probability is $1 - e^{-2/\bar{X}}$. Since $I_{(0,2)}(X_1)$ is an unbiased estimator of $P(X \le 2)$, let us find the joint pdf of $Z = X_1$ and Y by first letting $V = X_1 + X_2, U = X_1 + X_2 + X_3 + \ldots$. The Jacobian is one; then we integrate out those other variables obtaining

$$g(z, y; \theta) = \frac{(y-z)^{n-2}e^{y/\theta}}{(n-2)!\theta^n},\ 0 < z < y < \infty.$$

Since the pdf of Y is

$$g_2(y; \theta) = \frac{y^{n-1}e^{-y/\theta}}{(n-1)!\theta^n},\ 0 < y < \infty,$$

we have that the conditional pdf of Z, given $Y = y$, is

$$\begin{aligned}
h(z|y) &= \frac{(n-1)(y-z)^{n-2}}{y^{n-1}}, \ 0 < z < y. \\
E\left[I_{(0,2)}(Z)|y\right] &= \int_0^\infty \left\{\left[I_{(0,2)}(z)\right](n-1)(y-z)^{n-2}/y^{n-1}\right\} dy \\
&= 1 - \left(\frac{y-2}{y}\right)^{n-1} = 1 - (1-2/y)^{n-1}.
\end{aligned}$$

That is, the MVUE estimator is

$$\left(1 - \frac{2/\bar{X}}{n}\right)^{n-1}.$$

Of course, this is approximately equals to the mle when n is large.

7.6.11 The function of interest is $g(\theta) = \theta(1-\theta)$. Note, though, that $g'(1/2) = 0$; hence, the Δ procedure cannot be used. Expand $g(\theta)$ into a Taylor series about 1/2, i.e.,

$$g(\theta) = g(1/2) + 0 + g''(1/2)\frac{(\theta - (1/2))^2}{2} + R_n.$$

Evaluating this expression at $\overline{X}$, we have

$$\overline{X}(1-\overline{X}) = \frac{1}{4} + (-2)\frac{(\overline{X} - (1/2))^2}{2} + R_n.$$

That is,

$$\frac{n((1/4) - \overline{X}(1-\overline{X})}{1/4} = \frac{n(\overline{X} - (1/2))^2}{1/4} + R_n^*.$$

As on Page 216, we can show that the remainder goes to 0 in probability. Note that the first term on the right side goes to the $\chi^2(1)$ distribution as $n \to \infty$. Hence, so does the left side.

7.7.3

$$\begin{aligned}
f(x,y) &= \exp\left\{\left[\frac{-1}{2(1-\rho^2)\sigma_1^2}\right]x^2 + \left[\frac{-1}{2(1-\rho^2)\sigma_2^2}\right]y^2 + \left[\frac{\rho}{(1-\rho^2)\sigma_1\sigma_2}\right]xy\right. \\
&\quad + \left[\frac{\mu_1}{(1-\rho)\sigma_1^2} - \frac{\rho\mu_2}{(1-\rho^2)\sigma_1\sigma_2}\right]x + \left[\frac{\mu_2}{(1-\rho^2)\sigma_2^2} - \frac{\rho\mu_1}{(1-\rho^2)\sigma_1\sigma_2}\right]y \\
&\quad \left. + q(\mu_1,\mu_2,\sigma_1,\sigma_2,\rho)\right\}
\end{aligned}$$

Hence $\sum X_i^2, \sum Y_i^2, \sum X_iY_i, \sum X_i, \sum Y_i$ are joint complete sufficient statistics. Of course, the other five provide a one-to-one transformation with these five; so they are also joint complete and sufficient statistic.

7.7.4 Thus $K_1(x) = cK_2(x) + c_1$. Substituting this for $K_1(x)$ in the first expression we obtain an expression of the form of the second.

7.7.6

(a)
$$\int_{\theta_1-\theta_2}^{\theta_1+\theta_2}\int_{\theta_1-\theta_2}^{y_n}\frac{u(y_1,y_n)(y_n-y_1)^{n-2}}{(2\theta_2)^n}dy_1dy_n \equiv 0$$
for all $-\infty < \theta_1 < \infty$, $0 < \theta_2 < \infty$. Multiply by $(2\theta_2)^n$ and differentiate first w.r.t. θ_1 and then w.r.t. θ_2 (this is not easy). This finally yields
$$u(\theta_1-\theta_2, \theta_1+\theta_2) = 0, \text{ for all } (\theta_1, \theta_2),$$
which implies that
$$u(y_1, y_2) = 0, \text{ a.e.}$$

(b) $E(Y_1) = (\theta_1-\theta_2) + (2\theta_2)/(n+1)$, $E(Y_n) = (\theta_1+\theta_2) - (2\theta_2)/(n+1)$. So
$$E\left(\frac{Y_1+Y_n}{2}\right) = \theta_1 \text{ and } E(Y_n - Y_1) = 2\theta_2 - 4\theta_2/(n+1) = \theta_2\left(\frac{2n-2}{n+1}\right).$$
That is, $(Y_1+Y_n)/2$ and $[(n+1)(Y_n-Y_1)]/[2(n-1)]$ are those MVUE estimates.

7.7.9 Part (a): Consider the following function of the sufficient and complete statistics
$$\begin{aligned}\mathbf{W} &= \sum_{i=1}^n(\mathbf{X}_i-\overline{\mathbf{X}})(\mathbf{X}_i-\overline{\mathbf{X}})' \\ &= \sum_{i=1}^n \mathbf{X}_i\mathbf{X}_i' - n\overline{\mathbf{X}}\overline{\mathbf{X}}'.\end{aligned}$$
Recall that the variance-covariance matrix of a random vector $\mathbf{Z}$ can be expressed as
$$\operatorname{cov}(\mathbf{Z}) = E[\mathbf{Z}\mathbf{Z}'] - E[\mathbf{Z}]E[\mathbf{Z}]'.$$
In the notation of the example, we have
$$E[\sum_{i=1}^n \mathbf{X}_i\mathbf{X}_i'] = \sum_{i=1}^n E[\mathbf{X}_i\mathbf{X}_i'] = n\mathbf{\Sigma} + n\boldsymbol{\mu}\boldsymbol{\mu}'.$$
But the random vector $\overline{\mathbf{X}}$ has mean $\boldsymbol{\mu}$ and variance-covariance matrix $n^{-1}\mathbf{\Sigma}$. Hence,
$$E[\overline{\mathbf{X}}\overline{\mathbf{X}}'] = n^{-1}\mathbf{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}'.$$
Putting these last two results together
$$E[\mathbf{W}] = (n-1)\mathbf{\Sigma},$$
i.e., $\mathbf{S} = (n-1)^{-1}\mathbf{W}$ is an unbiased estimator of $\mathbf{\Sigma}$. Thus the $(i,j)th$ entry of $\mathbf{S}$ is the MVUE of σ_{ij}.

7.7.12 The order statistics are sufficient and complete and $\overline{X}$ is a function of them. Further, $\overline{X}$ is unbiased. Hence, $\overline{X}$ is the MVUE of μ.

7.8.1

(c) We know that $Y = \sum X_i$ is sufficient for θ and the mle is $\hat{\theta} = \bar{X}/3 = Y/3n$, which is one-to-one and hence also sufficient for θ.

7.8.2

(a)

$$\prod_{i=1}^{n} \left(\frac{1}{2\theta}\right) I_{[-\theta,\theta]}(x_i) = \left(\frac{1}{2\theta}\right)^n I_{[-\theta,y_n]}(y_1) I_{[y_1,\theta]}(y_n);$$

by the factorization theorem, the pair (Y_1, Y_n) is sufficient for θ.

(b) $L = \left(\frac{1}{2\theta}\right)^n$, provided $-\theta \le y_1$ and $y_n \le \theta$. That is, $-y_1 \le \theta$ and $y_n \le \theta$. We want to make θ as small as possible and satisfy these two restrictions; hence $\hat{\theta} = \max(-Y_1, Y_n)$.

(c) It is easy to show from the joint pdf Y_1 and Y_n that the pdf of $\hat{\theta}$ is $g(z;\theta) = nz^{n-1}/\theta^n, 0 \le z \le \theta$, zero elsewhere. Hence

$$L/g(z;\theta) = \frac{1}{2^n(nz^{n-1})}, \quad -z = -\hat{\theta} \le x_i \le \hat{\theta} = z,$$

which is free of θ.

7.8.7 For illustration $Y_{n-2} - Y_3, \min(-Y_1, Y_n)/\max(Y_1, Y_n)$ and $(Y_2 - Y_1)/\sum(Y_i - Y_1)$, respectively.

7.9.3 From previous results (Chapter 3), we know that Z and Y have a bivariate normal distribution. Thus they are independent if and only if their covariance is equal to zero; that is

$$\sum_{i=1}^{n} a_i \sigma^2 = 0 \text{ or, equivalently, } \sum_{i=1}^{n} a_i = 0.$$

If $\sum a_i = 0$, note that $\sum a_i X_i$ is location-invariant because $\sum a_i(x_i + d) = \sum a_i x_i$.

7.9.5 Of course, R is a scale-invariant statistic, and thus R and the complete sufficient statistic $\sum_1^n Y_i$ for θ are independent. Since $M_1(t) = E[\exp(tnY_1)] = (1 - \theta t)^{-1}$ for $t < 1/\theta$, and $M_2(t) = E[\exp(t\sum_1^n Y_i)] = (1 - \theta t)^{-n}$ we have

$$M_1^{(k)}(0) = \theta^k \Gamma(k+1) \text{ and } M_2^{(k)}(0) = \theta^k \Gamma(n+k)/\Gamma(n).$$

According to the result of Exercise 7.9.4 we now have $E(R^k) = M_1^{(k)}(0)/M_2^{(k)}(0) = \Gamma(k+1)\Gamma(n)/\Gamma(n+k)$. These are the moments of a beta distribution with $\alpha = 1$ and $\beta = n - 1$.

7.9.7 The two ratios are location- and scale-invariant statistics and thus are independent of the joint complete and sufficient statistic for the location and scale parameters, namely $\bar{X}$ and S^2.

7.9.9

(a) Here R is a scale-invariant statistic and hence independent of the complete and sufficient statistic, $\sum X_i^2$, for θ, the scale parameter.

(b) While the numerator, divided by θ, is $\chi^2(2)$ and the denominator, divided by θ, is $\chi^2(5)$, they are not independent and hence 5R/2 does not have an F-distribution.

(c) It is easy to get the moment of the numerator and denominator and thus the quotient of the corresponding moments to show that R has a beta distribution.

7.9.13 (a). Ignoring constants, the log of the likelihood is

$$l(\theta) \propto 3n \log \theta - \theta \sum_{i=1}^{n} x_i.$$

Taking the partial derivative of this expression with respect to θ, shows that the mle of θ is $3n/Y$. As we show below, it is biased.

(b). Immediate, because this pdf is a member of the regular exponential class.

(c). Because Y has a $\Gamma(3n, \theta^{-1})$ distribution, we have

$$\begin{aligned} E[Y^{-1}] &= \int_0^\infty \frac{1}{\Gamma(3n)\theta^{-3n}} y^{3n-1)-1} e^{-\theta y}\, dy \\ &= \int_0^\infty \frac{1}{\Gamma(3n)\theta^{-3n}} \theta^{-3n+2-1} z^{(3n-1)-1} e^{-z}\, dz \\ &= \frac{\theta}{3n-1}, \end{aligned}$$

where we used the substitution $z = \theta y$. Hence, the MVUE is $(3n-1)/Y$. Also, the mle is biased.

(d). The mgfs of X_1 and Y are $(1-\theta^{-1}t)^{-3}$ and $(1-\theta^{-1}t)^{-3n}$, respectively. It follows that θX_1 and θY have distributions free of θ. Hence, so does $X_1/Y = (X_1\theta)/(Y\theta)$. So by Theorem 7.9.1, X_1/Y and Y are independent.

(e). Let $T = X_1/Y = X_1/(X_1 + Z)$, where $Z = \sum_{i=2}^n X_i$. Let $S = Y = X_1 + Z$. Then the inverse transformation is $x_1 = st$ and $z = s(1-t)$ with spaces $0 < t < 1$ and $0 < s < \infty$. It is easy to see that the Jacobian is $J = s$. Because X_1 has a $\Gamma(3, 1/\theta)$ distribution, Z has a $\Gamma(3(n-1), 1/\theta)$ distribution, and X_1 and Z are independent, we have

$$\begin{aligned} f_{T,S}(t,s) &= f_{X_1}(st) f_Z(s(1-t))s \\ &= \frac{\theta^{3n}}{2\Gamma(3(n-1))} \left\{ t^{3-1}(1-t)^{3(n-1)-1} \right\} s^{3n-1} e^{-\theta s}. \end{aligned}$$

Based on the function of t within the braces, we see immediately that $T = X_1/Y$ has a beta distribution with parameters 3 and $3(n-1)$.

Chapter 8

Optimal Tests of Hypotheses

8.1.4

$$
\begin{aligned}
\frac{\left[\frac{1}{\sqrt{(2\pi)(1)}}\right]^n \exp\left[-\frac{\sum x_i^2}{(2)(1)}\right]}{\left[\frac{1}{\sqrt{(2\pi)(2)}}\right]^n \exp\left[-\frac{\sum x_i^2}{(2)(2)}\right]} &\leq k, \\
\exp\left[\frac{-\sum x_i^2}{(2)(2)}\right] &\leq k/(\sqrt{(}2)^n, \quad \sum x_i^2 \geq -4\log\left[k/(\sqrt{2})^n\right] = c.
\end{aligned}
$$

Since $\sum X_i$ is $\chi^2(r=10)$ under H_0, we take $c = 18.3$. Yes. Yes.

8.1.5

$$
\frac{1^n}{(2x_1)(2x_2)\dots(2x_n)} \leq k \text{ implies that } c = \frac{1}{2^n k} \leq \prod_{i=1}^n x_i.
$$

8.1.8

$$
\frac{1^n}{[6x_1(1-x_1)][6x_2(1-x_2)]\cdots[6x_n(1-x_n)]} \leq k \text{ implies that } c = \frac{1}{6^n k} \leq \prod_{i=1}^n [x_i(1-x_i)].
$$

8.1.10

$$
\frac{\frac{(0.1)^{\sum x_i} e^{-n(0.1)}}{x_1!x_2!\cdots x_n!}}{\frac{(0.5)^{\sum x_i} e^{-n(0.5)}}{x_1!x_2!\cdots x_n!}} \leq k; \quad \frac{e^{n(0.4)}}{k} \leq 5^{\sum x_i}; \quad c \leq \sum_{i=1}^n x_i.
$$

If $n = 10$ and $c = 3$; $\gamma(\theta) = P_\theta(\sum X_i \geq 3)$. So $\alpha = \gamma(0.1) = 0.08$ and $\gamma(0.5) = 0.875$.

8.2.2 The pdf of Y_4 is

$$
g(y;\theta) = 4y^3/\theta^4, \ 0 < y < \theta;
$$

So

$$\begin{aligned}
P(Y_4 \le 1/2 \text{ or } Y_4 \ge 1) &= \int_0^{1/2} 4y^3/\theta^4\,dy = \left(\frac{1}{2\theta}\right)^4, \quad \text{if } \theta < 1 \\
&= \int_0^{1/2} 4y^3/\theta^4\,dy + \int_1^{\theta} 4y^3/\theta^4\,dy \\
&= \left(\frac{1}{2\theta}\right)^4 + 1 - \frac{1}{\theta^4} = 1 - \frac{15}{16\theta^4}, \quad \text{if } 1 < \theta.
\end{aligned}$$

8.2.3

$$\begin{aligned}
\gamma(\theta) &= P_\theta(\bar{X} \ge 3/5) = P_\theta\left(\frac{\bar{X}-\theta}{2/5} \ge \frac{3/5-\theta}{2/5}\right) \\
&= 1 - \Phi\left(\frac{3-5\theta}{5}\right).
\end{aligned}$$

8.2.6 If $\theta > \theta'$, then we want to use a critical region of the from $\sum x_i^2 \ge c$. If $\theta < \theta'$, the critical region is like $\sum x_i^2 \le c$. That is, we cannot find one test which will be best for each type of alternative.

8.2.9 Let $X_1, X_2, \ldots, X_n$ be a random sample with the common Bernoulli pmf with parameter as given in the problem. Based on Example 8.2.5, the UMP test rejects H_0 if $Y \ge c$, $Y = \sum_{i=1}^n X_i$. In general, Y has a binomial(n,θ) distribution. To determine n we solve two simultaneous equations, one involving level and the other power. The level equation is

$$\begin{aligned}
0.05 &= \gamma(1/20) = P_{1/20}\left[\frac{Y-(n/20)}{\sqrt{19n/400}} \ge \frac{c-(n/20)}{\sqrt{19n/400}}\right] \\
&\doteq P\left[Z \ge \frac{c-(n/20)}{\sqrt{19n/400}}\right],
\end{aligned}$$

where by the Central Limit Theorem Z has a standard normal distribution. Hence, we get the equation

$$\frac{c-(n/20)}{\sqrt{19n/400}} = 1.645. \tag{8.0.1}$$

Likewise from the desired power $\gamma(1/10) = 0.90$, we obtain the equation

$$\frac{c-(n/20)-(n/10)}{\sqrt{9n/100}} = -1.282. \tag{8.0.2}$$

Solving (8.0.1) and (8.0.2) simultaneously, gives the solution $n = 122$.

8.2.10 The mgf of $Y = \sum_{i=1}^n X_i$ is $(1-\theta t)^{-n}, t < 1/\theta$, which is gamma $(\alpha = n, \beta = \theta)$. Thus, with the uniformly most powerful critical region of the form $\sum x_i \geq c$, we have the power function

$$\gamma(\theta) = \int_c^\infty \frac{1}{\Gamma(n)\theta^n} y^{n-1} e^{-y/\theta}\, dy.$$

8.2.12 (a)

$$\begin{aligned} \frac{\left(\frac{1}{2}\right)^{\sum x_i} \left(\frac{1}{2}\right)^{5-\sum x_i}}{\theta^{\sum x_i}(1-\theta)^{5-\sum x_i}} &\leq k, \quad \text{with } \theta < 1/2. \\ \left[\left(\frac{1/2}{\theta}\right)\left(\frac{1-\theta}{1/2}\right)\right]^{\sum x_i} &\leq k[2(1-\theta)]^5. \end{aligned}$$

Since the quantity in brackets on the left side is clearly greater than one, this inequality is of the form $\sum x_i \leq c$.

(b) $P(Y \leq 1,$ when $\theta = 1/2) = \sum_{y=0}^1 \binom{5}{y} \left(\frac{1}{2}\right)^5 = \frac{6}{32}$.

(c) $P(Y \leq 0,$ when $\theta = 1/2) = \left(\frac{1}{2}\right)^5 = \frac{1}{32}$.

(d) Alway reject H_0 if $Y = 0$. If $Y = 1$, select a random number from $0, 1, \ldots, 9$ and if it is 0 or 1, reject H_0. Thus

$$\alpha = \frac{1}{32} + \left(\frac{5}{32}\right)\left(\frac{2}{10}\right) = \frac{2}{32}.$$

8.3.5 Say $H_0 : \theta = \theta_0$ and $H_1 : \theta = \theta_1$. The likelihood ration test is

$$\lambda = \frac{f(x_1;\theta_0)f(x_2;\theta_0)\cdots f(x_n;\theta_0)}{\max[f(x_1;\theta_i)f(x_2;\theta_i)\cdots f(x_n;\theta_i), i = 0, 1]} \leq k$$

If the maximum in the denominator occurs when $i = 0$, the $\lambda = 1$ and we do not reject. If that maximum occurs when $i = 1$, then

$$\lambda = \frac{f(x_1;\theta_0)\cdots f(x_n;\theta_0)}{f(x_1;\theta_1)\cdots f(x_n;\theta_1)} \leq k$$

which is the critical region given by the Neyman-Pearson theorem.

8.3.6

$$\lambda = \frac{\left(\frac{1}{\sqrt{2\pi}}\right)^n \exp\left[-\sum(x_i - \theta')^2/2\right]}{\left(\frac{1}{\sqrt{2\pi}}\right)^n \exp\left[-\sum(x_i - \bar{x})^2/2\right]} \leq k$$

is equivalent to

$$\exp\left\{\left[-\sum(x_i - \bar{x})^2 - n(\bar{x} - \theta')^2 + \sum(x_i - \bar{x})^2\right]/2\right\} \leq c_1$$

and thus

$$n(\bar{x} - \theta) \geq c_2 \text{ and } |\bar{x} - \theta'| \geq c_3.$$

8.3.9 Since $\sum |x_i - \theta|$ is minimized when $\hat{\theta} = \text{median}(X_i) = Y_3$, the likelihood ratio test is

$$\lambda = \frac{\left(\frac{1}{2}\right)^5 \exp\left[-\sum |x_i - \theta_0|\right]}{\left(\frac{1}{2}\right)^5 \exp\left[-\sum |x_i - y_3|\right]} \leq k.$$

This is the equivalent to

$$\exp\left[-\sum |x_i - y_3| - 5|y_3 - \theta_0| + \sum |x_i - y_3|\right] \leq k$$

and

$$|y_3 - \theta_0| \geq c.$$

8.3.11 The likelihood function for this problem is

$$L(\theta) = \theta^n \left[\prod_{i=1}^{n} (1 - x_i)\right]^{\theta - 1}.$$

(a) For $\theta_1 < \theta_2$, the ratio of the likelihoods is

$$\frac{L(\theta_1)}{L(\theta_2)} = \left(\frac{\theta_1}{\theta_2}\right)^n \left[\prod_{i=1}^{n} (1 - x_i)\right]^{\theta_1 - \theta_2}.$$

This has decreasing monotone-likelihood-ratio in the statistic $\prod_{i=1}^{n}(1 - x_i)$. Hence, the UMP test, rejects H_0 if $\prod_{i=1}^{n}(1 - x_i) \geq c$.

(b) Taking the partial derivative of the log of the likelihood, yields the mle estimator:

$$\widehat{\theta} = \frac{n}{-\log \prod_{i=1}^{n}(1 - x_i)}.$$

The likelihood ratio test statistic is

$$\Lambda = \frac{1}{\widehat{\theta}^n \left(\prod_{i=1}^{n}(1 - x_i)\right)^{\widehat{\theta} - 1}}.$$

8.3.15 We obtain the cdf of X by conditioning on I_ϵ. Using independence between I_ϵ and Z and Y, we have

$$\begin{aligned} P(X \leq x | I_\epsilon = 0) &= P(Z \leq x) = \Phi(x) \\ P(X \leq x | I_\epsilon = 1) &= P(Y \leq x) = \Phi\left(\frac{x - \mu_c}{\sigma_c}\right). \end{aligned}$$

Hence, the cdf followed by the pdf are:

$$\begin{aligned} P(X \leq x) &= (1 - \epsilon)\Phi(x) + \epsilon \Phi\left(\frac{x - \mu_c}{\sigma_c}\right) \\ f_X(x) &= (1 - \epsilon)\phi(x) + \frac{\epsilon}{\sigma_c}\phi\left(\frac{x - \mu_c}{\sigma_c}\right). \end{aligned}$$

The mean and second moment of X are:

$$\begin{aligned} E(X) &= E(1-I_\epsilon)E(Z) + E(I_\epsilon)E(Y) = \epsilon\mu_c \\ E(X^2) &= E(1-I_\epsilon)E(Z^2) + E(I_\epsilon)E(Y^2) = (1-\epsilon) + \epsilon(\sigma_c^2 + \mu_c^2), \end{aligned}$$

where for the second moment we used the fact that the square of an indicator is the indicator and that the cross product is 0 with probability 1. Hence, the variance of X is: $(1-\epsilon) + \epsilon(\sigma_c^2 + \mu_c^2) - \epsilon^2\mu_c^2$.

8.4.2

$$\begin{aligned} \frac{0.2}{0.9} \approx k_0 &< \frac{(0.02)^{\sum x_i} e^{-n(0.02)}}{(0.07)^{\sum x_i} e^{-n(0.07)}} < k_1 \approx \frac{0.8}{0.1} \\ \frac{2}{9} &< \left(\frac{2}{7}\right)^{\sum x_i} e^{(0.05)n} < 8 \\ c_1(n) &= \frac{\log(2/9) - (0.05)n}{\log(2/7)} > \sum x_i > \frac{\log 8 - (0.05)n}{\log(2/7)} = c_0(n). \end{aligned}$$

8.4.4

$$\begin{aligned} \frac{0.02}{0.98} &< \frac{(0.01)^{\sum x_i}(0.99)^{100n - \sum x_i}}{(0.05)^{\sum x_i}(0.95)^{100n - \sum x_i}} < \frac{0.98}{0.02} \\ -\log 49 &< (\sum x_i)\log\left(\frac{19}{99}\right) + 100n\log\left(\frac{99}{95}\right) < \log 49 \\ \frac{[-100\log(99/95)]n - \log 49}{\log(19/99)} &> \sum x_i > \frac{-100\log(99/95)] + \log 49}{\log(19/99)} \end{aligned}$$

or, equivalently,

$$\frac{\log 49}{\log(99/19)} > \sum\left[x_i - 100\frac{\log(99/95)}{\log(99/19)}\right] > \frac{-\log 49}{\log(99/19)}.$$

8.5.2 (a) and (b) are straightforward.

(c) (1) $P(\sum X_i \geq c; \theta = 1/2) = (2)P(\sum X_i < c; \theta = 1)$ where $\sum X_i$ is Poisson (10θ). Using the Poisson tables, we find, with $c = 6$, the left side is too large, namely $1 - 0.616 > (2)(0.067)$. With $c = 7$, the left side is too small, namely $1 - 0.762 < 2(0.130)$ or, equivalently, $0.238 < 0.260$. To make this last inequality an equality, we need part of the probability that $Y = 6$, namely 0.146 and 0.063 under the respective hypotheses. So $0.238 + 0.146p = 0.260 - 2(0.063)p$ and $p = 0.08$.

8.5.4 Define $g(c)$ as follows and then take its derivative:

$$\begin{aligned} g(c) &= \Phi(c-78) - 3 + 3\Phi(c-75) \\ g'(c) &= \phi(c-78) + 3\phi(c-75) \end{aligned}$$

We want to solve $g(c) = 0$. If c_0 is an initial guess at the solution, then the next guess via Newton's algorithm is

$$c = c_0 - \frac{g(c_0)}{g'(c_0)}.$$

Here is an R function which performs a Newton step for this problem. If $c_0 = 75$ is chosen, in a few steps it is quite close to 76.8.

```
newtstp = function(c0){
     gc0 = pnorm(c0-78) - 3 + 3*pnorm(c0-75)
     gpc0 = dnorm(c0-78) + 3*dnorm(c0-75)
     c = c0 - (gc0/gpc0)
     gc = pnorm(c-78) - 3 + 3*pnorm(c-75)
     list(c0=c0,gc0=gc0,c=c,gc=gc)
}
```

8.5.5

$$\begin{aligned}
\frac{\frac{1}{(1)(5)}\exp\left(-\frac{x}{1}-\frac{y}{5}\right)}{\frac{1}{(3)(2)}\exp\left(-\frac{x}{3}-\frac{y}{2}\right)} &= \frac{6}{5}\exp\left(-\frac{2x}{3}+\frac{3y}{10}\right) \le k \\
-\frac{2x}{3}+\frac{3y}{10} &\le \log\frac{5k}{6} = c
\end{aligned}$$

leads to classification as to (x, y) coming from the second distribution.

Chapter 9

Inferences about Normal Models

Remark In both 9.1.3 and 9.1.5, we can use the two-sample result that

$$\sum_{j=1}^{2}\sum_{i=1}^{n_j}(X_{ij}-\bar{X}_{..})^2=\sum_{j=1}^{2}\sum_{i=1}^{n_j}(X_{ij}-\bar{X}_{.j})^2+\sum_{j=1}^{2}n_j(X_{.j}-\bar{X}_{..})^2.$$

Of course, with the usual normal assumptions, the terms on the right side (once divided by σ^2 are chi-squared variables with n_1+n_2-2 and one degrees of freedom, respectively; and they are independent.

9.1.3 Let the two samples be X_1 and $(X_2,\ldots,X_{n-1})$. Then, since $(X_1-X_1)^2=0$,

$$\sum_{i=1}^{n}(X_i-\bar{X})^2=\sum_{i=2}^{n}(X_i-\bar{X}')+[(X_1-\bar{X})^2+(n-1)(\bar{X}'-\bar{X})^2].$$

If we write $\bar{X}=[X_1+(n-1)\bar{X}']/n$, it is easy to show that the second term on the right side is equal to $(n-1)(X_1-\bar{X}')^2/n$, and it is $\chi^2(1)$ after being divided by σ^2.

9.1.5 First take as the two samples X_1,X_2,X_3 and X_4. The result in 9.1.3 yields

$$\sum_{i=1}^{4}(X_i-\bar{X})^2=\sum_{i=1}^{3}\left(X_i-\frac{X_1+X_2+X_3}{3}\right)^2+\frac{3}{4}\left(X_4-\frac{X_1+X_2+X_3}{3}\right)^2.$$

Apply the result again to the first term on the right side using the two samples X_1,X_2 and X_3. The last step is taken using the two samples of X_1 and X_2.

9.2.1 It is easy to show the first equality by writing

$$\sum_{j=1}^{b}\sum_{i=1}^{a_j}(X_{ij}-\bar{X}_{..})^2=\sum\sum[(X_{ij}-\bar{X}_{.j})+(\bar{X}_{.j}-\bar{X}_{..})]^2$$

and squaring the binomial on the right side (the sum of the cross product term clearly equals zero).

9.2.3 For this problem the random variables X_{ij} are iid with variance σ^2. We express the covariance of interest into its four terms and then by using independence we obtain the following simplification for each term:

$$\begin{aligned}
\mathrm{cov}(X_{ij}, \overline{X}_{\cdot j}) &= \mathrm{cov}\left(X_{ij}, \frac{1}{a_j}\sum_{l=1}^{a_j} X_{lj}\right) = \mathrm{cov}\left(X_{ij}, \frac{1}{a_j}X_{ij}\right) = \frac{\sigma^2}{a_j} \\
\mathrm{cov}(X_{ij}, \overline{X}_{\cdot\cdot}) &= \mathrm{cov}\left(X_{ij}, \frac{1}{N}\sum_{k=1}^{b}\sum_{l=1}^{a_k} X_{lk}\right) = \mathrm{cov}\left(X_{ij}, \frac{1}{N}X_{ij}\right) \\
&= \frac{\sigma^2}{N} \\
\mathrm{cov}(\overline{X}_{\cdot j}, \overline{X}_{\cdot j}) &= \frac{\sigma^2}{a_j} \\
\mathrm{cov}(\overline{X}_{\cdot j}, \overline{X}_{\cdot\cdot}) &= \mathrm{cov}\left(\overline{X}_{\cdot j}, \frac{1}{N}\sum_{k=1}^{b}\sum_{l=1}^{a_k} X_{lk}\right) \\
&= \mathrm{cov}\left(\overline{X}_{\cdot j}, \frac{1}{N}\sum_{l=1}^{a_j} X_{lj}\right) = \mathrm{cov}\left(\overline{X}_{\cdot j}, \frac{a_j}{N}\overline{X}_{\cdot j}\right) \\
&= \frac{a_j}{N}\frac{\sigma^2}{a_j} = \frac{\sigma^2}{N}.
\end{aligned}$$

Hence,

$$\mathrm{cov}(X_{ij} - \overline{X}_{\cdot j}, \overline{X}_{\cdot j} - \overline{X}_{\cdot\cdot}) = \frac{\sigma^2}{a_j} - \frac{\sigma^2}{N} - \frac{\sigma^2}{a_j} + \frac{\sigma^2}{N} = 0.$$

9.2.5 This can be thought of as a two-sample problem in which the first sample is the first sample and the second is a combination of the last $(b-1)$ samples. The difference of the two means, namely bd, is estimated by

$$\sum_{j=2}^{b} \bar{X}_{\cdot j}/(b-1) - \bar{X}_{\cdot 1} = \bar{X}'_{\cdot\cdot} - \bar{X}_{\cdot 1} = b\hat{d};$$

hence the estimator, $\hat{d}$ of d given in the book. Using the result of 9.1.3,

$$\begin{aligned}
\sum_{j=1}^{b}(a)(\bar{X}_{\cdot j} - \bar{X}_{\cdot\cdot})^2 &= \sum_{j=2}^{b}(a)(\bar{X}_{\cdot j} - \bar{X}'_{\cdot\cdot})^2 + \left(\frac{b-1}{b}\right)(\bar{X}'_{\cdot\cdot} - \bar{X}_{\cdot 1})^2 \\
&= Q_6 + Q_7, \quad \text{say.}
\end{aligned}$$

Accordingly,

$$\frac{Q_7/1}{(Q_3 + Q_6)/[b(a-1) + b - 2]}$$

has an $F(1, ab-2)$ distribution.

9.3.1

$$\begin{aligned} E[\exp(t)(Y_1+\cdots+Y_n)] &= \prod_{i=1}^{n}\{(1-2t)^{-r_i/2}\exp[t\theta_i/(1-2t)]\} \\ &= (1-2t)^{-(r_1+r_2+\cdots+r_n)/2}\exp[t(\sum\theta_i)/(1-2t)], \end{aligned}$$

which is the mgf of a $\chi^2(\sum r_i, \sum\theta_i)$ distribution.

9.3.2

$$\begin{aligned} \psi(t) &= \log M(t) = (-r/2)\log(1-2t) + t\theta/(1-2t) \\ \psi'(t) &= \frac{r}{1-2t} + \frac{\theta[(1-2t)-t(-2)]}{(1-2t)^2} = \frac{r}{1-2t} + \frac{\theta}{(1-2t)^2} \\ \psi''(t) &= \frac{2r}{(1-2t)^2} + \frac{4\theta}{(1-2t)^3}. \end{aligned}$$

Thus mean $= \psi'(0) = r+\theta$ and variance $= \psi''(0) = 2r+4\theta$.

9.3.6 Substituting μ_j for X_{ij} we see that the non-centrality parameters are

$$\begin{aligned} \theta_3 &= \sum\sum(\mu_j-\mu_j)^2 = 0, \\ \theta_4 &= \sum(a_j)(\mu_j-\bar{\mu}_.)^2, \quad \text{where } \bar{\mu}_. = \sum(a_j)\mu_j/\sum a_j. \end{aligned}$$

Thus, Q_3' and Q_4' are independent; and

$$Q_3'/\sigma^2 \text{ is } \chi^2(\sum a_j - b, 0),$$

$$Q_4'/\sigma^2 \text{ is } \chi^2(b-1, \theta_4),$$

$$F = \frac{Q_4'(b-1)}{Q_3'/(\sum a_j - b)} \text{ is } F(b-1, \sum a_j - b, \theta_4).$$

9.4.1 $P(A_1\cup A_2) = P(A_1)+P(A_2)-P(A_1\cap A_2) \le P(A_1)+P(A_2)$.
Thus

$$P[(A_1\cup A_2)\cup A_3] \le P(A_1\cup A_2)+P(A_3) \le P(A_1)+P(A_2)+P(A_3),$$

and so on. Also

$$\begin{aligned} P(A_1^*\cap A_2^*\cap\cdots\cap A_k^*) &= 1-P(A_1\cup A_2\cup\ldots\cup A_k) \\ &\ge 1-\sum_{i=1}^{k}P(A_i). \end{aligned}$$

9.4.3 In the case of simultaneous testing, a Type I error occurs iff at least one of the individual test rejects when all the hypotheses are true ($\cap H_0$). Choose the critical regions $C_{i,\alpha/m}$, $i=1,2,\ldots,m$. Then by Boole's inequality

$$\begin{aligned} P(\text{Type I Error}) &= P_{\cap H_0}\left[\cup_{i=1}^{m}C_{i,\alpha/m}\right] \\ &\le \sum_{i=1}^{m}P_{\cap H_0}\left[C_{i,\alpha/m}\right] = \sum_{i=1}^{m}\frac{\alpha}{m} = \alpha. \end{aligned}$$

9.5.1 Write the left side as

$$\sum_{j=1}^{b}\sum_{i=1}^{a}[(X_{ij}-\bar{X}_{i.}-\bar{X}_{.j}+\bar{X}_{..})+(\bar{X}_{.j}-\bar{X}_{..})]^2;$$

then square the binomial, and it is easy to see that the sum of the cross-product term equals zero.

9.5.3 We want to minimize

$$\begin{aligned}
&\sum\sum(x_{ij}-\mu-\alpha_i-\beta_j)^2\\
&= \sum\sum[(x_{ij}-\bar{x}_{i.}-\bar{x}_{.j}+\bar{x}_{..})+(\bar{x}_{..}-\mu)+(\bar{x}_{i.}-\bar{x}_{..}-\alpha_i)+(\bar{x}_{.j}-\bar{x}_{..}-\beta_j)]^2\\
&= \sum\sum(x_{ij}-\bar{x}_{i.}-\bar{x}_{.j}+\bar{x}_{..})^2+ab(\bar{x}_{..}-\mu)^2+\sum_i(b)(\bar{x}_{i.}-\bar{x}_{..}-\alpha_i)^2.\\
&\quad+\sum_j(a)(\bar{x}_{.j}-\bar{x}_{..}-\beta_j)^2.
\end{aligned}$$

To do this, we can make the last three terms equal to zero by taking $\hat{\mu}=\bar{x}_{..}$, $\hat{\alpha}_i=\bar{x}_{i.}-\bar{x}_{..}$, $\hat{\beta}_j=\bar{x}_{.j}-\bar{x}_{..}$. For example,

$$\begin{aligned}
\text{var}(\hat{\alpha}_1) &= \text{var}\left(\bar{X}_{1.}-\frac{\bar{X}_{1.}+\cdots+\bar{X}_{a.}}{a}\right)\\
&= \text{var}\left[\frac{(a-1)\bar{X}_{1.}-\cdots-\bar{X}_{a.}}{a}\right]\\
&= \left[\left(\frac{a-1}{a}\right)^2+(a-1)\left(\frac{-1}{a}\right)^2\right]\frac{\sigma^2}{n}\\
&= \left(\frac{a-1}{a}\right)\left(\frac{\sigma^2}{n}\right).
\end{aligned}$$

9.5.6 This can be worked in a manner similar to 9.5.3.

9.6.4 Write

$$\widehat{\eta}_0=\widehat{\alpha}+\widehat{\beta}(x_0-\overline{x})=[1\ x_0-\overline{x}]\begin{bmatrix}\widehat{\alpha}\\ \widehat{\beta}\end{bmatrix}.$$

Then by expression (9.6.6) and Theorem 3.5.1, it follows that $\widehat{\eta}_0$ has a normal distribution with mean η_0 and variance

$$\begin{aligned}
\text{Var}(\widehat{\eta}_0) &= [1\ x_0-\overline{x}]\,\sigma^2\begin{bmatrix}\frac{1}{n} & 0\\ 0 & [\sum_{i=1}^n(x_i-\overline{x})^2]^{-1}\end{bmatrix}\begin{bmatrix}1\\ x_0-\overline{x}\end{bmatrix}\\
&= \sigma^2\left[\frac{1}{n}+\frac{(x_0-\overline{x})^2}{\sum_{i=1}^n(x_i-\overline{x})^2}\right].
\end{aligned}$$

Based on the distribution of $\widehat{\eta}$, the independence between $\widehat{\sigma}^2$ and $\widehat{\eta}_0$ ($\widehat{\eta}_0$ is a function of $\widehat{\alpha}$ and $\widehat{\beta}$), and the distribution of $\widehat{\sigma}^2$, the following random variable

$$t = \frac{\widehat{\eta}_0 - \eta_0}{\widehat{\sigma}\sqrt{(1/n) + (x_0 - \overline{x})^2 / \sum_{i=1}^n (x_i - \overline{x})^2}}$$

has a student t-distribution with $n-2$ degrees of freedom. The desired confidence interval easily follows from this result.

9.6.7 Let $\theta = \gamma^2$. Then

$$\begin{aligned}
L &= \prod_{i=1}^n \left(\frac{1}{2\pi\theta x_i^2}\right)^{1/2} \mathrm{e}^{-(y_i-\beta x_i)^2/(2\theta x_i^2)} \\
\log L &= d(x_1,\ldots,x_n) - \frac{n}{2}\log\theta - \sum_{i=1}^n \frac{(y_i-\beta x_i)^2}{2\theta x_i^2} \\
\frac{\partial \log L}{\partial \beta} &= \sum_{i=1}^n \frac{(y_i-\beta x_i)(x_i)}{\theta x_i^2} = 0, \quad \hat{\beta} = \frac{1}{n}\sum\left(\frac{Y_i}{x_i}\right) \\
\frac{\partial \log L}{\partial \theta} &= \frac{-n}{2\theta} + \left[\sum_{i=1}^n \frac{(y_i-\beta x_i)^2}{2x_i^2}\right]/\theta^2 = 0, \\
\hat{\theta} &= \sum \frac{(Y_i - \hat{\beta}x_i)^2}{x_i^2}/n = \hat{\gamma}^2.
\end{aligned}$$

9.6.9 We wish to minimize

$$\sum_{j=1}^b \sum_{i=1}^a (x_{ij} - \mu_j)^2 = \sum\sum (x_{ij} - \bar{x}_{.j})^2 + \sum (a)(\bar{x}_{.j} - \mu_j)^2.$$

That is, we wish to minimize

$$\begin{aligned}
K(c,d) &= \sum_{j=1}^b \{\bar{x}_{.j} - c - d[j - (b+1)/2]\}^2 \\
&= b(\bar{x}_{..} - c)^2 + \sum_{j=1}^b \{\bar{x}_{.j} - \bar{x}_{..} - d[j-(b+1)/2]\}^2.
\end{aligned}$$

Clearly, we want $\hat{c} = \bar{X}_{..}$. Moreover, $\frac{\partial K}{\partial d} = 0$ yields

$$\hat{d} = \sum_{j=1}^b [j - (b+1)/2](\bar{X}_{.j} - \bar{X}_{..}) / \sum_{j=1}^b [j-(b+1)/2]^2.$$

The F-statistic (with one and $ab-2$ degrees of freedom) is

$$\frac{\hat{d}^2 \sum_{j=1}^b (a)[j-(b+1)/2]^2/1}{\sum\sum\{X_{ij} - \bar{X}_{..} - \hat{d}[j-(b+1)/2]\}^2/(ab-2)}.$$

9.6.11 Let $\boldsymbol{\theta} \in V$ then $\boldsymbol{\theta} = \alpha\mathbf{1} + \beta\mathbf{x}_c$, for some α and β, where $\mathbf{1}$ is an $n \times 1$ vector of ones.

(a). Note that

$$\begin{aligned} \|\mathbf{Y} - \theta\|^2 &= \|\mathbf{Y} - (\alpha + \beta\mathbf{x}_c)\|^2 \\ &= \sum_{i=1}^{n} \{Y_i - [\alpha + \beta(x_i - \overline{x})]\}^2. \end{aligned}$$

Hence, by the method of LS, $\widehat{\boldsymbol{\theta}} = \widehat{\alpha} + \widehat{\beta}$.

(c). For $\boldsymbol{\theta} \in V$, we have

$$\begin{aligned} \boldsymbol{\theta}'(\mathbf{Y} - \widehat{\boldsymbol{\theta}}) &= (\alpha + \beta\mathbf{x}_c)'(\mathbf{Y} - \widehat{\alpha}\mathbf{1} - \widehat{\beta}\mathbf{x}_c) \\ &= \alpha n\overline{y} - \alpha n\widehat{\alpha} - 0 + \beta\mathbf{x}_c'\mathbf{Y} - 0 - \beta\widehat{\beta}\mathbf{x}_c'\mathbf{x}_c \\ &= \alpha n\widehat{\alpha} - \alpha n\widehat{\alpha} + \beta\widehat{\beta}\mathbf{x}_c'\mathbf{x}_c - \beta\widehat{\beta}\mathbf{x}_c'\mathbf{x}_c = 0. \end{aligned}$$

For $\mathbf{v} \in V$, we have $\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta} \in V$. Furthermore from above, we obtain the identity

$$\begin{aligned} \|\mathbf{Y} - \boldsymbol{\theta}\|^2 &= \|\mathbf{Y} - \widehat{\boldsymbol{\theta}} + \widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}\|^2 \\ &= \|\mathbf{Y} - \widehat{\boldsymbol{\theta}}\|^2 + \|\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}\|^2, \end{aligned}$$

for all $\boldsymbol{\theta} \in V$. In particular the identity is true for $\boldsymbol{\theta} = \mathbf{0}$ whose substitution in the above identity shows that the angle between $\widehat{\boldsymbol{\theta}}$ and $\mathbf{Y} - \widehat{\boldsymbol{\theta}}$ is a right angle.

9.6.14 We have

$$\begin{aligned} E(\widehat{\boldsymbol{\beta}}) &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}\boldsymbol{\beta} = \boldsymbol{\beta}, \\ \text{Var}(\widehat{\boldsymbol{\beta}}) &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\sigma^2\mathbf{I}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}. \end{aligned}$$

9.6.16 The linear model for Y_i is

$$Y_i = \mu + e_i, \quad i = 1, 2, \ldots, n,$$

where $\text{Var}(e_i) = \gamma^2 x_i^2$. Let $Z_i = Y_i/x_i$. Then the model for z_i is

$$Z_i = \mu\frac{1}{x_i} + e_i^*, \quad i = 1, 2, \ldots, n,$$

where $\text{Var}(e_i) = \gamma^2$. Now obtain the LS fit of μ,

$$\widehat{\mu} = \frac{\sum_{i=1}^{n} Z_i/x_i}{\sum_{i=1}^{n} 1/x_i^2},$$

with the corresponding estimate of variance,

$$\widehat{\gamma}^2 = \frac{\sum_{i=1}^{n}(Z_i - \widehat{\mu})^2}{n-1}.$$

It follows that $(n-1)\widehat{\gamma}^2/\gamma^2$ has a $\chi^2(n-1)$ distribution, from which it is easy to construct a test of $H_0 : \gamma = 1$.

9.7.1

$$\begin{aligned}
\sum(X_i-\bar{X})(Y_i-\bar{Y}) &= \sum X_iY_i - \bar{X}\sum Y_i - \bar{Y}\sum X_i + n\bar{X}\bar{Y} \\
&= \sum X_iY_i - \bar{X}(n\bar{Y}) - \bar{Y}(n\bar{X}) + n\bar{X}\bar{Y} \\
&= \sum X_iY_i - n\bar{X}\bar{Y}.
\end{aligned}$$

9.7.4 Here T has the t-distribution with $(n-2)$ df; that is,

$$h(t) = \frac{\Gamma[(n-1)/2]}{\sqrt{\pi(n-2)}\Gamma[(n-2)/2]}\frac{1}{[1+t^2/(n-2)]^{(n-1)/2}},$$

$-\infty < t < \infty$. Since

$$\begin{aligned}
t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}, \quad \frac{dt}{dr} &= \frac{\sqrt{1-r^2}\sqrt{n-2} - r\sqrt{n-2}\left(\frac{1}{2}\right)(1-r^2)^{-1/2}(-2r)}{1-r^2} \\
&= \frac{\sqrt{n-2}[(1-r^2)+r^2]}{(1-r^2)^{3/2}} = \frac{\sqrt{n-2}}{(1-r^2)^{3/2}},
\end{aligned}$$

we have

$$\begin{aligned}
g(r) &= \frac{\Gamma[(n-1)/2](1-r^2)^{(n-1)/2}}{\sqrt{\pi(n-2)}\Gamma[(n-2)/2]}\frac{\sqrt{n-2}}{(1-r^2)^{3/2}}, \quad -1 < r < 1 \\
&= \frac{\Gamma[(n-1)/2]}{\sqrt{\pi}\Gamma[(n-2)/2]}(1-r^2)^{(n-4)/2}, \quad -1 < r < 1.
\end{aligned}$$

9.7.5 We know that both $\overline{X} \to \mu_X$ and $\widehat{\sigma}_X \to \sigma_X$ in probability. By the Weak Law of Large Numbers

$$\frac{1}{n}\sum_{i=1}^{n} X_iY_i \xrightarrow{P} E[XY].$$

Putting these results together, it follows that $r \to \rho$ in probability.

9.8.2

$$\begin{aligned}
\boldsymbol{AVA} &= \begin{bmatrix} \frac{1}{\sigma_1^2(1-\rho^2)} & \frac{-\rho}{\sigma_1\sigma_2(1-\rho^2)} \\ \frac{-\rho}{\sigma_1\sigma_2(1-\rho^2)} & \frac{1}{\sigma_2^2(1-\rho^2)} \end{bmatrix} \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} \begin{bmatrix} \frac{1}{\sigma_1^2(1-\rho^2)} & \frac{-\rho}{\sigma_1\sigma_2(1-\rho^2)} \\ \frac{-\rho}{\sigma_1\sigma_2(1-\rho^2)} & \frac{1}{\sigma_2^2(1-\rho^2)} \end{bmatrix} \\
&= \frac{1}{(1-\rho^2)^2}\begin{bmatrix} 1-\rho^2 & 0 \\ 0 & 1-\rho^2 \end{bmatrix} \begin{bmatrix} \frac{1}{\sigma_1^2} & \frac{-\rho}{\sigma_1\sigma_2} \\ \frac{-\rho}{\sigma_1\sigma_2} & \frac{1}{\sigma_2^2} \end{bmatrix} = \boldsymbol{A}
\end{aligned}$$

Hence $\boldsymbol{X}'\boldsymbol{AX}$ is $\chi^2(2, \boldsymbol{\mu A \mu})$. Since $\boldsymbol{\mu}'\boldsymbol{A\mu}$ is positive definite, $\mu_1 = \mu_2 = 0$ is a necessary and sufficient condition for the noncentrality to equal zero.

9.8.3 It is easy to see that $\boldsymbol{A}^2 = \boldsymbol{A}$ and $\text{tr}(\boldsymbol{A}) = 2$. Moreover $\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}/8$ equals, when $\boldsymbol{x}' = (4, 4, 4)$,

$$[(4)(1/2)(16) + 16]/8 = 6;$$

so we have that the quadratic form is $\chi^2(2, 6)$.

9.8.5 For Parts (a) and (b), let $\mathbf{X}' = (X_1, X_2, \ldots, X_n)$. Note that

$$\text{Var}(\mathbf{X}) = \sigma^2[\rho\mathbf{J} + (1-\rho)\mathbf{I}],$$

where $\mathbf{J}$ is the $n \times n$ matrix of all ones, which can be written as $\mathbf{J} = \mathbf{1}\mathbf{1}'$, and $\mathbf{1}$ is a $n \times 1$ vector of ones.

(a). Note that $\overline{X} = \mathbf{1}'\mathbf{X}$. Hence,

$$\begin{aligned} \text{Var}(\overline{X}) &= \frac{\sigma^2}{n^2}\mathbf{1}'\left[\rho\mathbf{J} + (1-\rho)\mathbf{I}\right]\mathbf{1} \\ &= \frac{\sigma^2}{n^2}\left[\rho n^2 + (1-\rho)n\right] \\ &= \sigma^2\left[\rho + \frac{1-\rho}{n}\right]. \end{aligned}$$

(b). Note that

$$(n-1)S^2 = \mathbf{X}'\left(\mathbf{I} - \frac{1}{n}\mathbf{J}\right)\mathbf{X}.$$

Hence, using Theorem 9.8.1,

$$\begin{aligned} E[(n-1)S^2] &= E\left[\mathbf{X}'\left(\mathbf{I} - \frac{1}{n}\mathbf{J}\right)\mathbf{X}\right] \\ &= \text{tr}\left(\mathbf{I} - \frac{1}{n}\mathbf{J}\right)\boldsymbol{\Sigma} + \mu^2\mathbf{1}'\left(\mathbf{I} - \frac{1}{n}\mathbf{J}\right)\mathbf{1} \\ &= \text{tr}\left(\mathbf{I} - \frac{1}{n}\mathbf{J}\right)\sigma^2\left[\rho\mathbf{1}\mathbf{1}' + (1-\rho)\mathbf{I}\right] + 0 \\ &= \sigma^2\text{tr}\left[0 + (1-\rho)\left(\mathbf{I} - \frac{1}{n}\mathbf{J}\right)\right] \\ &= \sigma^2(1-\rho)(n-1). \end{aligned}$$

Hence, $E[S^2/(1-\rho)] = \sigma^2$.

9.8.8 In the hint, take $\boldsymbol{\Gamma}$ to be the matrix of eigenvectors such that $\boldsymbol{\Gamma}'\boldsymbol{\Lambda}\boldsymbol{\Gamma}$ is the spectral decomposition of $\mathbf{A}$.

9.8.10 Let $\boldsymbol{\Gamma}'\boldsymbol{\Lambda}\boldsymbol{\Gamma}$ be the spectral decomposition of $\mathbf{A}$. In this problem, $\boldsymbol{\Lambda}^2 = \boldsymbol{\Lambda}$ because the diagonal elements of $\boldsymbol{\Lambda}$ are 0s and 1s. Then because $\boldsymbol{\Gamma}$ is orthogonal,

$$\mathbf{A}^2 = \boldsymbol{\Gamma}'\boldsymbol{\Lambda}\boldsymbol{\Gamma}\boldsymbol{\Gamma}'\boldsymbol{\Lambda}\boldsymbol{\Gamma} = \boldsymbol{\Gamma}'\boldsymbol{\Lambda}^2\boldsymbol{\Gamma} = \boldsymbol{\Gamma}'\boldsymbol{\Lambda}\boldsymbol{\Gamma} = \mathbf{A}.$$

9.9.1 The product of the matrices is *not* equal to the zero matrix. Hence they are dependent.

9.9.3

$$\begin{bmatrix} a_1^2 & a_1a_2 & a_1a_3 & a_1a_4 \\ a_2a_1 & a_2^2 & a_2a_3 & a_2a_4 \\ a_3a_1 & a_3a_2 & a_3^2 & a_3a_4 \\ a_4a_1 & a_4a_2 & a_4a_3 & a_4 \end{bmatrix} \begin{bmatrix} 0 & 1/2 & 0 & 0 \\ 1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1/2 \\ 0 & 0 & -1/2 & 0 \end{bmatrix} = \mathbf{0}$$

requires, among other things, that

$$a_1^2 = 0, \;\; a_2^2 = 0, \;\; a_3^2 = 0, \;\; a_4^2 = 0.$$

Thus $a_1 = a_2 = a_3 = a_4 = 0$.

9.9.4 Yes, $\boldsymbol{A} = \boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}$ and $\bar{X}^2$ and independent. The matrix of $\bar{X}^2$ is $(1/n)^2\boldsymbol{P}$. So $\boldsymbol{A}\boldsymbol{P} = \mathbf{0}$ means that the sum of each row (column) of $\boldsymbol{A}$ must equal zero.

9.9.5 The joint mgf is

$$E[\exp(t_1Q_1+t_2Q_2+\cdots+t_kQ_k)] = |\boldsymbol{I}-2t_1\sigma^2\boldsymbol{A}_1-2t_2\sigma^2\boldsymbol{A}_2-\cdots-2t_k\boldsymbol{A}_k|^{-1/2}.$$

The preceding can be proved by following Section 9.9 of the text. Now $E[\exp(t_iQ_i)] = |\boldsymbol{I} - 2t_i\sigma^2\boldsymbol{A}_i|^{-1/2}, i = 1, 2, \ldots, k$. If $\boldsymbol{A}_i\boldsymbol{A}_j = 0, i \neq j$ (which means pairwise independence), we have $\prod_{i=1}^k(\boldsymbol{I} - 2t_i\sigma^2\boldsymbol{A}_i) = \boldsymbol{I} - 2t_1\sigma^2\boldsymbol{A}_1 - \cdots - 2t_k\sigma^2\boldsymbol{A}_k$. The determinant of the product of several square matrices of the same order is the product of the determinants. Thus $\prod_{i=1}^k|\boldsymbol{I}-2t_i\sigma^2\boldsymbol{A}_i| = |I - 2t_1\sigma^2\boldsymbol{A}_1 - \cdots - 2t_k\sigma^2\boldsymbol{A}_k|$ which is a necessary and sufficient condition for mutual independence of $Q_1, Q_2, \ldots, Q_k$.

9.9.6 If $\boldsymbol{b}'\boldsymbol{X}$ and $\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}$ are independent, then $\boldsymbol{b}'\boldsymbol{A} = \mathbf{0}$ and thus $(\boldsymbol{b}\boldsymbol{b}')\boldsymbol{A} = \mathbf{0}$ which implies that $\boldsymbol{X}'\boldsymbol{b}\boldsymbol{b}'\boldsymbol{X}$ and $\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}$ and independent. Conversely, if the two quadratic forms are independent, the $(\boldsymbol{b}\boldsymbol{b}')\boldsymbol{A} = \mathbf{0}$ and $(\boldsymbol{b}'\boldsymbol{b})\boldsymbol{b}'\boldsymbol{A} = 0$. Because $\boldsymbol{b}'\boldsymbol{b}$ is a nonzero scalar, we have $\boldsymbol{b}'\boldsymbol{A} = \mathbf{0}$ which implies the independence of $\boldsymbol{b}'$ and $\boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}$.

9.9.7 Let $\boldsymbol{A}, \boldsymbol{A}_1, \boldsymbol{A}_2$ represent, respectively, the matrices of $Q, Q_1,$ and $, Q_2$. Let $\boldsymbol{L}'(\boldsymbol{A}_1 + \boldsymbol{A}_2)\boldsymbol{L} = \text{diag}\{\alpha_1, \ldots, \alpha_r, 0, \ldots, 0\}$ where r is the rank of $\boldsymbol{A}_1 + \boldsymbol{A}_2$. Since both Q_1 and Q_2 are nonnegative quadratic forms, then

(a) $\alpha_i > 0, \;\; i = 1, 2, \ldots, r$;

(b) $\boldsymbol{L}'(\boldsymbol{A}_1+\boldsymbol{A}_2)\boldsymbol{L}\boldsymbol{L}'\boldsymbol{A}\boldsymbol{L} = 0$ implies $\boldsymbol{L}'\boldsymbol{A}\boldsymbol{L} = \begin{bmatrix} 0 & 0 \\ 0 & \boldsymbol{B} \end{bmatrix}$ where $\boldsymbol{B}$ is $(n-r)$;

(c) $\boldsymbol{L}'\boldsymbol{A}_j\boldsymbol{L} = \begin{bmatrix} \boldsymbol{B}_j & 0 \\ 0 & 0 \end{bmatrix}$, where $\boldsymbol{B}_j$ is r by r, $j = 1, 2$. Thus $\boldsymbol{L}'\boldsymbol{A}_j\boldsymbol{L}\boldsymbol{L}'\boldsymbol{A}\boldsymbol{L} = 0$ and $\boldsymbol{A}_j\boldsymbol{A} = 0, \; j = 1, 2$.

9.9.10 (a) Because the covariance matrix is $\sigma^2\boldsymbol{I}$ and thus all of the correlation coefficients are equal to zero.

(b) The P linear forms $\hat{\boldsymbol{\beta}}$ have a p-variate normal distribution with mean matrix $(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta} = \boldsymbol{\beta}$ and covariance matrix

$$(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'(\sigma^2)\boldsymbol{I}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1} = \sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-1}.$$

(c) Write the left side as

$$[(\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}) + \boldsymbol{X}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})]'[(\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}) + \boldsymbol{X}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})]$$

and carry out the necessary algebra to show that this equals the right side. It is helpful to note that

$$(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})\boldsymbol{X}'(\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}) = (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})'(\boldsymbol{X}'\boldsymbol{Y} - \boldsymbol{X}'\boldsymbol{Y}) = \mathbf{0}$$

(d)

$$(1/\sigma^2)(\boldsymbol{X}'\boldsymbol{X})\sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-1}(1/\sigma^2)(\boldsymbol{X}'\boldsymbol{X}) = (1/\sigma^2)(\boldsymbol{X}'\boldsymbol{X}).$$

Chapter 10

Nonparametric and Robust Statistics

10.1.2 The cdf of $X - a$ is $F(t + a)$, where $F(t)$ is the cdf of X. By symmetry, we have that the cdf of $-(X - a)$ is

$$P[-(X - a) \leq t] = P[X \geq a - t] = P[X \leq a + t] = F(t + a).$$

10.1.3 See Section 5.9 on bootstrap procedures.

10.1.4 Part (b): For property (i), let $Y = aX$. Then

$$F_Y(t) = F_X(t/a).$$

It is easy to show that

$$F_Y^{-1}(u) = aF_X^{-1}(u),$$

from which we have immediately that

$$\begin{aligned} F_Y^{-1}(3/4) &= aF_X^{-1}(3/4) \\ F_Y^{-1}(1/4) &= aF_X^{-1}(1/4). \end{aligned}$$

Thus $\xi_{Y,3/4} - \xi_{Y,1/4} = a(\xi_{X,3/4} - \xi_{X,1/4})$.

10.2.3 If students do not have access to a computer, then have them do normal (Central Limit Theorem) approximations.

(a). The level of the test is

$$P_{H_0}[S \geq 16] = P[\text{bin}(25, 1/2) \geq 16] = 0.1148.$$

(b). The probability of success here is

$$p = P[X > 0] = P[Z > -.5] = 0.6915.$$

Hence, the power of the sign test is

$$P_{0.6915}[S \geq 16] = P[\text{bin}(25, 0.6915) \geq 16] = 0.7836.$$

(c). To obtain the test, solve for k in the equation

$$0.1148 = P_{H_0}[\overline{X}/(1/\sqrt{25}) \geq k] = P[Z \geq k],$$

where Z has a standard normal distribution. The solution is $k = 1.20$. The power of the this test to detect 0.5 is

$$P_{\mu=0.5}[\overline{X}/(1/\sqrt{25}) \geq 1.20] = P[Z \geq 1.20 - (.5/(1/\sqrt{25}))] = 0.9032.$$

10.2.4 Recall that

$$\tau_{X,S} = \frac{1}{2f(\xi_{X,0.5})}.$$

We shall show that Properties (i) and (ii) on page 518 are true. For (i), let $Y = aX$, $a > 0$. First, $f_Y(t) = (1/a)f_X(t/a)$. Then, since the median is a location parameter (functional), $\xi_{Y,0.5} = a\xi_{X,0.5}$. Hence,

$$\tau_{Y,S} = \frac{1}{2f_Y(\xi_{Y,0.5})} = \frac{1}{2(1/a)f_X(a\xi_{X,0.5}/a)} = a\tau_{X,S}.$$

For (ii), let $Y = X + b$. Then $f_Y(t) = f_X(t - b)$. Also, since the median is a location parameter (functional), $\xi_{Y,0.5} = \xi_{X,0.5} + b$. Hence,

$$\tau_{Y,S} = \frac{1}{2f_Y(\xi_{Y,0.5})} = \frac{1}{2f_X(\xi_{X,0.5} + b - b)} = \tau_{X,S}.$$

10.2.8 The t-test rejects H_0 in favor of H_1 if $\overline{X}/(\sigma/\sqrt{n}) > z_\alpha$.

(a). The power function is

$$\gamma_t(\theta) = P_\theta\left[\frac{\overline{X}}{\sigma/\sqrt{n}} > z_\alpha\right] = 1 - \Phi\left[z_\alpha - \frac{\sqrt{n}\theta}{\sigma}\right].$$

(b). Hence,

$$\gamma_t'(\theta) = \phi\left[z_\alpha - \frac{\sqrt{n}\theta}{\sigma}\right]\frac{\sqrt{n}}{\sigma} > 0.$$

(c). Here $\theta_n = \delta/\sqrt{n}$. Thus,

$$\gamma_t(\delta/\sqrt{n}) = 1 - \Phi\left[z_\alpha - \frac{\delta}{\sigma}\right].$$

(d). Write $\theta^* = \sqrt{n}\theta^*/\sqrt{n}$. Then we need to solve the following equation for n:

$$\gamma^* = \gamma_t(\theta^*) = 1 - \Phi\left[z_\alpha - \frac{\sqrt{n}\theta^*}{\sigma}\right].$$

After simplification, we get

$$\sqrt{n} = \frac{(z_\alpha - z_{\gamma^*})\sigma}{\theta^*}.$$

10.3.1 Expanding the product, we obtain

$$m_T(s) = \frac{1}{8}\left[e^{-6s} + e^{-4s} + e^{-2s} + e^{0} + e^{2s} + e^{4s} + e^{6s}\right],$$

from which the distribution can be read.

10.3.4 Property (1) follows because all the terms in the sum are nonnegative and $R|v_i| > 0$, for all i. Property (2), follows because ranks of absolute values are invariant to a constant multiple. For the third property, following the hint we have

$$\begin{aligned}
\|\mathbf{u}+\mathbf{v}\| &\leq \sum_{i=1}^{n} R|u_i+v_i||u_i| + \sum_{i=1}^{n} R|u_i+v_i||v_i| \\
&\leq \sum_{j=1}^{n} j|u|_{(j)} + \sum_{j=1}^{n} j|v|_{(j)} \\
&= \sum_{j=1}^{n} j|u|_{i_j} + \sum_{j=1}^{n} j|v|_{i_j} \\
&= \sum_{i=1}^{n} R|u_i||u_i| + \sum_{i=1}^{n} R|v_i||v_i| = \|\mathbf{u}\| + \|\mathbf{v}\|,
\end{aligned}$$

where the permutation i_j denotes the permutation of the antiranks.

10.3.5 Note that the definition of $\widehat{\theta}$ should read

$$\widehat{\theta} = \text{Argmin}\|\mathbf{X} - \theta\|.$$

Write the norm in terms of antiranks; that is,

$$\|\mathbf{X} - \theta\| = \sum_{j=1}^{n} j|X_{i_j} - \theta|.$$

Taking the partial of the right-side with respect to θ, we get

$$\frac{\partial}{\partial\theta}\|\mathbf{X} - \theta\| = -\sum_{j=1}^{n} j\,\text{sgn}(X_{i_j} - \theta) = -\sum_{i=1}^{n} R|X_i - \theta|\text{sgn}(X_i - \theta).$$

Setting this equation to 0, we see that it is equivalent to the equation

$$2T^{+}(\theta) - \frac{n(n+1)}{2} = 0,$$

which leads to the Hodges-Lehmann estimate; see expression (10.3.10).

10.4.3 Write

$$\sqrt{n}(\overline{Y} - \overline{X} - \Delta) = \sqrt{\frac{n}{n_2}}\sqrt{n_2}(\overline{Y} - \mu_Y) - \sqrt{\frac{n}{n_1}}\sqrt{n_1}(\overline{X} - \mu_X).$$

By the Central Limit Theorem, the terms on the right-side converge in distribution to $N(0, \sigma^2/\lambda_2)$ and $N(0, \sigma^2/\lambda_1)$ distributions, respectively. Using independence between the samples leads to the asymptotic distribution given in expression (10.4.28).

10.4.4 From the asymptotic distribution of U, we obtain the equation

$$\begin{aligned} \frac{\alpha}{2} &= P_\Delta[U(\Delta) \leq c] = P_\Delta[U(\Delta) \leq c + (1/2)] \\ &\doteq P\left[Z \leq \left\{(c + (1/2) - (n_1 n_2/2))/\sqrt{n_1 n_2 (n+1)/12}\right\}\right]. \end{aligned}$$

Setting the term in braces to $-z_{\alpha/2}$ yields the desired result.

10.4.5 Using $\Delta > 0$, we get the following implication which implies that $F_Y(y) \leq F_X(y)$:

$$Y \leq y \Leftrightarrow X + \Delta \leq y \Leftrightarrow X \leq y - \Delta \Rightarrow X \leq y.$$

10.5.3 The value of s_a^2 for Wilcoxon scores is

$$\begin{aligned} s_a^2 &= 12 \sum_{i=1}^{n} \left[\frac{i}{n+1} - \frac{1}{2}\right]^2 \\ &= \frac{12}{(n+1)^2} \left\{\sum_{i=1}^{n} i^2 - (n+1) \sum_{i=1}^{n} i + \frac{n(n+1)^2}{4}\right\} \\ &= \frac{n(n-1)}{n+1}. \end{aligned}$$

10.5.5 Use the change of variables $u = \Phi(x)$ to obtain

$$\begin{aligned} \int_0^1 \Phi^{-1}(u)\, du &= \int_{-\infty}^{\infty} x\phi(x)\, dx = 0 \\ \int_0^1 (\Phi^{-1}(u))^2\, du &= \int_{-\infty}^{\infty} x^2 \phi(x)\, dx = 1. \end{aligned}$$

10.5.10 For this problem

$$\tau_\varphi^{-1} = \int_0^1 \Phi^{-1}(u) \left\{-\frac{f'(F^{-1}(u))}{f(F^{-1}(u))}\right\} du.$$

Without loss of generality assume that $\mu = 0$. Then $f(x) = (1/\sqrt{2\pi}\sigma) \exp\{-x^2/2\sigma^2\}$. It follows that

$$\frac{f'(x)}{f(x)} = -\frac{x}{\sigma^2}.$$

Furthermore, because $F(t) = \Phi(t/\sigma)$ we get $F^{-1}(u) = \sigma\Phi^{-1}(u)$. Substituting this into the expression which defines τ_φ^{-1}, we obtain $\tau_\varphi^{-1} = \sigma^{-1}$.

10.5.12 The Riemann approximation yields

$$1 = \int_0^1 \varphi^2(u)\, du \doteq \sum_{i=1}^{n} \varphi^2\left(\frac{i}{n+1}\right)\frac{1}{n}.$$

10.5.15 Let $F(t)$ be the common cdf of X_i.

(a). Without loss of generality assume that $\theta = 0$. Let $0 < u < 1$ be an arbitrary but fixed u. Let $t = F^{-1}(1-u)$. Then

$$\varphi(1-u) = -\frac{f'(F^{-1}(1-u))}{f(F^{-1}(1-u))} = -\frac{f'(t)}{f(t)}. \qquad (10.0.1)$$

But $F^{-1}(1-u) = t$ implies, by symmetry about 0, that $u = 1 - F(t) = F(-t)$. Because $f'(t)$ and $f(t)$ are odd and even functions, respectively, we have

$$-\varphi(u) = \frac{f'(F^{-1}(u))}{f(F^{-1}(u))} = \frac{f'(-t)}{f(-t)} = -\frac{f'(t)}{f(t)}. \qquad (10.0.2)$$

By (10.0.1) and (10.0.2) the result follows.
Also, by (10.5.40), $\varphi(1/2) = -\varphi(1/2)$. So $\varphi(1/2) = 0$

(b). Since $(u+1)/2 > 1/2$ and $\varphi(u)$ is nondecreasing

$$\varphi^+(u) = \varphi((u+1)/2) \geq \varphi(0) = 0.$$

(e). Let i_j denote the permutation of antiranks. Then we can write W_{φ^+} as

$$W_{\varphi^+} = \sum_{j=1}^{n} \text{sgn}(X_{i_j}) a^+(i_j).$$

By the discussion on page 532, $\text{sgn}(X_{i_j})$ are iid with pmf $p(-1) = p(1) = 1/2$. Hence, the statistic W_{φ^+} is distribution-free under H_0.
The above expression can be used to find the null mean and variance of W_{φ^+} and to state its asymptotic distribution.

10.6.1 The following R code (driver and 4 functions) computes the four tests statistics based on the four respective score functions given in Exercise 10.6.1. In particular, the code returns the variances of the four tests. For sample sizes $n_1 = n_2 = 15$, the variances are: 2.419, 7.758, 2.419, and 2.419.

```
drive4 = function(x,y){
      n1 = length(x)
      n2 = length(y)
      n = n1 + n2
      cb = (1:n)/(n+1)
```

```
      const = (n1*n2)/(n*(n-1))
      p1 = phi1(cb)
      var1 = const*sum(p1^2)
      p2 = phi2(cb)
      var2 = const*sum(p2^2)
      p3 = phi1(cb)
      var3 = const*sum(p3^2)
      p4 = phi1(cb)
      var4 = const*sum(p4^2)
      vars = c(var1,var2,var3,var4)
      allxy = c(x,y)
      rall = rank(allxy)/(n+1)
      ind = c(rep(0,n1),rep(1,n2))
      s1 = sum(ind*phi1(rall))
      s2 = sum(ind*phi2(rall))
      s3 = sum(ind*phi3(rall))
      s4 = sum(ind*phi4(rall))
      tests = c(s1,s2,s3,s4)
      ztests = tests/sqrt(vars)
      list(vars=vars,tests=tests,ztests=ztests)
}

phi1 = function(u){
    phi1 = 2*u - 1
    phi1
}

phi2 = function(u){
    phi2 = sign(2*u - 1)
    phi2
}

phi3 = function(u){
    n = length(u)
    phi3 = rep(0,n)
    for (i in 1:n){
         if(u[i] <= .25){phi3[i] = 4*u[i] - 1}
         if(u[i] > .75){phi3[i] = 4*u[i] - 3}
    }
    phi3
}

phi4 = function(u){
    n = length(u)
    phi4 = rep(.5,n)
    for (i in 1:n){
```

```
        if(u[i] <= .50){phi4[i] = 4*u[i] - (3/2)}
    }
    phi4
}
```

10.6.2 Based on the above code, the standardized test statistics for the 4 respective scores are: 1.555, 1.077, 0.850, and 0.839.

10.7.1 Note that the ranks are invariant to constant shifts. From Model (10.7.1), under β we have,

$$P_\beta(Y_i \le t) = P[\varepsilon \le t - \alpha - \beta(x_i - \overline{x})]. \tag{10.0.3}$$

Under $\beta = 0$, we have

$$P_0(Y_i + \beta(x_i - \overline{x}) \le t) = P[\varepsilon + \alpha + \beta(x_i - \overline{x}) \le t],$$

which is the same as (10.0.3).

10.7.4 The power function is

$$\gamma(\beta) = P_\beta[T_\varphi(0) \ge c_\alpha] = P_0[T_\varphi(-\beta) \ge c_\alpha].$$

Suppose $\beta_1 < \beta_2$ then, since T_φ is nonincreasing, $T_\varphi(-\beta_1) \le T_\varphi(-\beta_2)$. This leads to the implication

$$T_\varphi(-\beta_1) \ge c_\alpha \Rightarrow T_\varphi(-\beta_2) \ge c_\alpha.$$

From which we get, $\gamma(\beta_1) \le \gamma(\beta_2)$.

10.7.5 As in the last exercise, the power function is

$$\begin{aligned} \gamma(\beta_n) &= P_{\beta_n}\left[T_\varphi(0) \ge z_\alpha \sigma_{T_\varphi}\right] \\ &= P_{\beta_n}\left[\frac{T_\varphi(0) - E_{\beta_n}[T_\varphi(0)]}{\sigma_{T_\varphi}} \ge z_\alpha - \frac{E_{\beta_n}[T_\varphi(0)]}{\sigma_{T_\varphi}}\right]. \end{aligned}$$

In the last expression, the random variable on the leftside is approximately $N(0,1)$ and, using the discussion on page 569, the right-side reduces to $z_\alpha - \beta_1 c_T$. These approximations can be made rigorous in a more advanced course.

10.8.1 Write τ as

$$\tau = 2P[\mathrm{sgn}[(X_1 - X_2)(Y_1 - Y_2)]] - 1.$$

It is easy to show that the right-side is between 0 and 1.

10.8.3 The following results were obtained at the site `www.stat.wmich.edu/slab/RGLM`.

Procedure	$\widehat{\alpha}$	(SE)	$\widehat{\beta}$	(SE)
LS	206.2	(13.01)	0.0151	(0.0055)
Wilcoxon	211.0	(2.59)	0.0098	(0.0011)

The obvious outlier spoiled the LS fit and its standard errors.

10.8.5 Part (a). Note that the scores are centered and $\sum_{i=1}^{n} a^2(i) = s_a^2$. Hence, r_a is a correlation coefficient on the pairs $(a(R(X_i)), a(R(Y_i)))$, $i = 1, 2, \ldots, n$.

10.8.9 As with the other rank score correlations, $\sqrt{n-1} r_N$ has a null asymptotic $N(0, 1)$ distribution. The following R-code computes r_N and its corresponding z-test statistic:

```
rn1089 = function(){
data=matrix(scan("olymp3.dat"),ncol=2,byrow=T)
x=data[,1]
y=data[,2]
n=length(x)
rx=rank(x)/(n+1)
ry=rank(y)/(n+1)
sx=qnorm(rx)
sy=qnorm(ry)
sa2 = sum(sx^2)
rn = sum(sx*sy)/sa2
zn = sqrt(n-1)*rn
list(rn=rn,zn=zn)
}
```

10.9.1 With $\boldsymbol{\eta} = \theta \mathbf{1}$, write

$$\|\mathbf{Y} - \boldsymbol{\eta}\|_{LS}^2 = \sum_{i=1}^{n} (y_i - \theta)^2.$$

Now take the partial derivative with respect to θ, set the result to 0, and solve for θ. This yields $\widehat{\theta} = \overline{y}$ and, hence, $\widehat{\boldsymbol{\eta}} = \overline{y}\mathbf{1}$.

10.9.4 Note that

$$F_{x,\epsilon}(t) - F_X(t) = \begin{cases} \epsilon[-F_X(t)] & \text{if } t < x \\ \epsilon[1 - F_X(t)] & \text{if } t \geq x. \end{cases}$$

In either case the expression in brackets is less that or equal to 1 in absolute value.

10.9.7 Let $V(F)$ denote the variance functional of the cdf $F(t)$.

(a). Let F_n denote the empirical cdf of $Y_1 - \overline{y}, \ldots, Y_n - \overline{Y}$. Then $V(F_n)$ solves

$$0 = \int_{-\infty}^{\infty} [t^2 - V(F_n)]\, dF_n(t) = \frac{1}{n} \sum_{i=1}^{n} (Y_i - \overline{Y})^2 - V(F_n).$$

(b). The functional at the contaminated cdf solves

$$\begin{aligned} 0 &= \int_{-\infty}^{\infty} [t^2 - V(F_{y,\epsilon})]\, dF_{y,\epsilon}(t) \\ &= (1-\epsilon)\int_{-\infty}^{\infty} [t^2 - V(F_{y,\epsilon})]\, dF(t) + \epsilon \int_{-\infty}^{\infty} [t^2 - V(F_{y,\epsilon})]\, d_{\Delta_y}(t) \end{aligned}$$

Taking the partial of both sides with respect to ϵ, we get

$$\begin{aligned} 0 &= -\int_{-\infty}^{\infty} [t^2 - V(F_{y,\epsilon})]\, dF(t) - (1-\epsilon)\int_{-\infty}^{\infty} dF(t)\frac{\partial V}{\partial \epsilon} \\ &\quad + \int_{-\infty}^{\infty} [t^2 - V(F_{y,\epsilon})]\, d_{\Delta_y}(t) + \epsilon \frac{\partial B}{\partial \epsilon}, \end{aligned}$$

where the last partial derivative is not needed. Evaluation of the last expression at $\epsilon = 0$ yields

$$\left.\frac{\partial V}{\partial \epsilon}\right|_{\epsilon=0} = y^2 - \sigma^2.$$

10.9.9 Recall that θ is the true median. We then have

$$E[IF(Y;\widehat{\theta}_{L_1})] = \frac{1}{2f_Y(\theta)}\frac{1}{2} - \frac{1}{2f_Y(\theta)}\frac{1}{2} = 0$$

and

$$E[IF^2(Y;\widehat{\theta}_{L_1})] = \frac{1}{4f_Y^2(\theta)}\frac{1}{2} + \frac{1}{4f_Y^2(\theta)}\frac{1}{2} = \frac{1}{4f_Y^2(\theta)}.$$

10.9.12 Part (a): Note that

$$\begin{aligned} \frac{\partial}{\partial \beta}\sum_{i=1}^{n} a(i)(Y - \mathbf{x}_c'\boldsymbol{\beta})_{(i)} &= -\sum_{i=1}^{n} a(i)[\mathbf{x}_c]_{(i)} \\ &= -\sum_{i=1}^{n} a((R(Y_i - \mathbf{x}_{ci}'\boldsymbol{\beta}))\mathbf{x}_{ci}, \end{aligned}$$

where the notation $[\mathbf{x}_c]_{(i)}$ means the $\mathbf{x}_c$ associated with $(Y - \mathbf{x}_c'\boldsymbol{\beta})_{(i)}$.

Part (c): From the Wilcoxon normal equations determined in the last exercise, it is clear that $\widehat{\boldsymbol{\beta}}_W$ is chosen so that the vector of ranked-scores of the residuals, $a(R(\mathbf{Y} - \mathbf{X}_c\widehat{\boldsymbol{\beta}}_W))$, is orthogonal to the range of of the matrix $\mathbf{X}_c$.

10.9.15 Start with the right-side, i.e.,

$$
\begin{aligned}
\sum_{i,j} |v_i - v_j| &= \sum_{i,j} |v_{(i)} - v_{(j)}| \\
&= 2\sum_{i<j} |v_{(i)} - v_{(j)}| \\
&= 2\sum_{i<j} (v_{(j)} - v_{(i)}) \\
&= 2\left\{\sum_{j=2}^{n}\sum_{i=1}^{j-1} v_{(j)} - \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} v_{(i)}\right\} \\
&= 2\left\{\sum_{j=2}^{n}(j-1)v_{(j)} - \sum_{i=1}^{n-1}(n-i)v_{(i)}\right\} \\
&= 4\sum_{j=2}^{n}\left(j - \frac{n+1}{2}\right) v_{(j)} \\
&= \frac{2}{\sqrt{3}(n+1)}\sum_{j=1}^{n}\sqrt{12}\left(\frac{j}{n+1} - \frac{1}{2}\right) v_{(j)}.
\end{aligned}
$$

10.9.16 Part (b). We know that $F(e_i)$ has an uniform$(0,1)$ distribution. Hence, since the scores are standardized,

$$
\begin{aligned}
E[\varphi(F(e_i))] &= \int_0^1 \varphi(u)\,du = 0 \\
E[\varphi^2(F(e_i))] &= \int_0^1 \varphi^2(u)\,du = 1.
\end{aligned}
$$

Chapter 11

Bayesian Statistics

11.2.2

$$\begin{aligned} k(\theta|x_1, x_2, \ldots, x_n) &\propto \theta^{\sum x_i}(1-\theta)^{n-\sum x_i}\theta^{\alpha-1}(1-\theta)^{\beta-1} \\ &= \theta^{\sum x_i+\alpha-1}(1-\theta)^{n-\sum x_i+\beta-1}, \end{aligned}$$

which is the pdf of a beta $(\alpha^* = \sum x_i + \alpha, \beta^* = n - \sum x_i + \beta)$ and, with $y = \sum x_i$, is the same as that of Example 11.2.2.

11.2.4 Considering Example 11.2.1, we know that the posterior distribution of the parameter, given the data, is gamma $[\alpha^* = \sum x_i + \alpha, \beta^* = \beta/(n\beta + 1)]$. With $Y = \sum X_i$ and square error loss, we want our Bayes estimator to be the conditional mean of the parameter, given the data. That is,

$$\alpha^*\beta^* = (Y+\alpha)\beta/(n\beta+1) = \frac{(Y/n)n + (\alpha\beta)(1/\beta)}{n + (1/\beta)},$$

which is the weighted average of $\overline{X} = Y/n$ and the prior mean of $\alpha\beta$.

11.2.6

$$\begin{aligned} k(\theta_1, \theta_2|y_1, y_2) &\propto \theta_1^{y_1}\theta_2^{y_2}(1-\theta_1-\theta_2)^{n-y_1-y_2}\theta_1^{\alpha_1-1}\theta_2^{\alpha_2-1}(1-\theta_1-\theta_2)^{\alpha_3-1} \\ &= \theta_1^{y_1+\alpha_1-1}\theta_2^{y_2+\alpha_2-1}(1-\theta_1-\theta_2)^{n-y_1-y_2+\alpha_3-1}, \end{aligned}$$

which is Dirichlet with $\alpha_1^* = y_1 + \alpha_1$, $\alpha_2^* = y_2 + \alpha_2$, $\alpha_3^* = n - y_1 - y_2 + \alpha_3$. The two conditional means are

$$\tfrac{y_1+\alpha_1}{n+\alpha_1+\alpha_2+\alpha_3} \text{ and } \tfrac{y_2+\alpha_2}{n+\alpha_1+\alpha_2+\alpha_3}.$$

11.2.8

(a) $$E\left[\left(\theta - \frac{10+Y}{45}\right)^2\right] = \left(\theta - \frac{10+30\theta}{45}\right)^2 + \left(\frac{1}{45}\right)^2 30\theta(1-\theta)$$

(a) $$E\left[\left(\theta - \frac{10+Y}{45}\right)^2\right] < \frac{\theta(1-\theta)}{30}$$

requires that

$$k(\theta) = \left(\frac{\theta}{3} - \frac{2}{9}\right)^2 - \frac{1}{54}\theta(1-\theta) < 0.$$

Find the two zeroes of $k(\theta)$, one of which is greater (less) than $2/3$.

11.2.9 The conditional pdf of the parameter, given $Y_4 = y_4$, is

$$h(\theta|y_4) \propto \left(\frac{4y_4^3}{\theta^4}\right)\left(\frac{2}{\theta^3}\right) \propto \frac{1}{\theta^7}, \quad y_4 < \theta \text{ and } 1 < \theta.$$

This means that

$$h(\theta|y_4) = \begin{cases} \frac{6}{\theta^7} & y_4 < 1 < \theta \\ \frac{6y_4^6}{\theta^7} & 1 < y_4 < \theta. \end{cases}$$

With the absolute value loss function, our Bayes decision is the median of the posterior distribution, namely

$$\int_m^\infty \frac{6}{\theta^7}\,d\theta = 1/2 \Rightarrow m = 2^{1/6} \text{ when } y_4 < 1,$$

and

$$\int_m^\infty \frac{6y_4^6}{\theta^7}\,d\theta = 1/2 \Rightarrow m = 2^{1/6}y_4 \text{ when } 1 < y_4.$$

11.3.2 The Bayes model is

$$\begin{aligned} X|\theta &\sim \Gamma(3, 1/\theta), \quad \theta > 0 \\ \Theta &\sim \Gamma(10, 2). \end{aligned}$$

(a). The posterior pdf simplifies to

$$k(\theta|\mathbf{x}) \propto \theta^{40-1}\exp\{-\theta[(1/2) + n\overline{x}]\},$$

which is the pdf of a $\Gamma(40, 1/[(1/2) + n\overline{x}])$ distribution.

(b). Squared error loss implies the Bayes estimate is the mean of the posterior; i.e., $40/[(1/2) + n\overline{x}]$.

(d). Note that $2[(1/2) + n\overline{x}]\Theta$ has a $\chi^2(80)$ distribution.

11.3.4 Let $\tau = u(\theta)$. By the chain rule we have

$$\frac{\partial \log f(x;\theta)}{\partial \tau} = \frac{\partial \log f(x;\theta)}{\partial \theta}\frac{\partial \theta}{\partial \tau}$$

Squaring both sides and taking expectations leads to

$$I(\tau) = I(\theta)\left(\frac{\partial\theta}{\partial\tau}\right)^2.$$

By the transformation rule the prior for τ is

$$h_2(\tau) = h(\theta)\left|\frac{\partial\theta}{\partial\tau}\right| \propto \sqrt{I(\theta)}\left|\frac{\partial\theta}{\partial\tau}\right| = \sqrt{I(\tau)}.$$

11.37 By Exercise 11.3.4, the Jeffreys prior is proportional to the square root of the information which, by Page 6.2.1, is $1/\sqrt{\theta(1-\theta)}$. Hence, the prior is a beta$(1/2, 1/2)$ distribution. For Part (b), the posterior pdf is

$$\begin{aligned} k(\theta|\mathbf{x}) &\propto \theta^{n\overline{x}}(1-\theta)^{n-n\overline{x}}\theta^{(1/2)-1}(1-\theta)^{(1/2)-1} \\ &\propto \theta^{n\overline{x}+(1/2)-1}(1-\theta)^{n-n\overline{x}+(1/2)-1}. \end{aligned}$$

Hence, the posterior distribution is beta$(n\overline{x}+(1/2), n-n\overline{x}+(1/2))$.

11.4.1 It is easy to show that the inverse conditional cdf of X given $Y = y$ is $F_{X|Y}^{-1}(u) = y - \log(1-u)$.

(a). Hence, the algorithm is of the form:

(0).	Generate U_1 and U_2 iid uniform$(0,1)$
(1).	Generate $Y = \log(1/(1-U_1)$
(2).	Generate $X = Y + \log(1/(1-U_2)$.

(b). For n large, generate $X_1, X_2, \ldots X_n$. Then $\overline{X}$ is an estimator of $E(X)$.

(c). The following R function will compute the algorithm.

```
condsim2<-function(nsims){
collect<-rep(0,nsims)
for(i in 1:nsims)
   {y<--log(1-runif(1))
    collect[i]<--log(1-runif(1))+y
   }
collect
}
```

11.4.3 Both marginal pdfs and the conditional pdf are given in the example.

(a). Use conditional expectation, i.e.

$$E(X) = E[E(X|Y)] = E(1+Y) = \frac{3}{2}.$$

(b) The cdf of X is

$$F_X(x) = 1 - 2e^{-x} + e^{-2x}, \quad x > 0.$$

For $0 < u < 1$, the inverse of this cdf is the solution of the equation

$$e^{-2x} - 2e^{-x} + (1-u) = 0.$$

This is a quadratic equation in e^{-x} with the solution $1-\sqrt{u}$, (the other solution cannot be true). This leads to the inverse of the cdf which is

$$F_X^{-1}(u) = \log[1/(1-\sqrt{u})].$$

Based on this, X can be generated by $\log[1/(1-\sqrt{U})]$, where U has a uniform $(0,1)$ distribution.

11.4.7 For this exercise a computer is not needed.

(a). The constant of proportionality K solves the equation

$$1 = K \int_0^1 y^{\alpha-1}(1-y)^{\beta-1} \left\{ \sum_{x=0}^{n} \binom{n}{x} y^x (1-y)^{n-x} \right\} dy,$$

which is easily determined to be $K = \Gamma(\alpha+\beta)/(\Gamma(\alpha)\Gamma(\beta))$.

(b) from the joint pdf, we have

$$f(x|y) \propto \binom{n}{x} y^x (1-y)^{n-x}.$$

Hence, $X|Y$ is binomial(n, Y). Likewise,

$$f(y|x) \propto y^{x+\alpha-1}(1-y)^{n-x+\beta-1};$$

so $Y|X$ is beta$(x+\alpha, n-x+\beta)$.

(c). The Gibbs sampler algorithm is: for $i = 1, 2, \ldots, m$

(1). Generate $Y_i|X_{i-1} \sim$ beta$(\alpha + X_{i-1}, n - X_{i-1} + \beta)$

(2). Generate $X_i|Y_i \sim$ binomial(n, Y_i) .

11.4.8 Here is R-code which runs the Gibbs sampler of the last exercise:

```
gibbser3 = function(alpha,beta,nt,m,n){
x0 = 1
yc = rep(0,m+n)
xc = c(x0,rep(0,m-1+n))
for(i in 2:(m+n)){yc[i] = rbeta(1,xc[i-1]+alpha,nt-xc[i-1]+beta)
              xc[i] = rbinom(1,nt,yc[i])}
y1=yc[1:m]
y2=yc[(m+1):(m+n)]
x1=xc[1:m]
x2=xc[(m+1):(m+n)]
list(y1 = y1,y2=y2,x1=x1,x2=x2)
}
```

To determine the mean of X, use the joint pdf to find that $E(X) = n(\alpha/(\alpha+\beta))$.

11.5.3 The Bayes model is

$$\begin{aligned} X|p &\sim \text{bin}(n,p), \quad 0 < p < 1 \\ p|\theta &\sim h(p|\theta) = \theta p^{\theta-1}, \quad \theta > 0 \\ \theta &\sim \Gamma(1,a), \quad a \text{ specified.} \end{aligned}$$

(a). The conditional pdf of p given y and θ is

$$g(p|y,\theta) \propto p^{y+\theta-1}(1-p)^{n-y+1-1},$$

which is the pdf of a beta$(y+\theta, n-y+1)$ distribution.

(b). The conditional pdf of θ given y and p is

$$g(\theta|y,p) \propto \theta \exp\{-\theta[(1/a) - \log p]\},$$

which is the pdf of a $\Gamma(2, [(1/a) - \log p]^{-1})$ distribution.

(c). The Gibbs sampler algorithm is for $i = 1, 2, \ldots, m$

$$\begin{array}{ll} (1). & \text{Generate } P_i|y, \Theta_{i-1} \sim \text{beta}(y + \Theta_{i-1}, n-y+1) \\ (2). & \text{Generate } \Theta_i|y, P_i \sim \Gamma\left(2, \left[\frac{1}{a} - \log P_i\right]^{-1}\right). \end{array}$$

11.5.5 Recall that

$$g(y,p|\theta) = g(y|p)g(p|\theta) = \binom{n}{y} p^{y+\theta-1}(1-p)^{n-y}\theta p^{\theta-1}.$$

Integrating out p, we have

$$\begin{aligned} g(y|\theta) &= \theta \int_0^1 \binom{n}{y} p^{y+\theta-1}(1-p)^{n-y+1-1}\,dp \\ &= \theta \binom{n}{y} \frac{\Gamma(y+\theta)\Gamma(n-y+1)}{\Gamma(y+\theta+n-y+1)}. \end{aligned}$$