INSTRUCTOR'S SOLUTIONS MANUAL

PROBABILITY AND STATISTICAL INFERENCE

NINTH EDITION

ROBERT V. HOGG

University of Iowa

Elliot A. Tanis

Hope College

Dale L. Zimmerman

University of Iowa

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 São 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: 978-0-321-91379-1
ISBN-10: 0-321-91379-5

www.pearsonhighered.com

PEARSON

Contents

Preface

This solutions manual provides answers for the even-numbered exercises in *Probability and Statistical Inference*, 9th edition, by Robert V. Hogg, Elliot A. Tanis, and Dale L. Zimmerman. Complete solutions are given for most of these exercises. You, the instructor, may decide how many of these solutions and answers you want to make available to your students. Note that the answers for the odd-numbered exercises are given in the textbook.

All of the figures in this manual were generated using *Maple*, a computer algebra system. Most of the figures were generated and many of the solutions, especially those involving data, were solved using procedures that were written by Zaven Karian from Denison University. We thank him for providing these. These procedures are available free of charge for your use. They are available for down load at http://www.math.hope.edu/tanis/. Short descriptions of these procedures are provided on the "Maple Card." Complete descriptions of these procedures are given in *Probability and Statistics: Explorations with MAPLE*, second edition, 1999, written by Zaven Karian and Elliot Tanis, published by Prentice Hall (ISBN 0-13-021536-8). You can download a copy of this manual at http://www.math.hope.edu/tanis/MapleManual.pdf.

Our hope is that this solutions manual will be helpful to each of you in your teaching.

If you find an error or wish to make a suggestion, send these to Elliot Tanis, `tanis@hope.edu`, and he will post corrections on his web page, http://www.math.hope.edu/tanis/.

R.V.H.
E.A.T.
D.L.Z.

Chapter 1

Probability

1.1 Properties of Probability

1.1-2 Sketch a figure and fill in the probabilities of each of the disjoint sets.

Let $A = \{\text{insure more than one car}\}$, $P(A) = 0.85$.

Let $B = \{\text{insure a sports car}\}$, $P(B) = 0.23$.

Let $C = \{\text{insure exactly one car}\}$, $P(C) = 0.15$.

It is also given that $P(A \cap B) = 0.17$. Since $A \cap C = \phi$, $P(A \cap C) = 0$. It follows that $P(A \cap B \cap C') = 0.17$. Thus $P(A' \cap B \cap C') = 0.06$ and $P(A' \cap B' \cap C) = 0.09$.

1.1-4 **(a)** $S = \{$HHHH, HHHT, HHTH, HTHH, THHH, HHTT, HTTH, TTHH, HTHT, THTH, THHT, HTTT, THTT, TTHT, TTTH, TTTT$\}$;

(b) **(i)** 5/16, **(ii)** 0, **(iii)** 11/16, **(iv)** 4/16, **(v)** 4/16, **(vi)** 9/16, **(vii)** 4/16.

1.1-6 **(a)** $P(A \cup B) = 0.4 + 0.5 - 0.3 = 0.6$;

(b)
$$\begin{aligned} A &= (A \cap B') \cup (A \cap B) \\ P(A) &= P(A \cap B') + P(A \cap B) \\ 0.4 &= P(A \cap B') + 0.3 \\ P(A \cap B') &= 0.1; \end{aligned}$$

(c) $P(A' \cup B') = P[(A \cap B)'] = 1 - P(A \cap B) = 1 - 0.3 = 0.7$.

1.1-8 Let A ={lab work done}, B ={referral to a specialist},

$P(A) = 0.41$, $P(B) = 0.53$, $P([A \cup B]') = 0.21$.

$$\begin{aligned} P(A \cup B) &= P(A) + P(B) - P(A \cap B) \\ 0.79 &= 0.41 + 0.53 - P(A \cap B) \\ P(A \cap B) &= 0.41 + 0.53 - 0.79 \;=\; 0.15. \end{aligned}$$

1.1-10
$$\begin{aligned} A \cup B \cup C &= A \cup (B \cup C) \\ P(A \cup B \cup C) &= P(A) + P(B \cup C) - P[A \cap (B \cup C)] \\ &= P(A) + P(B) + P(C) - P(B \cap C) - P[(A \cap B) \cup (A \cap C)] \\ &= P(A) + P(B) + P(C) - P(B \cap C) - P(A \cap B) - P(A \cap C) \\ &\quad + P(A \cap B \cap C). \end{aligned}$$

1.1-12 **(a)** 1/3; **(b)** 2/3; **(c)** 0; **(d)** 1/2.

1.1-14 $P(A) = \dfrac{2[r - r(\sqrt{3}/2)]}{2r} = 1 - \dfrac{\sqrt{3}}{2}.$

1.1-16 Note that the respective probabilities are $p_0,\ p_1 = p_0/4,\ p_2 = p_0/4^2, \cdots$.

$$\begin{aligned} \sum_{k=0}^{\infty} \frac{p_0}{4^k} &= 1 \\ \frac{p_0}{1 - 1/4} &= 1 \\ p_0 &= \frac{3}{4} \end{aligned}$$

$$1 - p_0 - p_1 = 1 - \frac{15}{16} = \frac{1}{16}.$$

1.2 Methods of Enumeration

1.2-2 **(a)** $(4)(5)(2) = 40$; **(b)** $(2)(2)(2) = 8$.

1.2-4 **(a)** $4\dbinom{6}{3} = 80$;

(b) $4(2^6) = 256$;

(c) $\dfrac{(4 - 1 + 3)!}{(4 - 1)!3!} = 20$.

1.2-6 $S =\{$ DDD, DDFD, DFDD, FDDD, DDFFD, DFDFD, FDDFD, DFFDD, FDFDD, FFDDD, FFF, FFDF, FDFF, DFFF FFDDF, FDFDF, DFFDF, FDDFF, DFDFF, DDFFF $\}$ so there are 20 possibilities.

1.2-8 $3 \cdot 3 \cdot 2^{12} = 36{,}864$.

1.2-10
$$\begin{aligned} \binom{n-1}{r} + \binom{n-1}{r-1} &= \frac{(n-1)!}{r!(n-1-r)!} + \frac{(n-1)!}{(r-1)!(n-r)!} \\ &= \frac{(n-r)(n-1)! + r(n-1)!}{r!(n-r)!} = \frac{n!}{r!(n-r)!} = \binom{n}{r}. \end{aligned}$$

1.2-12
$$\begin{aligned} 0 &= (1-1)^n = \sum_{r=0}^{n} \binom{n}{r} (-1)^r (1)^{n-r} = \sum_{r=0}^{n} (-1)^r \binom{n}{r}. \\ 2^n &= (1+1)^n = \sum_{r=0}^{n} \binom{n}{r} (1)^r (1)^{n-r} = \sum_{r=0}^{n} \binom{n}{r}. \end{aligned}$$

1.2-14 $\dbinom{10 - 1 + 36}{36} = \dfrac{45!}{36!9!} = 886{,}163{,}135.$

1.2-16 **(a)** $\dfrac{\dbinom{19}{3}\dbinom{52-19}{6}}{\dbinom{52}{9}} = \dfrac{102{,}486}{351{,}325} = 0.2917;$

(b) $\dfrac{\dbinom{19}{3}\dbinom{10}{2}\dbinom{7}{1}\dbinom{3}{0}\dbinom{5}{1}\dbinom{2}{0}\dbinom{6}{2}}{\dbinom{52}{9}} = \dfrac{7{,}695}{1{,}236{,}664} = 0.00622.$

1.3 Conditional Probability

1.3-2 **(a)** $\dfrac{1041}{1456}$;

(b) $\dfrac{392}{633}$;

(c) $\dfrac{649}{823}$.

(d) The proportion of women who favor a gun law is greater than the proportion of men who favor a gun law.

1.3-4 **(a)** $P(\text{HH}) = \dfrac{13}{52} \cdot \dfrac{12}{51} = \dfrac{1}{17}$;

(b) $P(\text{HC}) = \dfrac{13}{52} \cdot \dfrac{13}{51} = \dfrac{13}{204}$;

(c) $P(\text{Non-Ace Heart, Ace}) + P(\text{Ace of Hearts, Non-Heart Ace})$

$$= \frac{12}{52} \cdot \frac{4}{51} + \frac{1}{52} \cdot \frac{3}{51} = \frac{51}{52 \cdot 51} = \frac{1}{52}.$$

1.3-6 Let $H = \{\text{died from heart disease}\}$; $P = \{\text{at least one parent had heart disease}\}$.

$$P(H \mid P') = \frac{N(H \cap P')}{N(P')} = \frac{110}{648}.$$

1.3-8 **(a)** $\dfrac{3}{20} \cdot \dfrac{2}{19} \cdot \dfrac{1}{18} = \dfrac{1}{1140}$;

(b) $\dfrac{\dbinom{3}{2}\dbinom{17}{1}}{\dbinom{20}{3}} \cdot \dfrac{1}{17} = \dfrac{1}{380}$;

(c) $\displaystyle\sum_{k=1}^{9} \frac{\dbinom{3}{2}\dbinom{17}{2k-2}}{\dbinom{20}{2k}} \cdot \frac{1}{20-2k} = \frac{35}{76} = 0.4605.$

(d) Draw second. The probability of winning is $1 - 0.4605 = 0.5395$.

1.3-10 **(a)** $P(A) = \dfrac{52}{52} \cdot \dfrac{51}{52} \cdot \dfrac{50}{52} \cdot \dfrac{49}{52} \cdot \dfrac{48}{52} \cdot \dfrac{47}{52} = \dfrac{8{,}808{,}975}{11{,}881{,}376} = 0.74141$;

(b) $P(A') = 1 - P(A) = 0.25859$.

1.3-12 **(a)** It doesn't matter because $P(B_1) = \dfrac{1}{18},\ P(B_5) = \dfrac{1}{18},\ P(B_{18}) = \dfrac{1}{18}$;

(b) $P(B) = \dfrac{2}{18} = \dfrac{1}{9}$ on each draw.

1.3-14 **(a)** $P(A_1) = 30/100$;

(b) $P(A_3 \cap B_2) = 9/100$;

(c) $P(A_2 \cup B_3) = 41/100 + 28/100 - 9/100 = 60/100$;

(d) $P(A_1 \mid B_2) = 11/41$;

(e) $P(B_1 \mid A_3) = 13/29$.

1.3-16 $\dfrac{3}{5} \cdot \dfrac{5}{8} + \dfrac{2}{5} \cdot \dfrac{4}{8} = \dfrac{23}{40}$.

1.4 Independent Events

1.4-2 **(a)**
$$\begin{aligned} P(A \cap B) &= P(A)P(B) = (0.3)(0.6) = 0.18; \\ P(A \cup B) &= P(A) + P(B) - P(A \cap B) \\ &= 0.3 + 0.6 - 0.18 \\ &= 0.72. \end{aligned}$$

(b) $P(A|B) = \dfrac{P(A \cap B)}{P(B)} = \dfrac{0}{0.6} = 0$.

1.4-4 Proof of **(b)**:
$$\begin{aligned} P(A' \cap B) &= P(B)P(A'|B) \\ &= P(B)[1 - P(A|B)] \\ &= P(B)[1 - P(A)] \\ &= P(B)P(A'). \end{aligned}$$

Proof of **(c)**:
$$\begin{aligned} P(A' \cap B') &= P[(A \cup B)'] \\ &= 1 - P(A \cup B) \\ &= 1 - P(A) - P(B) + P(A \cap B) \\ &= 1 - P(A) - P(B) + P(A)P(B) \\ &= [1 - P(A)][1 - P(B)] \\ &= P(A')P(B'). \end{aligned}$$

1.4-6
$$\begin{aligned} P[A \cap (B \cap C)] &= P[A \cap B \cap C] \\ &= P(A)P(B)P(C) \\ &= P(A)P(B \cap C). \end{aligned}$$

$$\begin{aligned} P[A \cap (B \cup C)] &= P[(A \cap B) \cup (A \cap C)] \\ &= P(A \cap B) + P(A \cap C) - P(A \cap B \cap C) \\ &= P(A)P(B) + P(A)P(C) - P(A)P(B)P(C) \\ &= P(A)[P(B) + P(C) - P(B \cap C)] \\ &= P(A)P(B \cup C). \end{aligned}$$

$$\begin{aligned} P[A' \cap (B \cap C')] &= P(A' \cap C' \cap B) \\ &= P(B)[P(A' \cap C') \mid B] \\ &= P(B)[1 - P(A \cup C \mid B)] \\ &= P(B)[1 - P(A \cup C)] \\ &= P(B)P[(A \cup C)'] \\ &= P(B)P(A' \cap C') \\ &= P(B)P(A')P(C') \\ &= P(A')P(B)P(C') \\ &= P(A')P(B \cap C') \end{aligned}$$

$$\begin{aligned} P[A' \cap B' \cap C'] &= P[(A \cup B \cup C)'] \\ &= 1 - P(A \cup B \cup C) \\ &= 1 - P(A) - P(B) - P(C) + P(A)P(B) + P(A)P(C) + \\ &\qquad P(B)P(C) - PA)P(B)P(C) \\ &= [1 - P(A)][1 - P(B)][1 - P(C)] \\ &= P(A')P(B')P(C'). \end{aligned}$$

1.4-8 $\dfrac{1}{6} \cdot \dfrac{2}{6} \cdot \dfrac{3}{6} + \dfrac{1}{6} \cdot \dfrac{4}{6} \cdot \dfrac{3}{6} + \dfrac{5}{6} \cdot \dfrac{2}{6} \cdot \dfrac{3}{6} = \dfrac{2}{9}$.

1.4-10 **(a)** $\frac{3}{4} \cdot \frac{3}{4} = \frac{9}{16}$;

(b) $\frac{1}{4} \cdot \frac{3}{4} + \frac{3}{4} \cdot \frac{2}{4} = \frac{9}{16}$;

(c) $\frac{2}{4} \cdot \frac{1}{4} + \frac{2}{4} \cdot \frac{4}{4} = \frac{10}{16}$.

1.4-12 **(a)** $\left(\frac{1}{2}\right)^3 \left(\frac{1}{2}\right)^2$;

(b) $\left(\frac{1}{2}\right)^3 \left(\frac{1}{2}\right)^2$;

(c) $\left(\frac{1}{2}\right)^3 \left(\frac{1}{2}\right)^2$;

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

1.4-14 **(a)** $1 - (0.4)^3 = 1 - 0.064 = 0.936$;

(b) $1 - (0.4)^8 = 1 - 0.00065536 = 0.99934464$.

1.4-16 **(a)** $\displaystyle\sum_{k=0}^{\infty} \frac{1}{5}\left(\frac{4}{5}\right)^{2k} = \frac{5}{9}$;

(b) $\frac{1}{5} + \frac{4}{5} \cdot \frac{3}{4} \cdot \frac{1}{3} + \frac{4}{5} \cdot \frac{3}{4} \cdot \frac{2}{3} \cdot \frac{1}{2} \cdot \frac{1}{1} = \frac{3}{5}$.

1.4-18 **(a)** 7; **(b)** $(1/2)^7$; **(c)** 63; **(d)** No! $(1/2)^{63} = 1/9{,}223{,}372{,}036{,}854{,}775{,}808$.

1.4-20 No.

1.5 Bayes' Theorem

1.5-2 **(a)**
$$\begin{aligned} P(G) &= P(A \cap G) + P(B \cap G) \\ &= P(A)P(G \mid A) + P(B)P(G \mid B) \\ &= (0.40)(0.85) + (0.60)(0.75) = 0.79; \end{aligned}$$

(b)
$$\begin{aligned} P(A \mid G) &= \frac{P(A \cap G)}{P(G)} \\ &= \frac{(0.40)(0.85)}{0.79} = 0.43. \end{aligned}$$

1.5-4 Let event B denote an accident and let A_1 be the event that age of the driver is 16–25. Then

$$\begin{aligned} P(A_1 \mid B) &= \frac{(0.1)(0.05)}{(0.1)(0.05) + (0.55)(0.02) + (0.20)(0.03) + (0.15)(0.04)} \\ &= \frac{50}{50 + 110 + 60 + 60} = \frac{50}{280} = 0.179. \end{aligned}$$

1.5-6 Let B be the event that the policyholder dies. Let A_1, A_2, A_3 be the events that the deceased is standard, preferred and ultra-preferred, respectively. Then

$$
\begin{aligned}
P(A_1 \mid B) &= \frac{(0.60)(0.01)}{(0.60)(0.01)+(0.30)(0.008)+(0.10)(0.007)} \\
&= \frac{60}{60+24+7} = \frac{60}{91} = 0.659; \\
P(A_2 \mid B) &= \frac{24}{91} = 0.264; \\
P(A_3 \mid B) &= \frac{7}{91} = 0.077.
\end{aligned}
$$

1.5-8 Let A be the event that the tablet is under warranty.

$$
\begin{aligned}
P(B_1 \mid A) &= \frac{(0.40)(0.10)}{(0.40)(0.10)+(0.30)(0.05)+(0.20)(0.03)+(0.10)(0.02)} \\
&= \frac{40}{40+15+6+2} = \frac{40}{63} = 0.635; \\
P(B_2 \mid A) &= \frac{15}{63} = 0.238; \\
P(B_3 \mid A) &= \frac{6}{63} = 0.095; \\
P(B_4 \mid A) &= \frac{2}{63} = 0.032.
\end{aligned}
$$

1.5-10 **(a)** $P(D^+) = (0.02)(0.92) + (0.98)(0.05) = 0.0184 + 0.0490 = 0.0674$;

(b) $P(A^- \mid D^+) = \dfrac{0.0490}{0.0674} = 0.727$; $P(A^+ \mid D^+) = \dfrac{0.0184}{0.0674} = 0.273$;

(c) $P(A^- \mid D^-) = \dfrac{(0.98)(0.95)}{(0.02)(0.08)+(0.98)(0.95)} = \dfrac{9310}{16+9310} = 0.998$;

$P(A^+ \mid D^-) = 0.002$.

(d) Yes, particularly those in part (b).

1.5-12 Let $D = \{\text{has the disease}\}$, $DP = \{\text{detects presence of disease}\}$. Then

$$
\begin{aligned}
P(D \mid DP) &= \frac{P(D \cap DP)}{P(DP)} \\
&= \frac{P(D)\cdot P(DP \mid D)}{P(D)\cdot P(DP \mid D) + P(D')\cdot P(DP \mid D')} \\
&= \frac{(0.005)(0.90)}{(0.005)(0.90)+(0.995)(0.02)} \\
&= \frac{0.0045}{0.0045+0.199} = \frac{0.0045}{0.0244} = 0.1844.
\end{aligned}
$$

1.5-14 Let $D = \{\text{defective roll}\}$. Then

$$
\begin{aligned}
P(I \mid D) &= \frac{P(I \cap D)}{P(D)} \\
&= \frac{P(I)\cdot P(D \mid I)}{P(I)\cdot P(D \mid I) + P(II)\cdot P(D \mid II)} \\
&= \frac{(0.60)(0.03)}{(0.60)(0.03)+(0.40)(0.01)} \\
&= \frac{0.018}{0.018+0.004} = \frac{0.018}{0.022} = 0.818.
\end{aligned}
$$

Chapter 2

Discrete Distributions

2.1 Random Variables of the Discrete Type

2.1-2 **(a)**

$$f(x) = \begin{cases} 0.6, & x = 1, \\ 0.3, & x = 5, \\ 0.1, & x = 10, \end{cases}$$

(b)

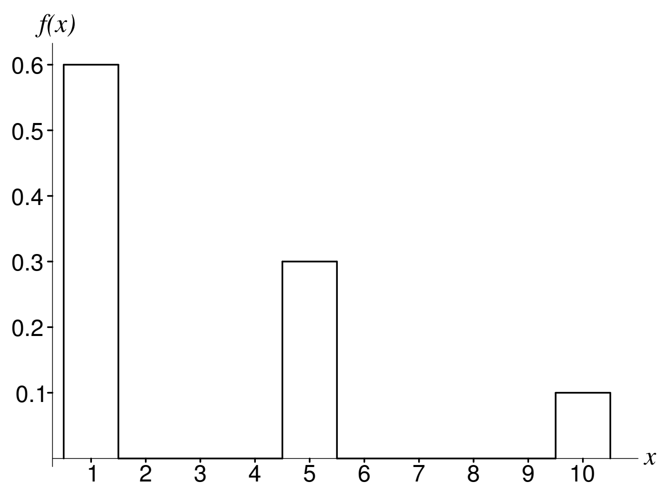


Figure 2.1–2: A probability histogram

2.1-4 **(a)** $f(x) = \dfrac{1}{10}, \qquad x = 0, 1, 2, \cdots, 9;$

(b)
$\mathcal{N}(\{0\})/150 = 11/150 = 0.073;$ $\mathcal{N}(\{5\})/150 = 13/150 = 0.087;$
$\mathcal{N}(\{1\})/150 = 14/150 = 0.093;$ $\mathcal{N}(\{6\})/150 = 22/150 = 0.147;$
$\mathcal{N}(\{2\})/150 = 13/150 = 0.087;$ $\mathcal{N}(\{7\})/150 = 16/150 = 0.107;$
$\mathcal{N}(\{3\})/150 = 12/150 = 0.080;$ $\mathcal{N}(\{8\})/150 = 18/150 = 0.120;$
$\mathcal{N}(\{4\})/150 = 16/150 = 0.107;$ $\mathcal{N}(\{9\})/150 = 15/150 = 0.100.$

(c)

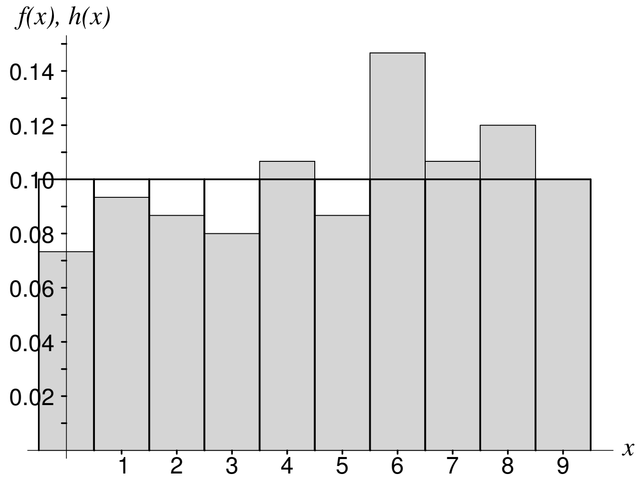


Figure 2.1–4: Michigan daily lottery digits

2.1-6 **(a)** $f(x) = \dfrac{6 - |\,7 - x\,|}{36}$, $x = 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12.$

(b)

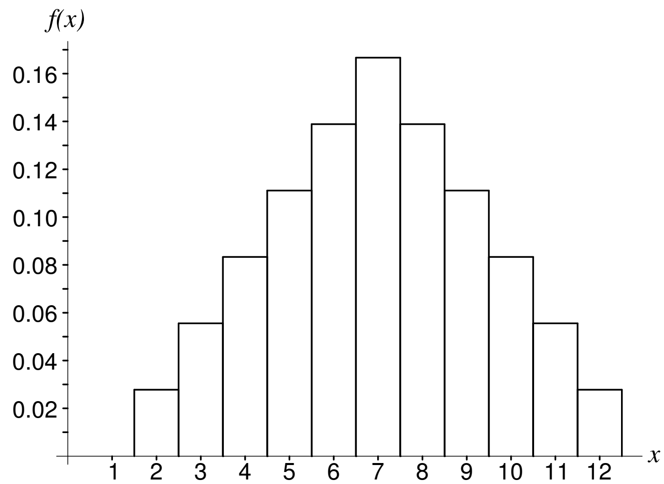


Figure 2.1–6: Probability histogram for the sum of a pair of dice

2.1-8 **(a)** The space of W is $S = \{0, 1, 2, 3, 4, 5, 6, 7\}$.

$P(W = 0) = P(X = 0, Y = 0) = \frac{1}{2} \cdot \frac{1}{4} = \frac{1}{8}$, assuming independence.

$P(W = 1) = P(X = 0, Y = 1) = \frac{1}{2} \cdot \frac{1}{4} = \frac{1}{8}$,

$P(W = 2) = P(X = 2, Y = 0) = \frac{1}{2} \cdot \frac{1}{4} = \frac{1}{8}$,

$P(W = 3) = P(X = 2, Y = 1) = \frac{1}{2} \cdot \frac{1}{4} = \frac{1}{8}$,

$P(W = 4) = P(X = 0, Y = 4) = \frac{1}{2} \cdot \frac{1}{4} = \frac{1}{8}$,

$P(W = 5) = P(X = 0, Y = 5) = \frac{1}{2} \cdot \frac{1}{4} = \frac{1}{8}$,

$P(W = 6) = P(X = 2, Y = 4) = \frac{1}{2} \cdot \frac{1}{4} = \frac{1}{8}$,

$P(W = 7) = P(X = 2, Y = 5) = \frac{1}{2} \cdot \frac{1}{4} = \frac{1}{8}$.

That is, $f(w) = P(W = w) = \frac{1}{8}, \quad w \in S$.

(b)

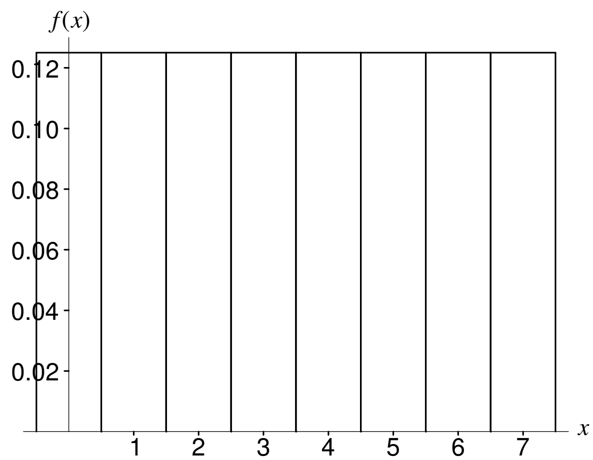


Figure 2.1–8: Probability histogram of sum of two special dice

2.1-10 **(a)** $\dfrac{\binom{3}{1}\binom{47}{9}}{\binom{50}{10}} = \dfrac{39}{98}$;

(b) $\displaystyle\sum_{x=0}^{1} \frac{\binom{3}{x}\binom{47}{10-x}}{\binom{50}{10}} = \frac{221}{245}$.

2.1-12
$$\begin{aligned} P(X \ge 4 \,|\, X \ge 1) &= \frac{P(X \ge 4)}{P(X \ge 1)} = \frac{1 - P(X \le 3)}{1 - P(X = 0)} \\ &= \frac{1 - [1 - 1/2 + 1/2 - 1/3 + 1/3 - 1/4 + 1/4 - 1/5]}{1 - [1 - 1/2]} = \frac{2}{5}. \end{aligned}$$

2.1-14 $P(X \ge 1) = 1 - P(X = 0) = 1 - \dfrac{\binom{3}{0}\binom{17}{5}}{\binom{20}{5}} = 1 - \dfrac{91}{228} = \dfrac{137}{228} = 0.60.$

2.1-16 **(a)** $P(2, 1, 6, 10)$ means that 2 is in position 1 so 1 cannot be selected. Thus

$$P(2, 1, 6, 10) = \frac{\binom{1}{0}\binom{1}{1}\binom{8}{5}}{\binom{10}{6}} = \frac{56}{210} = \frac{4}{15};$$

(b) $P(i, r, k, n) = \dfrac{\binom{i-1}{r-1}\binom{1}{1}\binom{n-i}{k-r}}{\binom{n}{k}}.$

2.2 Mathematical Expectation

2.2-2 $E(X) = (-1)\left(\dfrac{4}{9}\right) + (0)\left(\dfrac{1}{9}\right) + (1)\left(\dfrac{4}{9}\right) = 0;$

$E(X^2) = (-1)^2\left(\dfrac{4}{9}\right) + (0)^2\left(\dfrac{1}{9}\right) + (1)^2\left(\dfrac{4}{9}\right) = \dfrac{8}{9};$

$E(3X^2 - 2X + 4) = 3\left(\dfrac{8}{9}\right) - 2(0) + 4 = \dfrac{20}{3}.$

2.2-4
$$\begin{aligned} 1 &= \sum_{x=0}^{6} f(x) = \frac{9}{10} + c\left(\frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6}\right) \\ c &= \frac{2}{49}; \\ E(\text{Payment}) &= \frac{2}{49}\left(1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{3} + 3 \cdot \frac{1}{4} + 4 \cdot \frac{1}{5} + 5 \cdot \frac{1}{6}\right) = \frac{71}{490} \text{ units.} \end{aligned}$$

2.2-6 Note that $\displaystyle\sum_{x=1}^{\infty} \frac{6}{\pi^2 x^2} = \frac{6}{\pi^2}\sum_{x=1}^{\infty} \frac{1}{x^2} = \frac{6}{\pi^2}\frac{\pi^2}{6} = 1$, so this is a pdf

$$E(X) = \sum_{x=1}^{\infty} x\,\frac{6}{\pi^2 x^2} = \frac{6}{\pi^2}\sum_{x=1}^{\infty}\frac{1}{x} = +\infty$$

and it is well known that the sum of this harmonic series is not finite.

2.2-8 $E(|X - c|) = \dfrac{1}{7}\displaystyle\sum_{x \in S} |x - c|$, where $S = \{1, 2, 3, 5, 15, 25, 50\}$.

When $c = 5$,

$$E(|X - 5|) = \frac{1}{7}\left[(5 - 1) + (5 - 2) + (5 - 3) + (5 - 5) + (15 - 5) + (25 - 5) + (50 - 5)\right].$$

If c is either increased or decreased by 1, this expectation is increased by 1/7. Thus $c = 5$, the median, minimizes this expectation while $b = E(X) = \mu$, the mean, minimizes $E[(X-b)^2]$. You could also let $h(c) = E(\,|\,X - c\,|\,)$ and show that $h'(c) = 0$ when $c = 5$.

2.2-10 $(1)\cdot\frac{15}{36} + (-1)\cdot\frac{21}{36} = \frac{-6}{36} = \frac{-1}{6}$;

$(1)\cdot\frac{15}{36} + (-1)\cdot\frac{21}{36} = \frac{-6}{36} = \frac{-1}{6}$;

$(4)\cdot\frac{6}{36} + (-1)\cdot\frac{30}{36} = \frac{-6}{36} = \frac{-1}{6}$.

2.2-12 **(a)** The average class size is $\dfrac{(16)(25) + (3)(100) + (1)(300)}{20} = 50$;

(b)

$$f(x) = \begin{cases} 0.4, & x = 25, \\ 0.3, & x = 100, \\ 0.3, & x = 300, \end{cases}$$

(c) $E(X) = 25(0.4) + 100(0.3) + 300(0.3) = 130$.

2.3 Special Mathematical Expectations

2.3-2 **(a)**

$$\begin{aligned}
\mu &= E(X) \\
&= \sum_{x=1}^{3} x\,\frac{3!}{x!\,(3-x)!}\left(\frac{1}{4}\right)^x\left(\frac{3}{4}\right)^{3-x} \\
&= 3\left(\frac{1}{4}\right)\sum_{k=0}^{2}\frac{2!}{k!\,(2-k)!}\left(\frac{1}{4}\right)^k\left(\frac{3}{4}\right)^{2-k} \\
&= 3\left(\frac{1}{4}\right)\left(\frac{1}{4}+\frac{3}{4}\right)^2 = \frac{3}{4}; \\
E[X(X-1)] &= \sum_{x=2}^{3} x(x-1)\,\frac{3!}{x!\,(3-x)!}\left(\frac{1}{4}\right)^x\left(\frac{3}{4}\right)^{3-x} \\
&= 2(3)\left(\frac{1}{4}\right)^2\frac{3}{4} + 6\left(\frac{1}{4}\right)^3 \\
&= 6\left(\frac{1}{4}\right)^2 = 2\left(\frac{1}{4}\right)\left(\frac{3}{4}\right); \\
\sigma^2 &= E[X(X-1)] + E(X) - \mu^2 \\
&= (2)\left(\frac{3}{4}\right)\left(\frac{1}{4}\right) + \left(\frac{3}{4}\right) - \left(\frac{3}{4}\right)^2 \\
&= (2)\left(\frac{3}{4}\right)\left(\frac{1}{4}\right) + \left(\frac{3}{4}\right)\left(\frac{1}{4}\right) = 3\left(\frac{1}{4}\right)\left(\frac{3}{4}\right);
\end{aligned}$$

(b)
$$\begin{aligned}\mu &= E(X)\\ &= \sum_{x=1}^{4} x\,\frac{4!}{x!\,(4-x)!}\left(\frac{1}{2}\right)^x\left(\frac{1}{2}\right)^{4-x}\\ &= 4\left(\frac{1}{2}\right)\sum_{k=0}^{3}\frac{3!}{k!\,(3-k)!}\left(\frac{1}{2}\right)^k\left(\frac{1}{2}\right)^{3-k}\\ &= 4\left(\frac{1}{2}\right)\left(\frac{1}{2}+\frac{1}{2}\right)^3 = 2;\end{aligned}$$

$$\begin{aligned}E[X(X-1)] &= \sum_{x=2}^{4} x(x-1)\,\frac{4!}{x!\,(4-x)!}\left(\frac{1}{2}\right)^x\left(\frac{1}{2}\right)^{4-x}\\ &= 2(6)\left(\frac{1}{2}\right)^4+(6)(4)\left(\frac{1}{2}\right)^4+(12)\left(\frac{1}{2}\right)^4\\ &= 48\left(\frac{1}{2}\right)^4 = 12\left(\frac{1}{2}\right)^2;\\ \sigma^2 &= (12)\left(\frac{1}{2}\right)^2+\frac{4}{2}-\left(\frac{4}{2}\right)^2 = 1.\end{aligned}$$

2.3-4 $E[(X-\mu)/\sigma] = (1/\sigma)[E(X)-\mu] = (1/\sigma)(\mu-\mu) = 0;$

$E\{[(X-\mu)/\sigma]^2\} = (1/\sigma^2)E[(X-\mu)^2] = (1/\sigma^2)(\sigma^2) = 1.$

2.3-6 $f(1) = \frac{3}{8},\ f(2) = \frac{2}{8},\ f(3) = \frac{3}{8}$

$\mu = 1\cdot\frac{3}{8}+2\cdot\frac{2}{8}+3\cdot\frac{3}{8} = 2,$

$\sigma^2 = 1^2\cdot\frac{3}{8}+2^2\cdot\frac{2}{8}+3^2\cdot\frac{3}{8}-2^2 = \frac{3}{4}.$

2.3-8
$$\begin{aligned}E(X) &= \sum_{x=1}^{4} x\cdot\frac{2x-1}{16}\\ &= \frac{50}{16} = 3.125;\\ E(X^2) &= \sum_{x=1}^{4} x^2\cdot\frac{2x-1}{16}\\ &= \frac{85}{8};\\ \mathrm{Var}(X) &= \frac{85}{8}-\left(\frac{25}{8}\right)^2 = \frac{55}{64} = 0.8594;\\ \sigma &= \frac{\sqrt{55}}{8} = 0.9270.\end{aligned}$$

2.3-10 We have $N = N_1 + N_2$. Thus

$$\begin{aligned}
E[X(X-1)] &= \sum_{x=0}^{n} x(x-1)f(x) \\
&= \frac{\displaystyle\sum_{x=2}^{n} x(x-1)\,\frac{N_1!}{x!\,(N_1-x)!}\cdot\frac{N_2!}{(n-x)!\,(N_2-n+x)!}}{\dbinom{N}{n}} \\
&= N_1(N_1-1)\,\frac{\displaystyle\sum_{x=2}^{n} \frac{(N_1-2)!}{(x-2)!\,(N_1-x)!}\cdot\frac{N_2!}{(n-x)!\,(N_2-n+x)!}}{\dbinom{N}{n}}.
\end{aligned}$$

In the summation, let $k = x - 2$, and in the denominator, note that

$$\binom{N}{n} = \frac{N!}{n!(N-n)!} = \frac{N(N-1)}{n(n-1)}\binom{N-2}{n-2}.$$

Thus

$$\begin{aligned}
E[X(X-1)] &= \frac{N_1(N_1-1)}{\dfrac{N(N-1)}{n(n-1)}}\sum_{k=0}^{n-2}\frac{\dbinom{N_1-2}{k}\dbinom{N_2}{n-2-k}}{\dbinom{N-2}{n-2}} \\
&= \frac{N_1(N_1-1)(n)(n-1)}{N(N-1)}.
\end{aligned}$$

2.3-12 **(a)** $f(x) = \left(\dfrac{364}{365}\right)^{x-1}\left(\dfrac{1}{365}\right), \qquad x = 1, 2, 3, \ldots,$

(b)
$$\begin{aligned}
\mu &= \frac{1}{\dfrac{1}{365}} = 365, \\
\sigma^2 &= \frac{\dfrac{364}{365}}{\left(\dfrac{1}{365}\right)^2} = 132{,}860, \\
\sigma &= 364.500;
\end{aligned}$$

(c)
$$\begin{aligned}
P(X > 400) &= \left(\frac{364}{365}\right)^{400} = 0.3337, \\
P(X < 300) &= 1 - \left(\frac{364}{365}\right)^{299} = 0.5597.
\end{aligned}$$

2.3-14 $P(X \ge 100) = P(X > 99) = (0.99)^{99} = 0.3697.$

2.3-16 **(a)** $f(x) = (1/2)^{x-1}, \quad x = 2, 3, 4, \ldots;$

(b)
$$\begin{aligned} M(t) &= E[e^{tx}] = \sum_{x=2}^{\infty} e^{tx}(1/2)^{x-1} \\ &= 2\sum_{x=2}^{\infty}(e^t/2)^x \\ &= \frac{2(e^t/2)^2}{1-e^t/2} = \frac{e^{2t}}{2-e^t}, \qquad t < \ln 2; \end{aligned}$$

(c)
$$\begin{aligned} M'(t) &= \frac{4e^{2t}-e^{3t}}{(2-e^t)^2} \\ \mu &= M'(0) = 3; \\ M''(t) &= \frac{(2-e^t)^2(8e^{2t}-3e^{3t})-(4e^{2t}-e^{3t})2*(2-e^t)(-e^t)}{(2-e^t)^4} \\ \sigma^2 &= M''(0)-\mu^2 = 11-9 = 2; \end{aligned}$$

(d) **(i)** $P(X \le 3) = 3/4$; **(ii)** $P(X \ge 5) = 1/8$; **(iii)** $P(X = 3) = 1/4$.

2.3-18
$$\begin{aligned} P(X > k+j \,|\, X > k) &= \frac{P(X > k+j)}{P(X > k)} \\ &= \frac{q^{k+j}}{q^k} = q^j = P(X > j). \end{aligned}$$

2.4 The Binomial Distribution

2.4-2 $f(-1) = \dfrac{11}{18}, \qquad f(1) = \dfrac{7}{18};$

$$\mu = (-1)\frac{11}{18} + (1)\frac{7}{18} = -\frac{4}{18};$$

$$\sigma^2 = \left(-1+\frac{4}{18}\right)^2\left(\frac{11}{18}\right) + \left(1+\frac{4}{18}\right)^2\left(\frac{7}{18}\right) = \frac{77}{81}.$$

2.4-4 **(a)** X is $b(7,\ 0.15)$;

(b) **(i)** $P(X \ge 2) = 1 - P(X \le 1) = 1 - 0.7166 = 0.2834;$

(ii) $P(X = 1) = P(X \le 1) - P(X \le 0) = 0.7166 - 0.3206 = 0.3960;$

(iii) $P(X \le 3) = 0.9879.$

2.4-6 **(a)** X is $b(15, 0.75)$; $\qquad 15 - X$ is $b(15, 0.25)$;

(b) $P(X \ge 10) = P(15 - X \le 5) = 0.8516;$

(c) $P(X \le 10) = P(15 - X \ge 5) = 1 - P(15 - X \le 4) = 1 - 0.6865 = 0.3135;$

(d) $P(X = 10) = P(X \ge 10) - P(X \ge 11)$
$= P(15 - X \le 5) - P(15 - X \le 4) = 0.8516 - 0.6865 = 0.1651;$

(e) $\mu = (15)(0.75) = 11.25, \quad \sigma^2 = (15)(0.75)(0.25) = 2.8125; \quad \sigma = \sqrt{2.8125} = 1.67705.$

2.4-8 **(a)** $1 - 0.01^4 = 0.99999999$; **(b)** $0.99^4 = 0.960596$.

2.4-10 **(a)** X is $b(8,\ 0.90)$;

(b) **(i)** $P(X = 8) = P(8 - X = 0) = 0.4305;$

(ii) $P(X \le 6) = P(8 - X \ge 2)$
$= 1 - P(8 - X \le 1) = 1 - 0.8131 = 0.1869;$

(iii) $P(X \ge 6) = P(8 - X \le 2) = 0.9619.$

2.4-12 **(a)**

$$f(x) = \begin{cases} 125/216, & x = -1, \\ 75/216, & x = 1, \\ 15/216, & x = 2, \\ 1/216, & x = 3; \end{cases}$$

(b)

$$\begin{aligned} \mu &= (-1)\cdot\frac{125}{216} + (1)\cdot\frac{75}{216} + (2)\cdot\frac{15}{216} + (3)\cdot\frac{1}{216} = -\frac{17}{216}; \\ \sigma^2 &= E(X^2) - \mu^2 = \frac{269}{216} - \left(-\frac{17}{216}\right)^2 = 1.2392; \\ \sigma &= 1.11; \end{aligned}$$

(c) See Figure 2.4-12.

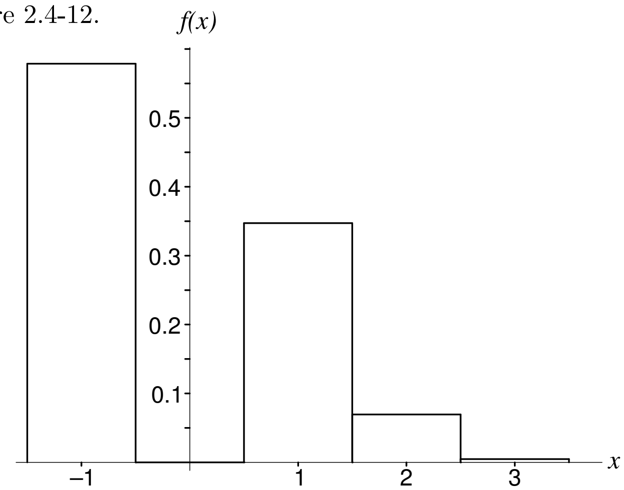


Figure 2.4–12: Losses in chuck-a-luck

2.4-14 Let X equal the number of winning tickets when n tickets are purchased. Then

$$\begin{aligned} P(X \ge 1) &= 1 - P(X = 0) \\ &= 1 - \left(\frac{9}{10}\right)^n. \end{aligned}$$

(a)

$$\begin{aligned} 1 - (0.9)^n &= 0.50 \\ (0.9)^n &= 0.50 \\ n \ln 0.9 &= \ln 0.5 \\ n &= \frac{\ln 0.5}{\ln 0.9} = 6.58 \end{aligned}$$

so $n = 7$.

(b)

$$\begin{aligned} 1 - (0.9)^n &= 0.95 \\ (0.9)^n &= 0.05 \\ n &= \frac{\ln 0.05}{\ln 0.09} = 28.43 \end{aligned}$$

so $n = 29$.

2.4-16 It is given that X is $b(10,\ 0.10)$. We are to find M so that

$P(1000X \le M) \ge 0.99)$ or $P(X \le M/1000) \ge 0.99$. From Appendix Table II,

$P(X \le 4) = 0.9984 > 0.99$. Thus $M/1000 = 4$ or $M = 4000$ dollars.

2.4-18 X is $b(5,\ 0.05)$. The expected number of tests is

$1\,P(X = 0) + 6\,P(X > 0) = 1\,(0.7738) + 6\,(1 - 0.7738) = 2.131.$

2.4-20 **(a)** **(i)** $b(5,\ 0.7)$; **(ii)** $\mu = 3.5, \sigma^2 = 1.05$; **(iii)** 0.1607;

(b) **(i)** geometric, $p = 0.3$; **(ii)** $\mu = 10/3, \sigma^2 = 70/9$; **(iii)** 0.51;

(c) **(i)** Bernoulli, $p = 0.55$; **(ii)** $\mu = 0.55, \sigma^2 = 0.2475$; **(iii)** 0.55;

(d) **(ii)** $\mu = 2.1, \sigma^2 = 0.89$; **(iii)** 0.7;

(e) **(i)** discrete uniform on $1, 2, \ldots, 10$; **(ii)** 5.5, 8.25; **(iii)** 0.2.

2.5 The Negative Binomial Distribution

2.5-2 $\displaystyle \binom{10-1}{5-1}\left(\frac{1}{2}\right)^5\left(\frac{1}{2}\right)^5 = \frac{126}{1024} = \frac{63}{512}.$

2.5-4 Let "being missed" be a success and let X equal the number of trials until the first success. Then $p = 0.01$.

$P(X \le 50) = 1 - 0.99^{50} = 1 - 0.605 = 0.395.$

2.5-6 **(a)**

$$\begin{aligned}
R(t) &= \ln(1 - p + pe^t),\\
R'(t) &= \left[\frac{pe^t}{1 - p + pe^t}\right]_{t=0} = p,\\
R''(t) &= \left[\frac{(1 - p + pe^t)(pe^t) - (pe^t)(pe^t)}{(1 - p + pe^t)^2}\right]_{t=0} = p(1 - p);
\end{aligned}$$

(b)

$$\begin{aligned}
R(t) &= n\ln(1 - p + pe^t),\\
R'(t) &= \left[\frac{npe^t}{1 - p + pe^t}\right]_{t=0} = np,\\
R''(t) &= n\left[\frac{(1 - p + pe^t)(pe^t) - (pe^t)(pe^t)}{(1 - p + pe^t)^2}\right]_{t=0} = np(1 - p);
\end{aligned}$$

(c)

$$\begin{aligned}
R(t) &= \ln p + t - \ln[1 - (1 - p)e^t],\\
R'(t) &= \left[1 + \frac{(1 - p)e^t}{1 - (1 - p)e^t}\right]_{t=0} = 1 + \frac{1 - p}{p} = \frac{1}{p},\\
R''(t) &= \left[(-1)\{1 - (1 - p)e^t\}^2\{-(1 - p)e^t\}\right]_{t=0} = \frac{1 - p}{p};
\end{aligned}$$

(d)

$$\begin{aligned}
R(t) &= r\left[\ln p + t - \ln\{1 - (1 - p)e^t\}\right],\\
R'(t) &= r\left[\frac{1}{1 - (1 - p)e^t}\right]_{t=0} = \frac{r}{p},\\
R''(t) &= r\left[(-1)\{1 - (1 - p)e^t\}^{-2}\{-(1 - p)e^t\}\right]_{t=0} = \frac{r(1 - p)}{p^2}.
\end{aligned}$$

2.5-8 $(0.7)(0.7)(0.3) = 0.147.$

2.5-10 **(a)** Let X equal the number of boxes that must be purchased. Then

$$E(X) = \sum_{x=1}^{12} \frac{1}{(13-x)/12} = \frac{86021}{2310} = 37.2385;$$

(b) $\dfrac{100 \cdot E(X)}{365} \approx 10.2.$

2.6 The Poisson Distribution

2.6-2 $\lambda = \mu = \sigma^2 = 3$ so $P(X=2) = 0.423 - 0.199 = 0.224.$

2.6-4
$$\begin{aligned} 3\,\frac{\lambda^1 e^{-\lambda}}{1!} &= \frac{\lambda^2 e^{-\lambda}}{2!} \\ e^{-\lambda}\lambda(\lambda-6) &= 0 \\ \lambda &= 6 \end{aligned}$$

Thus $P(X=4) = 0.285 - 0.151 = 0.134.$

2.6-6 $\lambda = (1)(50/100) = 0.5$, so $P(X=0) = e^{-0.5}/0! = 0.607.$

2.6-8 $np = 1000(0.005) = 5;$

(a) $P(X \le 1) \approx 0.040;$

(b) $P(X = 4,5,6) = P(X \le 6) - P(X \le 3) \approx 0.762 - 0.265 = 0.497.$

2.6-10 $\sigma = \sqrt{9} = 3,$

$P(3 < X < 15) = P(X \le 14) - P(X \le 3) = 0.959 - 0.021 = 0.938.$

2.6-12 Since $E(X) = 0.2$, the expected loss is $(0.02)(\$10{,}000) = \$2{,}000.$

Chapter 3

Continuous Distributions

3.1 Random Variables of the Continuous Type

3.1–2 $\mu = 0, \quad \sigma^2 = (1+1)^2/12 = 1/3.$

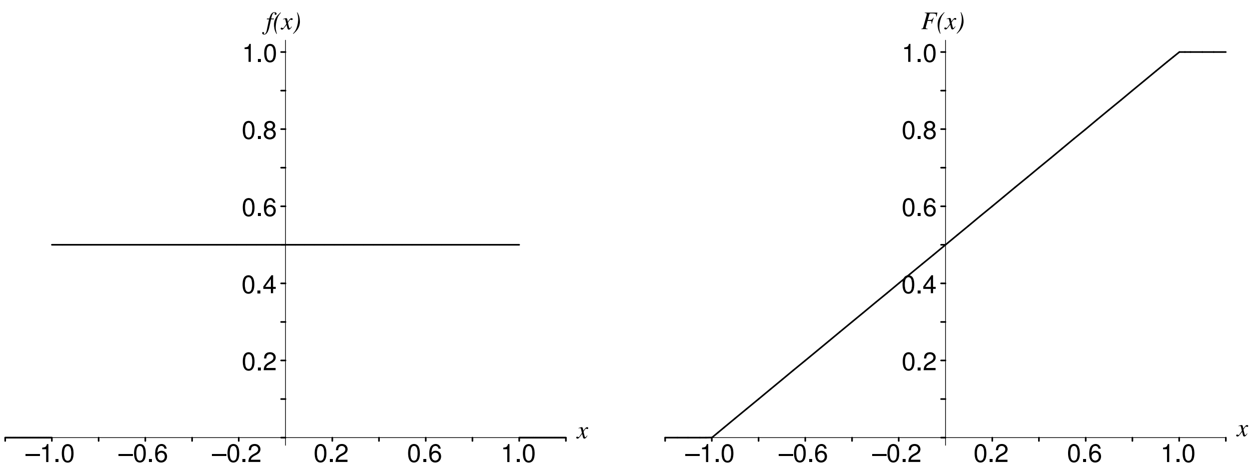


Figure 3.1–2: $f(x) = 1/2$ and $F(x) = (x+1)/2$

3.1–4 X is $U(4, 5)$;

(a) $\mu = 9/2$; **(b)** $\sigma^2 = 1/12$; **(c)** 0.5.

3.1–6

$$\begin{aligned}
E(\text{profit}) &= \int_0^n [x - 0.5(n-x)]\frac{1}{200}\,dx + \int_n^{200} [n - 5(x-n)]\frac{1}{200}\,dx \\
&= \frac{1}{200}\left[\frac{x^2}{2} + \frac{(n-x)^2}{4}\right]_0^n + \frac{1}{200}\left[6nx - \frac{5x^2}{2}\right]_n^{200} \\
&= \frac{1}{200}\left[-3.25n^2 + 1200n - 100000\right] \\
\text{derivative} &= \frac{1}{200}[-6.5n + 1200] \;=\; 0 \\
n &= \frac{1200}{6.5} \;\approx\; 185.
\end{aligned}$$

3.1–8 **(a)** **(i)**

$$\begin{aligned}\int_0^c x^3/4\,dx &= 1\\ c^4/16 &= 1\\ c &= 2;\end{aligned}$$

(ii)

$$\begin{aligned}F(x) &= \int_{-\infty}^x f(t)\,dt\\ &= \int_0^x t^3/4\,dt\\ &= x^4/16,\end{aligned}$$

$$F(x) = \begin{cases} 0, & -\infty < x < 0,\\ x^4/16, & 0 \le x < 2,\\ 1, & 2 \le x < \infty.\end{cases}$$

(iii)

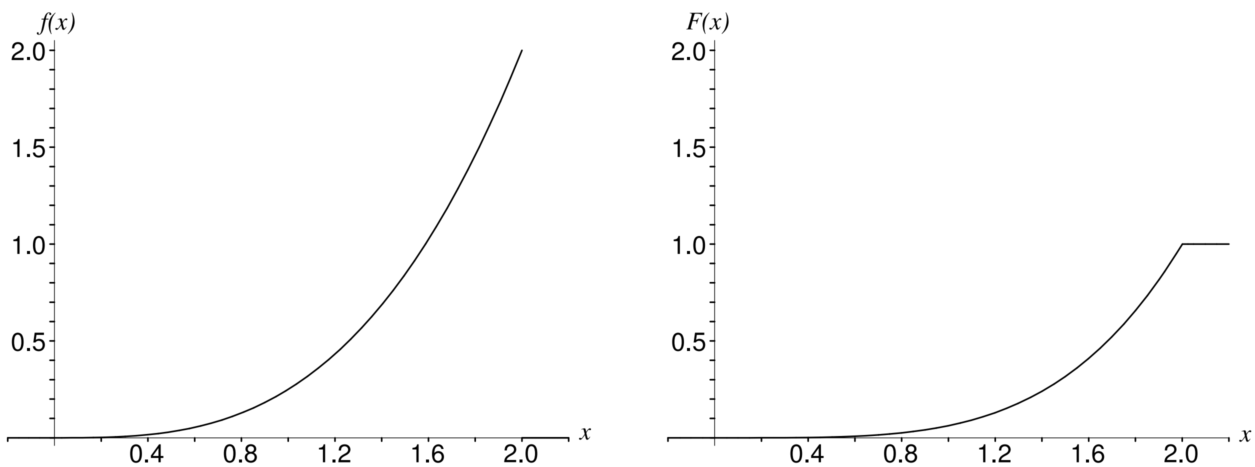


Figure 3.1–8: **(a)** Continuous distribution pdf and cdf

(iv)

$$\begin{aligned}\mu &= \int_0^2 (x)(x^3/4)\,dx = \frac{8}{5};\\ E(X^2) &= \int_0^2 (x^2)(x^3/4)\,dx = \frac{8}{3};\\ \sigma^2 &= \frac{8}{3} - \left(\frac{8}{5}\right)^2 = \frac{8}{75}.\end{aligned}$$

(b) **(i)**

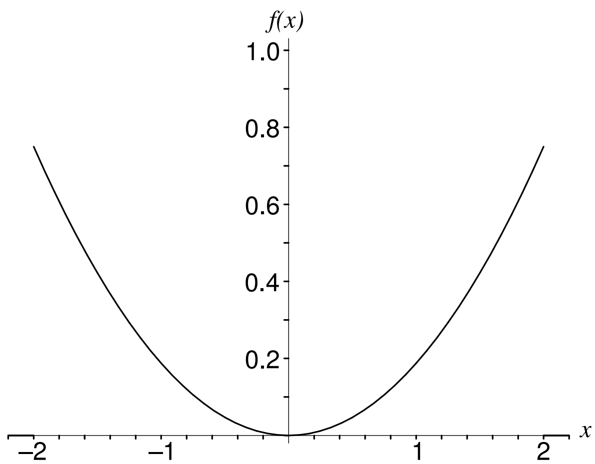
$$\begin{aligned}\int_{-c}^{c} (3/16)x^2\,dx &= 1\\ c^3/8 &= 1\\ c &= 2;\end{aligned}$$

(ii)

$$\begin{aligned}F(x) &= \int_{-\infty}^{x} f(t)\,dt\\ &= \int_{-2}^{x} (3/16)t^2\,dt\\ &= \left[\frac{t^3}{16}\right]_{-2}^{x}\\ &= \frac{x^3}{16}+\frac{1}{2},\end{aligned}$$

$$F(x)=\begin{cases}0, & -\infty < x < -2,\\ \dfrac{x^3}{16}+\dfrac{1}{2}, & -2 \le x < 2,\\ 1, & 2 \le x < \infty.\end{cases}$$

(iii)



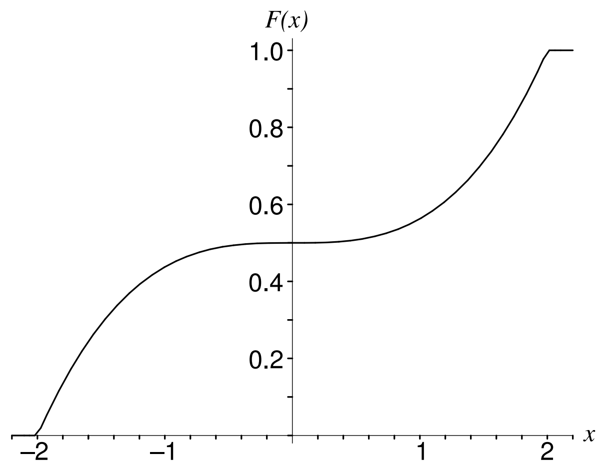


Figure 3.1–8: **(b)** Continuous distribution pdf and cdf

(iv)

$$\begin{aligned}\mu &= \int_{-2}^{2} (x)(3/16)(x^2)\,dx \;=\; 0;\\ \sigma^2 &= \int_{-2}^{2} (x^2)(3/16)(x^2)\,dx \;=\; \frac{12}{5}.\end{aligned}$$

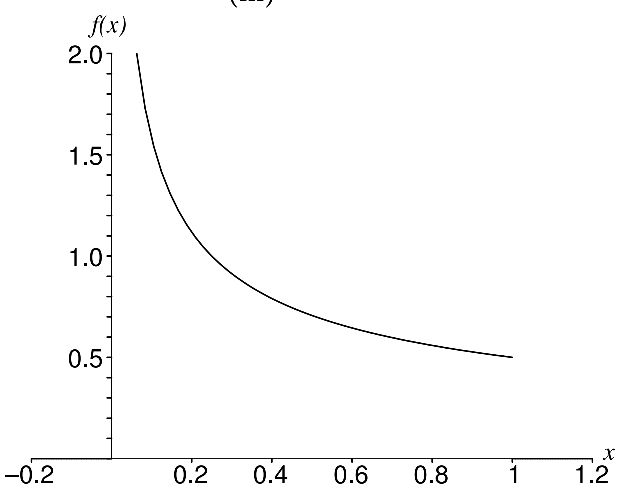
(c) **(i)**
$$\begin{aligned}\int_0^1 \frac{c}{\sqrt{x}}\,dx &= 1\\ 2c &= 1\\ c &= 1/2.\end{aligned}$$

The pdf in part (c) is unbounded.

(ii)
$$\begin{aligned}F(x) &= \int_{-\infty}^{x} f(t)\,dt\\ &= \int_0^x \frac{1}{2\sqrt{t}}\,dt\\ &= \left[\sqrt{t}\,\right]_0^x \;=\; \sqrt{x},\end{aligned}$$

$$F(x) = \begin{cases} 0, & -\infty < x < 0,\\ \sqrt{x}, & 0 \le x < 1,\\ 1, & 1 \le x < \infty.\end{cases}$$

(iii)



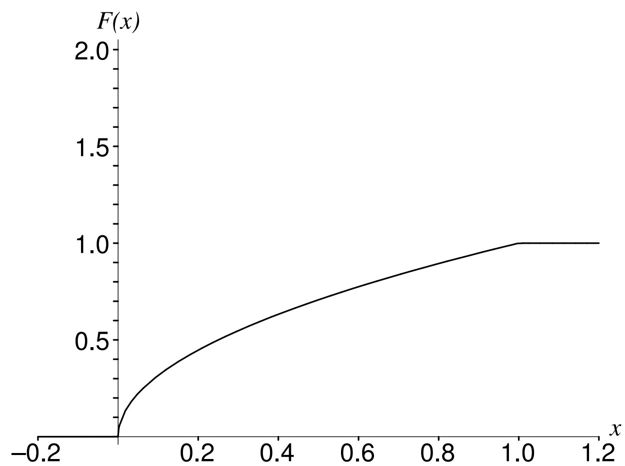


Figure 3.1–8: **(c)** Continuous distribution pdf and cdf

(iv)
$$\begin{aligned}\mu &= \int_0^1 (x)(1/2)/\sqrt{x}\,dx \;=\; \frac{1}{3};\\ E(X^2) &= \int_0^1 (x^2)(1/2)/\sqrt{x}\,dx \;=\; \frac{1}{5};\\ \sigma^2 &= \frac{1}{5} - \left(\frac{1}{3}\right)^2 \;=\; \frac{4}{45}.\end{aligned}$$

3.1–10 **(a)**
$$\begin{aligned}\int_1^\infty \frac{c}{x^2}\,dx &= 1\\ \left[\frac{-c}{x}\right]_1^\infty &= 1\\ c &= 1;\end{aligned}$$

(b) $E(X) = \displaystyle\int_1^\infty \frac{x}{x^2}\,dx = [\ln x]_1^\infty$, which is unbounded.

3.1–12 **(a)**

$$F(x) = \begin{cases} 0, & -\infty < x < -1, \\ (x^3+1)/2, & -1 \le x < 1, \\ 1, & 1 \le x < \infty. \end{cases}$$

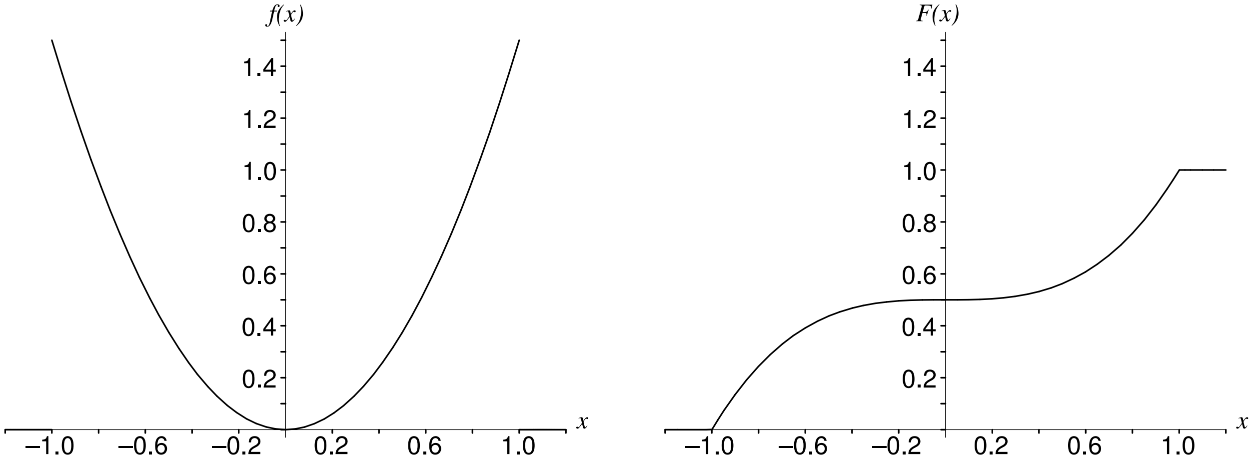


Figure 3.1–12: **(a)** $f(x) = (3/2)x^2$ and $F(x) = (x^3+1)/2$

(b)

$$F(x) = \begin{cases} 0, & -\infty < x < -1, \\ (x+1)/2, & -1 \le x < 1, \\ 1, & 1 \le x < \infty. \end{cases}$$

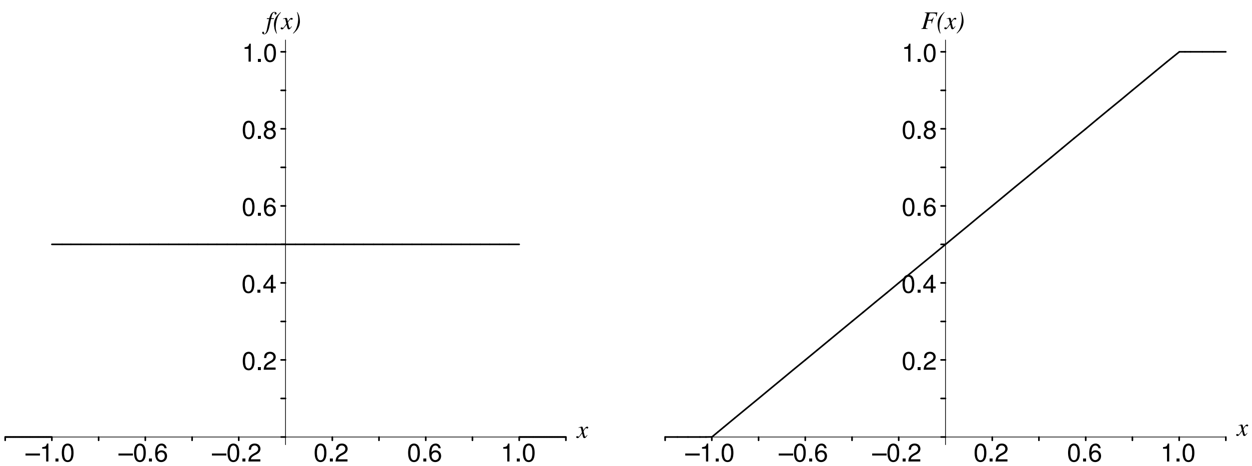


Figure 3.1–12: **(b)** $f(x) = 1/2$ and $F(x) = (x+1)/2$

(c)

$$F(x) = \begin{cases} 0, & -\infty < x < -1, \\ (x+1)^2/2, & -1 \le x < 0, \\ 1 - (1-x)^2/2, & 0 \le x < 1, \\ 1, & 1 \le x < \infty. \end{cases}$$

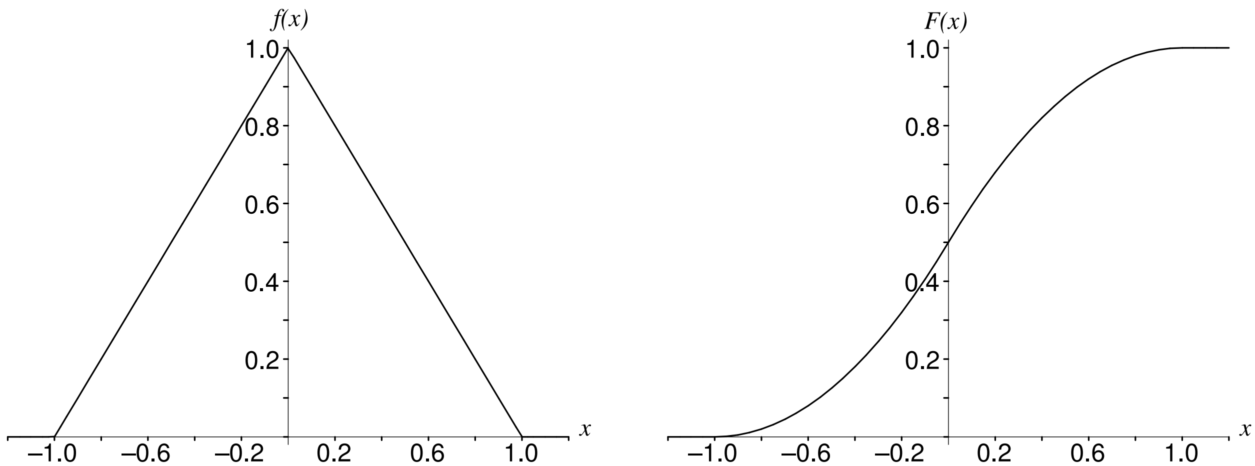


Figure 3.1–12: **(c)** $f(x)$ and $F(x)$ for Exercise **3.1-12(c)**

3.1–14 (b)

$$F(x) = \begin{cases} 0, & -\infty < x \le 0, \\ \dfrac{x}{2}, & 0 < x \le 1, \\ \dfrac{1}{2}, & 1 < x \le 2, \\ \dfrac{x}{2} - \dfrac{1}{2} & 2 \le x < 3, \\ 1, & 3 \le x < \infty; \end{cases}$$

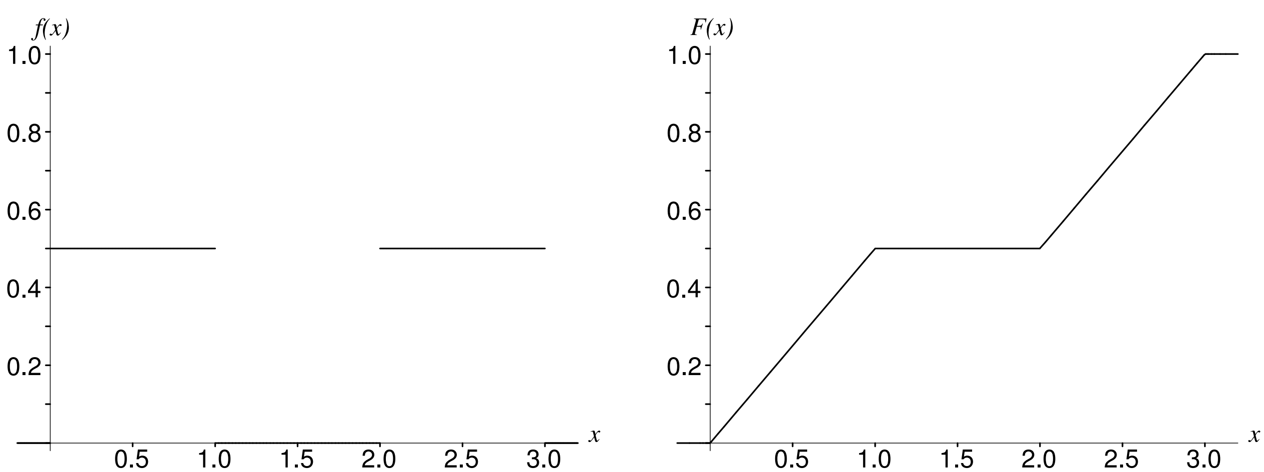


Figure 3.1–14: $f(x)$ and $F(x)$ for Exercise **3.1-14(a)** and **(b)**

(c)
$$\begin{aligned}\frac{q_1}{2} &= 0.25\\ q_1 &= 0.5,\end{aligned}$$

(d) $1 \le m \le 2$,

(e)
$$\begin{aligned}\frac{q_3}{2} - \frac{1}{2} &= 0.75\\ \frac{q_3}{2} &= \frac{5}{4}\\ q_3 &= \frac{5}{2}.\end{aligned}$$

3.1–16 $F(x) = (x+1)^2/4, \qquad -1 < x < 1.$

(a)
$$\begin{aligned}F(\pi_{0.64}) = (\pi_{0.64}+1)^2/4 &= 0.64\\ \pi_{0.64}+1 &= \sqrt{2.56}\\ \pi_{0.64} &= 0.6;\end{aligned}$$

(b)
$$\begin{aligned}(\pi_{0.25}+1)^2/4 &= 0.25\\ \pi_{0.25}+1 &= \sqrt{1.00}\\ \pi_{0.25} &= 0;\end{aligned}$$

(c)
$$\begin{aligned}(\pi_{0.81}+1)^2/4 &= 0.81\\ \pi_{0.81}+1 &= \sqrt{3.24}\\ \pi_{0.81} &= 0.8.\end{aligned}$$

3.1–18 $P(X > 2) = \displaystyle\int_2^\infty 4x^3 e^{-x^4}\,dx = \left[-e^{-x^4}\right]_2^\infty = e^{-16}.$

3.1–20 **(a)**
$$\begin{aligned}\int_0^1 x\,dx + \int_1^\infty \frac{c}{x^3}\,dx &= 1\\ \left[\frac{x^2}{2}\right]_0^1 - \left[\frac{c}{2x^2}\right]_1^\infty &= 1\\ \frac{1}{2} + \frac{c}{2} &= 1\\ c &= 1;\end{aligned}$$

(b) $E(X) = \displaystyle\int_0^1 x^2\,dx + \int_1^\infty \frac{1}{x^2}\,dx = \frac{4}{3};$

(c) the variance does not exist;

(d) $P(1/2 \le X \le 2) = \displaystyle\int_{1/2}^1 x\,dx + \int_1^2 \frac{1}{x^3}\,dx = \frac{3}{4}.$

3.2 The Exponential, Gamma, and Chi-Square Distributions

3.2–2 **(a)** $f(x) = \left(\dfrac{2}{3}\right) e^{-2x/3}, \qquad 0 \le x < \infty;$

(b) $P(X > 2) = \displaystyle\int_2^\infty \frac{2}{3} e^{-2x/3}\,dx = \left[-e^{-2x/3}\right]_2^\infty = e^{-4/3}.$

3.2–4 Let $F(x) = P(X \le x)$. Then

$$\begin{aligned} P(X > x + y \,|\, X > x) &= P(X > y) \\ \frac{1 - F(x+y)}{1 - F(x)} &= 1 - F(y). \end{aligned}$$

That is, with $g(x) = 1 - F(x)$, $g(x+y) = g(x)g(y)$. This functional equation implies that

$$1 - F(x) = g(x) = a^{cx} = e^{(cx)\ln a} = e^{bx}$$

where $b = c \ln a$. That is, $F(x) = 1 - e^{bx}$. Since $F(\infty) = 1$, b must be negative, say $b = -\lambda$ with $\lambda > 0$. Thus $F(x) = 1 - e^{-\lambda x}$, $0 \le x$, the distribution function of an exponential distribution.

3.2–6 **(a)**
$$\begin{aligned} P(X > 40) &= \int_{40}^{\infty} \frac{3}{100} e^{-3x/100}\, dx \\ &= \left[-e^{-3x/100}\right]_{40}^{\infty} = e^{-1.2}; \end{aligned}$$

(b) Flaws occur randomly so we are observing a Poisson process.

3.2–8 Either use integration by parts or

$$\begin{aligned} F(x) &= P(X \le x) \\ &= 1 - \sum_{k=0}^{\alpha-1} \frac{(\lambda x)^k e^{-\lambda x}}{k!}. \end{aligned}$$

Thus, with $\lambda = 1/\theta = 1/4$ and $\alpha = 2$,

$$\begin{aligned} P(X < 5) &= 1 - e^{-5/4} - \left(\frac{5}{4}\right) e^{-5/4} \\ &= 0.35536. \end{aligned}$$

3.2–10 The moment generating function of X is $M(t) = (1 - \theta t)^{-\alpha}$, $t < 1/\theta$. Thus

$$\begin{aligned} M'(t) &= \alpha\theta(1 - \theta t)^{-\alpha-1} \\ M''(t) &= \alpha(\alpha+1)\theta^2(1 - \theta t)^{-\alpha-2}. \end{aligned}$$

The mean and variance are

$$\begin{aligned} \mu &= M'(0) \;=\; \alpha\theta \\ \sigma^2 &= M''(0) - (\alpha\theta)^2 = \alpha(\alpha+1)\theta^2 - (\alpha\theta)^2 \\ &= \alpha\theta^2. \end{aligned}$$

3.2–12 **(a)** W has a gamma distribution with $\alpha = 7$, $\theta = 1/16$.

(b) Using Table III in the Appendix,

$$\begin{aligned} P(W \le 0.5) &= 1 - \sum_{k=0}^{6} \frac{8^k e^{-8}}{k!} \\ &= 1 - 0.313 = 0.687, \end{aligned}$$

because here $\lambda w = (16)(0.5) = 8$.

3.2–14 $a = 5.226$, $b = 21.03$.

3.2–16 Note that $\lambda = 5/10 = 1/2$ is the mean number of arrivals per minute. Thus $\theta = 2$ and the pdf of the waiting time before the eighth toll is

$$\begin{aligned} f(x) &= \frac{1}{\Gamma(8)2^8}\, x^{8-1} e^{-x/2} \\ &= \frac{1}{\Gamma\left(\dfrac{16}{2}\right)2^{16/2}}\, x^{16/2-1} e^{-x/2}, \qquad 0 < x < \infty, \end{aligned}$$

the pdf of a chi-square distribution with $r = 16$ degrees of freedom. Using Table IV,

$$P(X > 26.30) = 0.05.$$

3.2–18 **(a)** $\mu = \displaystyle\int_{80}^{\infty} x \cdot \frac{x-80}{50^2} e^{-(x-80)/50}\, dx$. Let $y = x - 80$. Then

$$\begin{aligned} \mu &= 80 + \int_0^{\infty} y \cdot \frac{1}{\Gamma(2)50^2} y^{2-1} e^{-y/50}\, dy \\ &= 80 + 2(50) \ = \ 180. \end{aligned}$$

$\mathrm{Var}(X) = \mathrm{Var}(Y) = 2(50^2) = 5000.$

(b)

$$\begin{aligned} f'(x) &= \frac{1}{50^2} e^{-(x-80)/50} - \frac{x-80}{50^2}\frac{1}{50} e^{-(x-80)/50} \ = \ 0 \\ 50 - x + 80 &= 0 \\ x &= 130. \end{aligned}$$

(c)

$$\begin{aligned} \int_{80}^{200} \frac{x-80}{50^2} e^{-(x-80)/50} dx &= \left[-\frac{x-80}{50} e^{-(x-80)/50} - e^{-(x-80)/50} \right]_{80}^{200} \\ &= \frac{-120}{50} e^{-120/50} - e^{-120/50} + 1 \\ &= 1 - \frac{17}{5} e^{-12/5} \ = \ 0.6916 \ = \ 69.16\%. \end{aligned}$$

3.2–20

$$\begin{aligned} E[v(T)] &= \int_0^3 100(2^{3-t} - 1) e^{-t/5}/5 dt \\ &= \int_0^3 -20 e^{-t/5} dt + 100 \int_0^3 e^{(3-t)\ln 2} e^{-t/5}/5 dt \\ &= -100(1 - e^{-0.6}) + 100 e^{3\ln 2} \int_0^3 e^{-t\ln 2} e^{-t/5}/5 dt \\ &= -100(1 - e^{-0.6}) + 100 e^{3\ln 2} \left[-\frac{e^{-(\ln 2 + 0.2)t}}{\ln 2 + 0.2} \right]_0^3 \ = \ \$121.734. \end{aligned}$$

3.2–22

$$\begin{aligned} F(x) &= \int_{-\infty}^{x} \frac{e^{-w}}{(1+e^{-w})^2}\, dw = \frac{1}{1+e^{-x}}, \qquad -\infty < x < \infty. \\ G(y) &= P\left[\frac{1}{1+e^{-X}} \le y\right] \ = \ P\left[X \le -\ln\left(\frac{1}{y} - 1\right)\right] \\ &= \frac{1}{1 + \left(\dfrac{1}{y} - 1\right)} \ = \ y, \qquad 0 < y < 1, \end{aligned}$$

the $U(0, 1)$ distribution function.

3.2–24 $P(X > 100 \,|\, X > 50) = P(X > 50) = 3/4.$

3.3 The Normal Distribution

3.3–2 **(a)** 0.3078: **(b)** 0.4959; **(c)** 0.2711; **(d)** 0.1646;

(e) 0.0526; **(f)** 0.3174; **(g)** 0.0456; **(h)** 0.0026.

3.3–4 **(a)** 1.282; **(b)**-1.645; **(c)** -1.66; **(d)** -1.82.

3.3–6 $M(t) = e^{166t+400t^2/2}$ so

(a) $\mu = 166$; **(b)** $\sigma^2 = 400$;

(c) $P(170 < X < 200) = P(0.2 < Z < 1.7) = 0.3761$;

(d) $P(148 \le X \le 172) = P(-0.9 \le Z \le 0.3) = 0.4338$.

3.3–8 We must solve $f''(x) = 0$. We have

$$\begin{aligned}
\ln f(x) &= -\ln(\sqrt{2\pi}\,\sigma) - (x-\mu)^2/2\sigma^2, \\
\frac{f'(x)}{f(x)} &= \frac{-2(x-\mu)}{2\sigma^2} \\
\frac{f(x)f''(x) - [f'(x)]^2}{[f(x)]^2} &= \frac{-1}{\sigma^2} \\
f''(x) &= f(x)\left\{\frac{-1}{\sigma^2} + \left[\frac{f'(x)}{f(x)}\right]\right\} \;=\; 0 \\
\frac{(x-\mu)^2}{\sigma^4} &= \frac{1}{\sigma^2} \\
x - \mu &= \pm\sigma \qquad \text{or} \qquad x = \mu \pm \sigma.
\end{aligned}$$

3.3–10
$$\begin{aligned}
G(y) &= P(Y \le y) \;=\; P(aX + b \le y) \\
&= P\left(X \le \frac{y-b}{a}\right) \qquad \text{if} \;\; a > 0 \\
&= \int_{-\infty}^{(y-b)/a} \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}\,dx
\end{aligned}$$

Let $w = ax + b$ so $dw = a\;dx$. Then

$$G(y) = \int_{-\infty}^{y} \frac{1}{a\sigma\sqrt{2\pi}}\, e^{-(w-b-a\mu)^2/2a^2\sigma^2}\;dw$$

which is the distribution function of the normal distribution $N(b + a\mu,\; a^2\sigma^2)$. The case when $a < 0$ can be handled similarly.

3.3–12 X is $N(500, 10000)$; so $[(X-500)^2/100]^2$ is $\chi^2(1)$ and

$$P\left[2.706 \le \left(\frac{X-500}{100}\right)^2 \le 5.204\right] = 0.975 - 0.900 = 0.075.$$

3.3–14
$$\begin{aligned}
G(x) &= P(X \le x) \\
&= P(e^Y \le x) \\
&= P(Y \le \ln x) \\
&= \int_{-\infty}^{\ln x} \frac{1}{\sqrt{2\pi}} e^{-(y-10)^2/2} dy \;=\; \Phi(\ln x - 10)
\end{aligned}$$

$$g(x) = G'(x) = \frac{1}{\sqrt{2\pi}} e^{-(\ln x - 10)^2/2} \frac{1}{x}, \qquad 0 < x < \infty.$$

$$\begin{aligned}
P(10{,}000 < X < 20{,}000) &= P(\ln 10{,}000 < Y < \ln 20{,}000) \\
&= \Phi(\ln 20{,}000 - 10) - \Phi(\ln 10{,}000 - 10) \\
&= 0.461557 - 0.214863 \;=\; 0.246694.
\end{aligned}$$

3.3–16 **(a)** $N(0,1)$; **(b)** $N(-1,1)$; **(c)** $N(2,1)$. Note the slopes of these graphs at 0.

3.4 Additional Models

3.4–2 With $b = \ln 1.1$,

$$\begin{aligned}
G(w) &= 1 - \exp\left[-\frac{a}{\ln 1.1} e^{w \ln 1.1} + \frac{a}{\ln 1.1}\right] \\
G(64) - G(63) &= 0.01 \\
a &= 0.00002646 = \frac{1}{37792.19477} \\
P(W \le 71 \,|\, 70 < W) &= \frac{P(70 < W \le 71)}{P(70 < W)} \\
&= 0.0217.
\end{aligned}$$

3.4–4
$$\begin{aligned}
\lambda(w) &= ae^{bw} + c \\
H(w) &= \textstyle\int_0^w (ae^{bt} + c)\, dt \\
&= \frac{a}{b}\left(e^{bw} - 1\right) + cw \\
G(w) &= 1 - \exp\left[-\frac{a}{b}\left(e^{bw} - 1\right) - cw\right], \qquad 0 < \infty \\
g(w) &= (ae^{bw} + c)e^{-\frac{a}{b}(e^{bw} - 1) - cw}, \qquad 0 < \infty.
\end{aligned}$$

3.4–6
(a) $1/4 - 1/8 = 1/8$; **(b)** $1/4 - 1/4 = 0$;
(c) $3/4 - 1/4 = 1/2$; **(d)** $1 - 1/2 = 1/2$;
(e) $3/4 - 3/4 = 0$; **(f)** $1 - 3/4 = 1/4$.

3.4–8 There is a discrete point of probability at $x = 0$, $P(X = 0) = 1/3$, and $F'(x) = (2/3)e^{-x}$ for $0 < x$. Thus

$$\begin{aligned}
\mu = E(X) &= (0)(1/3) + \int_0^\infty x(2/3)e^{-x}dx \\
&= (2/3)\left[-xe^{-x} - e^{-x}\right]_0^\infty \;=\; 2/3, \\
E(X^2) &= (0)^2(1/3) + \int_0^\infty x^2(2/3)e^{-x}dx \\
&= (2/3)[-x^2e^{-x} - 2xe^{-x} - 2e^{-x}]_0^\infty \;=\; 4/3,
\end{aligned}$$

so

$$\sigma^2 \;=\; \mathrm{Var}(X) \;=\; 4/3 - (2/3)^2 \;=\; 8/9.$$

3.4–10 $$T = \begin{cases} X, & X \le 4, \\ 4, & 4 < X; \end{cases}$$

$$\begin{aligned}
E(T) &= \int_0^4 x\left(\frac{1}{5}\right)e^{-x/5}\,dx + \int_4^\infty 4\left(\frac{1}{5}\right)e^{-x/5}\,dx \\
&= \left[-xe^{-x/5} - 5e^{-x/5}\right]_0^4 + 4\left[-e^{-x/5}\right]_4^\infty \\
&= 5 - 4e^{-4/5} - 5e^{-4/5} + 4e^{-4/5} \\
&= 5 - 5e^{-4/5} \;\approx\; 2.753.
\end{aligned}$$

3.4–12 **(a)**
$$\begin{aligned}
t &= \ln x \\
x &= e^t \\
\frac{dx}{dt} &= e^t \\
g(t) &= f(e^t)\frac{dx}{dt} = e^t e^{-e^t}, \qquad -\infty < t < \infty.
\end{aligned}$$

(b)
$$\begin{aligned}
t &= \alpha + \beta \ln w \\
\frac{dt}{dw} &= \frac{\beta}{w} \\
h(w) &= e^{\alpha+\beta\ln w}e^{-e^{\alpha+\beta\ln w}}\left(\frac{\beta}{w}\right) \\
&= \beta w^{\beta-1}e^{\alpha}e^{-w^{\beta}e^{\alpha}}, \qquad 0 < w < \infty.
\end{aligned}$$

3.4–14 **(a)** $((0.03)\displaystyle\int_{2/30}^1 6(1-x)^5\,dx = 0.0198;$

(b) $E(X) = (0.97)(0) + 0.03\displaystyle\int_0^1 x6(1-x)^5\,dx = 0.0042857;$

The expected payment is $E(X)\cdot[\$\,30{,}000] = \$\,128.57$.

3.4–16
$$\begin{aligned}
2500m\int_0^1 \frac{1}{10}e^{-x/10}\,dx + (m/2)2500\int_1^2 \frac{1}{10}e^{-x/10}\,dx &= 200\\
2500m[1-e^{-1/10}] + 1250m[e^{-1/10}-e^{-2/10}] &= 200\\
2500m - 1250me^{-1/10} - 1250me^{-2/10} &= 200\\
m &= \frac{4}{50-25e^{-1/10}-25e^{-2/10}}\\
&= 0.5788.
\end{aligned}$$

3.4–18 $P(X>x) = \displaystyle\int_x^\infty \left(\frac{t}{4}\right)^3 e^{-(t/4)^4}\,dt = e^{-(x/4)^4}$;

$$P(X>5\,|\,X>4) = \frac{P(X>5)}{P(X>4)} = \frac{e^{-625/256}}{e^{-1}} = e^{-369/256}.$$

3.4–20 **(a)** $\displaystyle\int_{40}^{60} \frac{2x}{50^2}e^{-(x/50)^2}\,dx = \left[-e^{-(x/50)^2}\right]_{40}^{60} = e^{-16/25} - e^{-36/25}$;

(b) $P(X>80) = \left[-e^{-(x/50)^2}\right]_{80}^{\infty} = e^{-64/25}$.

Chapter 4

Bivariate Distributions

4.1 Bivariate Distributions of the Discrete Type

4.1–2

$\frac{4}{16}$	4	• $\frac{1}{16}$	• $\frac{1}{16}$	• $\frac{1}{16}$	• $\frac{1}{16}$
$\frac{4}{16}$	3	• $\frac{1}{16}$	• $\frac{1}{16}$	• $\frac{1}{16}$	• $\frac{1}{16}$
$\frac{4}{16}$	2	• $\frac{1}{16}$	• $\frac{1}{16}$	• $\frac{1}{16}$	• $\frac{1}{16}$
$\frac{4}{16}$	1	• $\frac{1}{16}$	• $\frac{1}{16}$	• $\frac{1}{16}$	• $\frac{1}{16}$
		1	2	3	4
		$\frac{4}{16}$	$\frac{4}{16}$	$\frac{4}{16}$	$\frac{4}{16}$

(e) Independent, because $f_X(x)f_Y(y) = f(x,y)$.

4.1–4

$\frac{1}{25}$	12									• $\frac{1}{25}$
$\frac{1}{25}$	11									• $\frac{1}{25}$
$\frac{2}{25}$	10							• $\frac{1}{25}$		• $\frac{1}{25}$
$\frac{2}{25}$	9							• $\frac{1}{25}$		• $\frac{1}{25}$
$\frac{3}{25}$	8					• $\frac{1}{25}$		• $\frac{1}{25}$		• $\frac{1}{25}$
$\frac{2}{25}$	7					• $\frac{1}{25}$		• $\frac{1}{25}$		
$\frac{3}{25}$	6			• $\frac{1}{25}$		• $\frac{1}{25}$		• $\frac{1}{25}$		
$\frac{2}{25}$	5			• $\frac{1}{25}$		• $\frac{1}{25}$				
$\frac{3}{25}$	4	• $\frac{1}{25}$		• $\frac{1}{25}$		• $\frac{1}{25}$				
$\frac{2}{25}$	3	• $\frac{1}{25}$		• $\frac{1}{25}$						
$\frac{2}{25}$	2	• $\frac{1}{25}$		• $\frac{1}{25}$						
$\frac{1}{25}$	1	• $\frac{1}{25}$								
$\frac{1}{25}$	0	• $\frac{1}{25}$								
		0	1	2	3	4	5	6	7	8
		$\frac{1}{5}$		$\frac{1}{5}$		$\frac{1}{5}$		$\frac{1}{5}$		$\frac{1}{5}$

(c) Not independent, because $f_X(x)f_Y(y) \neq f(x,y)$ and also because the support is not rectangular.

4.1–6 $\dfrac{25!}{7!8!6!4!}(0.30)^7(0.40)^8(0.20)^6(0.10)^4 = 0.00405.$

4.1–8 **(a)** $f(x,y) = \dfrac{7!}{x!y!(7-x-y)!}(0.78)^x(0.01)^y(0.21)^{7-x-y}, \qquad 0 \le x+y \le 7;$

(b) X is $b(7, 0.78), \qquad x = 0, 1, \ldots, 7.$

4.2 The Correlation Coefficient

4.2–2 **(c)**

$$\begin{aligned}
\mu_X &= 0.5(0) + 0.5(1) \;=\; 0.5,\\
\mu_Y &= 0.2(0) + 0.6(1) + 0.2(2) \;=\; 1,\\
\sigma_X^2 &= (0-0.5)^2(0.5) + (1-0.5)^2(0.5) \;=\; 0.25,\\
\sigma_Y^2 &= (0-1)^2(0.2) + (1-1)^2(0.6) + (2-1)^2(0.2) \;=\; 0.4,\\
\mathrm{Cov}(X,Y) &= (0)(0)(0.2) + (1)(2)(0.2) + (0)(1)(0.3) \;+\\
&\quad\; (1)(1)(0.3) - (0.5)(1) = 0.2,\\
\rho &= \frac{0.2}{\sqrt{0.25}\sqrt{0.4}} \;=\; \sqrt{0.4};
\end{aligned}$$

(d) $y = 1 + \sqrt{0.4}\left(\dfrac{\sqrt{0.4}}{\sqrt{0.25}}\right)(x-0.5) = 0.6 + 0.8x.$

4.2–4 Note that X is $b(3,\ 1/6)$, Y is $b(3,\ 1/2)$ so

(a) $E(X) = 3(1/6) = 1/2;$
(b) $E(Y) = 3(1/2) = 3/2;$
(c) $\mathrm{Var}(X) = 3(1/6)(5/6) = 5/12;$
(d) $\mathrm{Var}(Y) = 3(1/2)(1/2) = 3/4;$
(e)

$$\begin{aligned}
\mathrm{Cov}(X,\,Y) &= 0 + (1)f(1,1) + 2f(1,2) + 2f(2,1) - (1/2)(3/2)\\
&= (1)(1/6) + 2(1/8) + 2(1/24) - 3/4\\
&= -1/4;
\end{aligned}$$

(f) $\rho = \dfrac{-1/4}{\sqrt{\dfrac{5}{12}\cdot\dfrac{3}{4}}} = \dfrac{-1}{\sqrt{5}}.$

4.2–6 **(b)**

$$\begin{array}{cc|ccc}
\frac{1}{6} & 2 & \bullet\,\frac{1}{6} & & \\
\frac{2}{6} & 1 & \bullet\,\frac{1}{6} & \bullet\,\frac{1}{6} & \\
\frac{3}{6} & 0 & \bullet\,\frac{1}{6} & \bullet\,\frac{1}{6} & \bullet\,\frac{1}{6} \\
\hline
 & & 0 & 1 & 2 \\
 & & \frac{3}{6} & \frac{2}{6} & \frac{1}{6}
\end{array}$$

(c) $\mathrm{Cov}(X,Y) = (1)(1)\left(\dfrac{1}{6}\right) - \left(\dfrac{2}{3}\right)\left(\dfrac{2}{3}\right) = \dfrac{1}{6} - \dfrac{4}{9} = \dfrac{-5}{18};$

$$\textbf{(d)}\quad \sigma_X^2 = \frac{2}{6}+\frac{4}{6}-\left(\frac{2}{3}\right)^2 = \frac{5}{9} = \sigma_Y^2,$$

$$\rho = \frac{-5/18}{\sqrt{(5/9)(5/9)}} = -\frac{1}{2};$$

$$\textbf{(e)}\quad y = \frac{2}{3}-\frac{1}{2}\sqrt{\frac{5/9}{5/9}}\left(x-\frac{2}{3}\right)$$

$$y = 1-\frac{1}{2}\,x.$$

4.2–8 **(a)** $f_X(1) = 0.15,\ f_X(2) = 0.25,\ f_X(3) = 0.45,\ f_X(4) = 0.15;$

$f_Y(1) = 0.35,\ f_Y(2) = 0.65;\ \ \mu_X = 2.60;\ \mu_Y = 1.65;$

$\sigma_X^2 = 0.8400;\ \sigma_Y^2 = 0.2275;$

(b) $\mathrm{Cov}(X,Y) = -0.0900;\ \rho = -0.2059;$

(c) $E(C) = \$34.70.$

4.2–10 Note that

$$\begin{aligned} h(v) &= E\{[(X-\mu_X)+v(Y-\mu_Y)]^2\} \\ &= E[(X-mu_X)2+2E[(X-\mu_X)(Y-\mu_Y)+E[(Y-\mu_Y)^2] \\ &= \sigma_X^2+2\mathrm{Cov}(X,Y)v+\sigma_Y^2v^2 \ \geq\ 0. \end{aligned}$$

Thus the discriminant of this quadratic must be less than or equal to 0. So we have

$$\begin{aligned} [\mathrm{Cov}(X,Y)]^2-\sigma_X^2\sigma_Y^2 &\leq 0 \\ \rho^2 &\leq 1 \\ -1 &\leq \rho \leq 1. \end{aligned}$$

4.3 Conditional Distributions

4.3–2

2	$\frac{1}{4}$	$\frac{3}{4}$	$g(x\,\|\,2)$
1	$\frac{3}{4}$	$\frac{1}{4}$	$g(x\,\|\,1)$
	1	2	

equivalently, $g(x\,|\,y) = \dfrac{3-2|x-y|}{4}$,

$x = 1, 2,$ for $y = 1$ or 2;

	$h(y\,\|\,1)$	$h(y\,\|\,2)$
2	$\frac{1}{4}$	$\frac{3}{4}$
1	$\frac{3}{4}$	$\frac{1}{4}$
	1	2

equivalently, $h(y\,|\,x) = \dfrac{3-2|x-y|}{4}$,

$y = 1, 2,$ for $x = 1$ or 2;

$\mu_{X|1} = 5/4,\ \mu_{X|2} = 7/4,\ \mu_{Y|1} = 5/4,\ \mu_{Y|2} = 7/4;$

$\sigma^2_{X|1} = \sigma^2_{X|2} = \sigma^2_{Y|1} = \sigma^2_{Y|2} = 3/16.$

4.3–4 **(a)** X is $b(400,\ 0.75)$;

(b) $E(X) = 300, \quad \text{Var}(X) = 75$;

(c) $b(300,\ 2/3)$;

(d) $E(Y) = 200, \quad \text{Var}(Y) = 200/3$.

4.3–6 **(a)** $P(X = 500) = 0.40, \quad P(Y = 500) = 0.35,$

$P(Y = 500 \,|\, X = 500) = 0.50, \quad P(Y = 100 \,|\, X = 500) = 0.25$;

(b) $\mu_X = 485, \quad \mu_Y = 510, \quad \sigma_X^2 = 118{,}275, \quad \sigma_Y^2 = 130{,}900$;

(c) $\mu_{X|Y=100} = 2400/7, \quad \mu_{Y|X=500} = 525$;

(d) $\text{Cov}(X, Y) = 49650$;

(e) $\rho = 0.399$.

4.3–8 **(a)** X and Y have a trinomial distribution with $n = 30, \quad p_X = 1/6, \quad p_Y = 1/6$.

(b) The conditional pmf of X, given $Y = y$, is

$$b\left(n - y,\ \frac{p_X}{1 - p_Y}\right) = b(30 - y,\ 1/5).$$

(c) Since $E(X) = 5$ and $\text{Var}(X) = 25/6$,

$$E(X^2) = \text{Var}(X) + [E(X)]^2 = 25/6 + 25 = 175/6.$$

Similarly, $E(Y) = 5,\ \text{Var}(Y) = 25/6,\ E(Y^2) = 175/6$. The correlation coefficient is

$$\rho = -\sqrt{\frac{(1/6)(1/6)}{(5/6)(5/6)}} = -1/5$$

so

$$E(XY) = -1/5\sqrt{(25/6)(25/6)} + (5)(5) = 145/6.$$

Thus

$$E(X^2 - 4XY + 3Y^2) = \frac{175}{6} - 4\left(\frac{145}{6}\right) + 3\left(\frac{175}{6}\right) = \frac{120}{6} = 20.$$

4.3–10 **(a)** $f(x, y) = 1/[10(10 - x)], \qquad x = 0, 1, \cdots, 9, \quad y = x, x + 1, \cdots, 9$;

(b) $f_Y(y) = \displaystyle\sum_{x=0}^{y} \frac{1}{10(10 - x)}, \qquad y = 0, 1, \ldots, 9$;

(c) $E(Y|x) = (x + 9)/2$.

4.4 Bivariate Distributions of the Continuous Type

4.4–2 **(a)**

$$\begin{aligned} f_X(x) &= \int_0^1 (x + y)\, dy \\ &= \left[xy + \frac{1}{2}y^2\right]_0^1 = x + \frac{1}{2}, \qquad 0 \le x \le 1; \\ f_Y(y) &= \int_0^1 (x + y)\, dx = y + \frac{1}{2}, \qquad 0 \le y \le 1; \\ f(x, y) &= x + y \ne \left(x + \frac{1}{2}\right)\left(y + \frac{1}{2}\right) = f_X(x) f_Y(y). \end{aligned}$$

(b) **(i)** $\mu_X = \displaystyle\int_0^1 \left(x + \frac{1}{2}\right) dx = \left[\frac{1}{3}x^3 + \frac{1}{4}x^2\right]_0^1 = \frac{7}{12}$;

(ii) $\mu_Y = \displaystyle\int_0^1 y\left(y+\frac{1}{2}\right)dy = \frac{7}{12};$

(iii)
$$\begin{aligned} E(X^2) &= \int_0^1 x^2\left(x+\frac{1}{2}\right)dx = \left[\frac{1}{4}x^4+\frac{1}{6}x^3\right]_0^1 = \frac{5}{12}, \\ \sigma_X^2 &= E(X^2)-\mu_X^2 = \frac{5}{12}-\left(\frac{7}{12}\right)^2 = \frac{11}{144}. \end{aligned}$$

(iv) Similarly, $\sigma_Y^2 = \dfrac{11}{144}$.

4.4–4 **(a)**
$$\begin{aligned} P\left(0 \le X \le \tfrac{1}{2}\right) &= \int_0^{\frac{1}{2}}\int_{x^2}^1 \frac{3}{2}\,dy\,dx \\ &= \int_0^{\frac{1}{2}} \frac{3}{2}(1-x^2)\,dx = \frac{11}{16}; \end{aligned}$$

(b)
$$\begin{aligned} P\left(\tfrac{1}{2} \le Y \le 1\right) &= \int_{\frac{1}{2}}^1\int_0^{\sqrt{y}} \frac{3}{2}\,dx\,dy \\ &= \int_{\frac{1}{2}}^1 \frac{3}{2}\sqrt{y}\,dy = 1-\left(\frac{1}{2}\right)^{3/2}; \end{aligned}$$

(c)
$$\begin{aligned} P\left(X \ge \tfrac{1}{2},\ Y \ge \tfrac{1}{2}\right) &= \int_{\frac{1}{2}}^1\int_{\frac{1}{2}}^{\sqrt{y}} \frac{3}{2}\,dx\,dy \\ &= \int_{\frac{1}{2}}^1 \frac{3}{2}\left(\sqrt{y}-\frac{1}{2}\right)dy \\ &= \frac{5}{8}-\left(\frac{1}{2}\right)^{3/2}; \end{aligned}$$

(d) X and Y are dependent.

4.4–6 **(a)** The variances are

$$\sigma_X^2 = E(X^2)-0^2 = \int_{-1}^1 x^2\cdot\frac{3}{2}x^2\,dx = \left[\frac{3x^5}{10}\right]_{-1}^1 = \frac{3}{5}$$

and

$$\begin{aligned} \sigma_Y^2 = E(Y^2)-0^2 &= \int_{-1}^1 y^2(1-|y|)\,dy \\ &= \int_{-1}^0 y^2(1+y)\,dy + \int_0^1 y^2(1-y)\,dy \\ &= \frac{1}{3}-\frac{1}{4}+\frac{1}{3}-\frac{1}{4} = \frac{1}{6}. \end{aligned}$$

(b)
$$\begin{aligned} P(-X \le Y) &= \int_0^1\int_{-y}^1 \frac{3}{2}x^2(1-y)\,dxdy + \int_{-1}^0\int_{-y}^1 \frac{3}{2}x^2(1+y)\,dxdy \\ &= \frac{1}{2}-\frac{1}{4}+\frac{1}{8}-\frac{1}{10}+\frac{1}{2}-\frac{1}{4}-\frac{1}{8}+\frac{1}{10} = \frac{1}{2}. \end{aligned}$$

4.4–8

$$\begin{aligned}
\mu_X = E(X) &= \int_{-1}^{0} x \cdot \frac{15}{7}\, x^2(1-x^2)\, dx + \int_0^1 x \cdot \frac{15}{7}\, x^2\, dx \\
&= \frac{15}{7}\left[\frac{x^4}{4} - \frac{x^6}{6}\right]_{-1}^{0} + \frac{15}{7}\left[\frac{x^4}{4}\right]_0^1 \\
&= \frac{15}{7}\left[\frac{-1}{4} + \frac{1}{6}\right] + \frac{15}{7}\left[\frac{1}{4}\right] \;=\; \frac{5}{14} \\
E(X^2) &= \int_{-1}^{0} x^2 \cdot \frac{15}{7}\, x^2(1-x^2)\, dx + \int_0^1 x^2 \cdot \frac{15}{7}\, x^2\, dx \\
&= \frac{15}{7}\left[\frac{x^5}{5} - \frac{x^7}{7}\right]_{-1}^{0} + \frac{15}{7}\left[\frac{x^5}{5}\right]_0^1 \\
&= \frac{15}{7}\left[\frac{1}{5} - \frac{1}{7}\right] + \frac{15}{7}\left[\frac{1}{5}\right] \;=\; \frac{27}{49} \\
\sigma_X^2 &= \frac{27}{49} - \left(\frac{5}{14}\right)^2 \;=\; \frac{83}{196}; \\
\mu_Y = E(Y) &= \textstyle\int_0^1 y \cdot \dfrac{10}{7}\,(y + y^4)\, dy \\
&= \frac{10}{7}\left[\frac{y^3}{3} + \frac{y^6}{6}\right]_0^1 \\
&= \frac{10}{7}\left[\frac{1}{3} + \frac{1}{6}\right] \;=\; \frac{5}{7} \\
E(Y^2) &= \textstyle\int_0^1 y^2 \cdot \dfrac{10}{7}\,(y + y^4)\, dy \\
&= \frac{10}{7}\left[\frac{y^4}{4} + \frac{y^7}{7}\right]_0^1 \\
&= \frac{10}{7}\left[\frac{1}{4} + \frac{1}{7}\right] \;=\; \frac{55}{98} \\
\sigma_Y^2 &= \frac{55}{98} - \left(\frac{5}{7}\right)^2 \;=\; \frac{5}{98}.
\end{aligned}$$

4.4–10 The area of the space is

$$\int_2^6 \int_1^{14-2t_2} dt_1 dt_2 = \int_2^6 (13 - 2t_2)\, dt_2 = 20;$$

Thus

$$\begin{aligned}
P(T_1 + T_2 > 10) &= \int_2^4 \int_{10-t_2}^{14-2t_2} \frac{1}{20} dt_1 dt_2 \\
&= \int_2^4 \frac{4 - t_2}{20} dt_2 \\
&= \left[-\frac{(4-t_2)^2}{40}\right]_2^4 \;=\; \frac{1}{10}.
\end{aligned}$$

4.4–12 $E[a_1u_1(X_1,\, X_2) + a_2u_2(X_1,\, X_2)]$

$$\begin{aligned}
&= \sum_{(x_1,x_2)}\sum_{\in R} [a_1u_1(x_1,\, x_2) + a_2u_2(x_1,\, x_2)]f(x_1,\, x_2) \\
&= a_1 \sum_{(x_1,\, x_2)}\sum_{\in R} u_1(x_1,\, x_2)f(x_1,\, x_2) + a_2 \sum_{(x_1,\, x_2)}\sum_{\in R} u_2(x_1,\, x_2)f(x_1,\, x_2) \\
&= a_1E[u_1(X_1,\, X_2)] + a_2E[u_2(X_1,\, X_2)].
\end{aligned}$$

4.4–14 **(a)**
$$\begin{aligned}
f_X(x) &= \int_x^1 8xy\,dy \;=\; 4x(1-x^2), \qquad 0 \le x \le 1, \\
f_Y(y) &= \int_0^y 8xy\,dx \;=\; 4y^3, \qquad 0 \le x \le 1;
\end{aligned}$$

(b)
$$\begin{aligned}
\mu_X &= \int_0^1 x4x(1-x^2)\,dx \;=\; \frac{8}{15}, \\
\mu_Y &= \int_0^1 y*4y^3\,dy \;=\; \frac{4}{5}, \\
\sigma_X^2 &= \int_0^1 (x-8/15)^2 4x(1-x^2)\,dx \;=\; \frac{11}{225}, \\
\sigma_Y^2 &= \int_0^1 (y-4/5)^2 * 4y^3\,dy \;=\; \frac{2}{75}, \\
\text{Cov}(X,Y) &= \int_0^1\int_x^1 (x-8/15)(y-4/5)8xy\,dy\,dx \;=\; \frac{4}{225}, \\
\rho &= \frac{4/225}{\sqrt{(11/225)(2/75)}} \;=\; \frac{2\sqrt{66}}{33};
\end{aligned}$$

(c) $y = \dfrac{20}{33} + \dfrac{4x}{11}$.

4.4–16 From Example 4.4–3, $\mu_X = \dfrac{1}{3}$, $\mu_Y = \dfrac{2}{3}$, and $E(Y^2) = \dfrac{1}{2}$;

$$E(X^2) = \int_0^1 2x^2(1-x)\,dx = \frac{1}{6}, \quad \sigma_X^2 = \frac{1}{6} - \left(\frac{1}{3}\right)^2 = \frac{1}{18}, \quad \sigma_Y^2 = \frac{1}{2} - \left(\frac{2}{3}\right)^2 = \frac{1}{18};$$

$$\text{Cov}(X,Y) = \int_0^1\int_x^1 2xy\,dy\,dx - \left(\frac{1}{3}\right)\left(\frac{2}{3}\right) \;=\; \frac{1}{4} - \frac{2}{9} \;=\; \frac{1}{36},$$

so

$$\rho = \frac{1/36}{\sqrt{1/18}\,\sqrt{1/18}} = \frac{1}{2}.$$

4.4–18 **(b)**

$$f_X(x) = \begin{cases} \displaystyle\int_0^x 1/8\,dy \;=\; x/8, & 0 \le x \le 2, \\ \displaystyle\int_{x-2}^x 1/8\,dy \;=\; 1/4, & 2 < x < 4, \\ \displaystyle\int_{x-2}^4 1/8\,dy \;=\; (6-x)/8, & 4 \le x \le 6; \end{cases}$$

(c) $f_Y(y) = \displaystyle\int_y^{y+2} 1/8\,dx = 1/4, \qquad 0 \le y \le 4;$

(d)

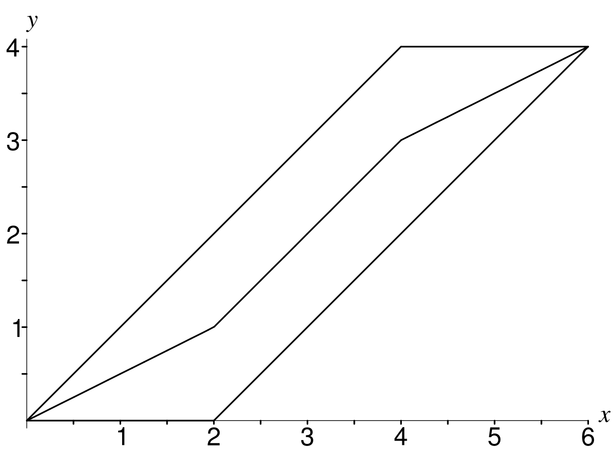
$$h(y\,|\,x) = \begin{cases} 1/x, & 0 \le y \le x, & 0 \le x \le 2, \\ 1/2, & x-2 < y < x, & 2 < x < 4, \\ 1/(6-x), & x-2 \le y \le 4, & 4 \le x \le 6; \end{cases}$$

(e) $g(x\,|\,y) = 1/2, \qquad y \le x \le y+2;$

(f)

$$E(Y\,|\,x) = \begin{cases} \displaystyle\int_0^x y\left(\frac{1}{x}\right)dy = \frac{x}{2}, & 0 \le x \le 2, \\ \displaystyle\int_{x-2}^x y\cdot\frac{1}{2}\,dy = \left[\frac{y^2}{4}\right]_{x-2}^x = x-1, & 2 < x < 4, \\ \displaystyle\int_{x-2}^4 \frac{y}{6-x}\,dy = \left[\frac{y^2}{2(6-x)}\right]_{x-2}^4 = \frac{x+2}{2}, & 4 \le x < 6; \end{cases}$$

(g) $E(X\,|\,y) = \displaystyle\int_y^{y+2} x\cdot\frac{1}{2}\,dx = \left[\frac{x^2}{4}\right]_y^{y+2} = y+1, \qquad 0 \le y \le 4;$



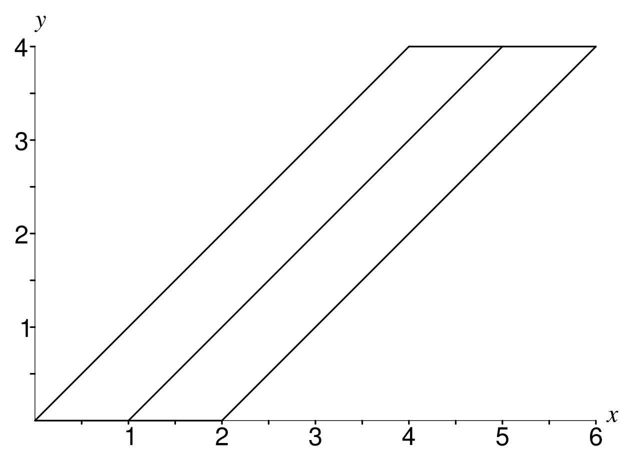


Figure 4.4–18: **(h)** $y = E(Y\,|\,x)$ **(i)** $x = E(X\,|\,y)$

4.4–20 **(a)** $f(x,y) = f_X(x)h(y\,|\,x) = 1\cdot\dfrac{1}{x+1} = \dfrac{1}{x+1}, \qquad 0 < y < x+1, \quad 0 < x < 1;$

(b) $E(Y\,|\,x) = \displaystyle\int_0^{x+1} y\left(\frac{1}{x+1}\right)dy = \left[\frac{y^2}{2(x+1)}\right]_0^{x+1} = \frac{x+1}{2};$

(c)

$$f_Y(y) = \begin{cases} \displaystyle\int_0^1 \frac{1}{x+1}\,dx = [\ln(x+1)]_0^1 = \ln 2, & 0 < y < 1, \\ \displaystyle\int_{y-1}^1 \frac{1}{x+1}\,dx = [\ln(x+1)]_{y-1}^1 = \ln 2 - \ln y, & 1 < y < 2. \end{cases}$$

4.5 The Bivariate Normal Distribution

4.5–2
$$\begin{aligned} q(x,y) &= \frac{[y-\mu_Y-\rho(\sigma_Y/\sigma_X)(x-\mu_X)]^2}{\sigma_Y^2(1-\rho^2)} + \frac{(x-\mu_X)^2}{\sigma_X^2} \\ &= \frac{1}{1-\rho^2}\left[\frac{(y-\mu_Y)^2}{\sigma_Y^2} - \frac{2\rho(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y}\right. \\ &\qquad \left. + \frac{\rho^2(x-\mu_X)^2}{\sigma_X^2} + (1-\rho^2)\frac{(x-\mu_X)^2}{\sigma_X^2}\right] \\ &= \frac{1}{1-\rho^2}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right) + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right] \end{aligned}$$

4.5–4 **(a)** $E(Y \mid X = 72) = 80 + \dfrac{5}{13}\left(\dfrac{13}{10}\right)(72-70) = 81$;

(b) $\text{Var}(Y \mid X = 72) = 169\left[1 - \left(\dfrac{5}{13}\right)^2\right] = 144$;

(c) $P(Y \le 84 \mid X = 72) = P\left(Z \le \dfrac{84-81}{12}\right) = \Phi(0.25) = 0.5987.$

4.5–6 **(a)** $P(18.5 < Y < 25.5) = \Phi(0.8) - \Phi(-1.2) = 0.6730$;

(b) $E(Y \mid x) = 22.7 + 0.78(3.5/4.2)(x - 22.7) = 0.65x + 7.945$;

(c) $\text{Var}(Y \mid x) = 12.25(1 - 0.78^2) = 4.7971$;

(d) $P(18.5 < Y < 25.5 \mid X = 23) = \Phi(1.189) - \Phi(-2.007) = 0.8828 - 0.0224 = 0.8604$;

(e) $P(18.5 < Y < 25.5 \mid X = 25) = \Phi(0.596) - \Phi(-2.60) = 0.7244 - 0.0047 = 0.7197.$

(f)

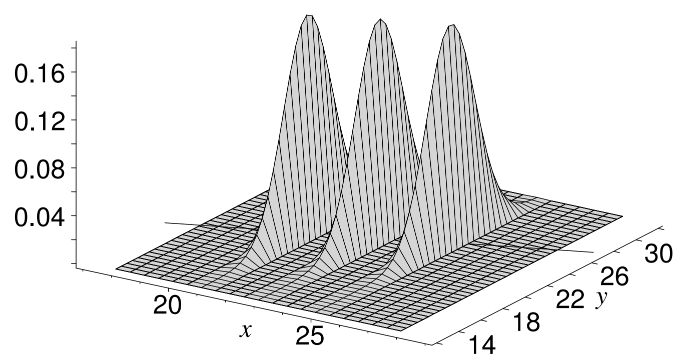


Figure 4.5–6: Conditional pdfs of Y, given $x = 21, 23, 25$

4.5–8 **(a)** $P(13.6 < Y < 17.2) = \Phi(0.55) - \Phi(-0.35) = 0.3456$;

(b) $E(Y \mid x) = 15 + 0(4/3)(x - 10) = 15$;

(c) $\text{Var}(Y \mid x) = 16(1 - 0^2) = 16$;

(d) $P(13.6 < Y < 17.2 \mid X = 9.1) = 0.3456.$

4.5–10 **(a)** $P(2.80 \le Y \le 5.35) = \Phi(1.50) - \Phi(0) = 0.4332$;

(b) $E(Y \mid X = 82.3) = 2.80 + (-0.57)\left(\dfrac{1.7}{10.5}\right)(82.3 - 72.30) = 1.877$;

$$\text{Var}(Y \mid X = 82.3) = 2.89[1 - (-0.57)^2] = 1.9510;$$

$$\begin{aligned} P(2.76 \le Y \le 5.34 \mid X = 82.3) &= \Phi(2.479) - \Phi(0.632) \\ &= 0.9934 - 0.7363 \ = \ 0.2571. \end{aligned}$$

4.5–12

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} \frac{1}{2\pi}\, e^{-(x^2+y^2)/2}\, dy + \int_{-\infty}^{\infty} \frac{1}{2\pi}\, e^{-(x^2+y^2)/2} xy e^{-(x^2+y^2-2)/2}\, dy \\ &= \frac{1}{\sqrt{2\pi}}\, e^{-x^2/2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}\, e^{-y^2/2}\, dy + 0 \\ &= \frac{1}{\sqrt{2\pi}}\, e^{-x^2/2}, \qquad -\infty < x < \infty. \end{aligned}$$

Note that the first integrand is the product of two $N(0,1)$ pdfs and the integral of a pdf is equal to 1. The second integral is an odd function so it is equal to 0.

Similarly, $f_Y(y) = \dfrac{1}{\sqrt{2\pi}}\, e^{-y^2/2}, \qquad -\infty < y < \infty.$

It also follows that $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y)\, dx\, dy = \int_{-\infty}^{\infty} f_X(x)\, dx = 1.$

Chapter 5

Distributions of Functions of Random Variables

5.1 Functions of One Random Variable

5.1–2 Here $x = \sqrt{y}$, $\dfrac{dx}{dy} = \dfrac{1}{2\sqrt{y}}$ and $0 < x < \infty$ maps onto $0 < y < \infty$. Thus

$$g(y) = \sqrt{y}\left|\frac{1}{2\sqrt{y}}\right| = \frac{1}{2}e^{-y/2}, \qquad 0 < y < \infty.$$

5.1–4 **(a)**

$$F(x) = \begin{cases} 0, & x < 0, \\ \displaystyle\int_0^x 2t\,dt = x^2, & 0 \le x < 1, \\ 1, & 1 \le x, \end{cases}$$

(b) Let $y = x^2$; so $x = \sqrt{y}$. Let Y be $U(0,1)$; then $X = \sqrt{Y}$ has the given x-distribution.

(c) Repeat the procedure outlined in part (b) 10 times.

5.1–6 It is easier to note that

$$\frac{dy}{dx} = \frac{e^{-x}}{(1+e^{-x})^2} \qquad \text{and} \qquad \frac{dx}{dy} = \frac{(1+e^{-x})^2}{e^{-x}}.$$

Say the solution of x in terms of y is given by x^*. Then the pdf of Y is

$$g(y) = \frac{e^{-x^*}}{(1+e^{-x^*})^2}\left|\frac{(1+e^{-x^*})^2}{e^{-x^*}}\right| = 1, \qquad 0 < y < 1,$$

as $-\infty < x < \infty$ maps onto $0 < y < 1$. Thus Y is $U(0,1)$.

5.1–8

$$\begin{aligned} x &= \left(\frac{y}{5}\right)^{10/7} \\ \frac{dx}{dy} &= \frac{10}{7}\left(\frac{y}{5}\right)^{3/7}\left(\frac{1}{5}\right) \\ f(x) &= e^{-x}, \qquad 0 < x < \infty \\ g(y) &= e^{-(y/5)^{10/7}}\left(\frac{2}{7}\right)\left(\frac{1}{5}\right)^{3/7} y^{3/7} \\ &= \frac{10/7}{5^{10/7}}y^{3/7}e^{-(y/5)^{10/7}}, \qquad 0 < y < \infty. \end{aligned}$$

(The reason for writing the pdf in that form is because Y has a Weibull distribution with $\alpha = 10/7$ and $\beta = 5$.)

5.1–10 Since $-1 < x < 3$, we have $0 \le y < 9$.

When $0 < y < 1$, then

$$x_1 = -\sqrt{y}, \quad \frac{dx_1}{dy} = \frac{-1}{2\sqrt{y}}; \qquad x_2 = \sqrt{y}, \quad \frac{dx_2}{dy} = \frac{1}{2\sqrt{y}}.$$

When $1 < y < 9$, then

$$x = \sqrt{y}, \quad \frac{dx}{dy} = \frac{1}{2\sqrt{y}}.$$

Thus

$$g(y) = \begin{cases} \dfrac{1}{4}\cdot\left|\dfrac{-1}{2\sqrt{y}}\right| + \dfrac{1}{4}\cdot\left|\dfrac{1}{2\sqrt{y}}\right| = \dfrac{1}{4\sqrt{y}}, & 0 < y < 1, \\ \dfrac{1}{4}\cdot\left|\dfrac{1}{2\sqrt{y}}\right| = \dfrac{1}{8\sqrt{y}}, & 1 \le y < 9. \end{cases}$$

5.1–12

$$\begin{aligned} E(X) &= \int_{-\infty}^{\infty} \frac{x}{\pi(1+x^2)}\,dx \\ &= \lim_{a\to-\infty}\left[\frac{1}{2\pi}\ln(1+x^2)\right]_a^0 + \lim_{b\to+\infty}\left[\frac{1}{2\pi}\ln(1+x^2)\right]_0^b \\ &= \frac{1}{2\pi}\left[\lim_{a\to-\infty}\{-\ln(1+a^2)\} + \lim_{b\to+\infty}\ln(1+b^2)\right]. \end{aligned}$$

$E(X)$ does not exist because neither of these limits exists.

5.1–14 X is $N(0,1)$ and $Y = |\,X\,|$. Let

$$\begin{aligned} x_1 &= -y, & -\infty < x_1 < 0, \\ x_2 &= y, & 0 < x_2 < \infty. \end{aligned}$$

Then

$$\frac{dx_1}{dy} = -1 \quad \text{and} \quad \frac{dx_2}{dy} = 1.$$

Thus the pdf of Y is

$$g(y) = \frac{1}{\sqrt{2\pi}}e^{-(-y^2)}|-1| + \frac{1}{\sqrt{2\pi}}e^{-y^2}|\,1\,| = \frac{2}{\sqrt{2\pi}}e^{-y^2}, \qquad 0 < y < \infty.$$

5.2 Transformations of Two Random Variables

5.2-2 **(a)** The joint pdf of X_1 and X_2 is

$$f(x_1, x_2) = \frac{1}{\Gamma\left(\frac{r_1}{2}\right)\Gamma\left(\frac{r_2}{2}\right)2^{(r_1+r_2)/2}}\, x_1^{r_1/2-1} x_2^{r_2/2-1}\, e^{-(x_1+x_2)/2},$$

$$0 < x_1 < \infty, \quad 0 < x_2 < \infty.$$

Let $Y_1 = (X_1/r_1)/(X_2/r_2)$ and $Y_2 = X_2$. The Jacobian of the transformation is $(r_1/r_2)y_2$. Thus

$$g(y_1, y_2) = \frac{1}{\Gamma\left(\frac{r_1}{2}\right)\Gamma\left(\frac{r_2}{2}\right)2^{(r_1+r_2)/2}}\left(\frac{r_1 x_1 x_2}{r_2}\right)^{r_1/2-1} x_2^{r_2/2-1} e^{-(y_2/2)(r_1 y_1/r_2+1)}\left(\frac{r_1 y_2}{r_2}\right),$$

$$0 < y_1 < \infty, \quad 0 < y_2 < \infty.$$

(b) The marginal pdf of Y_1 is $g_1(y_1) = \displaystyle\int_0^\infty g(y_1,\, y_2)\, dy_2$.

Make the change of variables $w = \dfrac{y_2}{2}\left(\dfrac{r_1 y_1}{r_2} + 1\right)$. Then

$$g_1(y_1) = \frac{\Gamma\left(\dfrac{r_1+r_2}{2}\right)\left(\dfrac{r_1}{r_2}\right)^{r_1/2} y_1^{r_1/2-1}}{\Gamma\left(\dfrac{r_1}{2}\right)\Gamma\left(\dfrac{r_2}{2}\right)\left(1+\dfrac{r_1 y_1}{r_2}\right)^{(r_1+r_2)/2}} \cdot 1, \qquad 0 < y_1 < \infty.$$

5.2-4 **(a)** $F_{0.05}(9, 24) = 2.30$;

(b) $F_{0.95}(9, 24) = \dfrac{1}{F_{0.05}(24, 9)} = \dfrac{1}{2.90} = 0.3448$;

(c) $P(W < 0.277) = P\left(\dfrac{1}{W} > \dfrac{1}{0.277}\right) = P\left(\dfrac{1}{W} > 3.61\right) = 0.025$;

$P(0.277 \le W \le 2.70) = P(W \le 2.70) - P(W \le 0.277) = 0.975 - 0.025 = 0.95.$

5.2-6

$$\begin{aligned}
F(w) &= P\left(\frac{X_1}{X_1+X_2} \le w\right), \qquad 0 < w < 1 \\
&= \int_0^\infty \int_{(1-w)x_1/w}^\infty \frac{x_1^{\alpha-1} x_2^{\beta-1} e^{-(x_1+x_2)/\theta}}{\Gamma(\alpha)\Gamma(\beta)\theta^{\alpha+\beta}}\, dx_2 dx_1 \\
f(w) = F'(w) &= \int_0^\infty \frac{-x_1^{\alpha-1}[(1-w)x_1/w]^{\beta-1}\, e^{-[x_1+(1-w)x_1/w]/\theta}}{\Gamma(\alpha)\Gamma(\beta)\theta^{\alpha+\beta}} \left(\frac{-1}{w^2}\right) x_1\, dx_1 \\
&= \frac{1}{\Gamma(\alpha)\Gamma(\beta)} \frac{(1-w)^{\beta-1}}{w^{\beta+1}} \int_0^\infty \frac{x_1^{\alpha+\beta-1}\, e^{-x_1/\theta w}}{\theta^{\alpha+\beta}}\, dx_1 \\
&= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha+\beta)} \frac{(\theta w)^{\alpha+\beta}}{w^{\beta+1}} \frac{(1-w)^{\beta-1}}{\theta^{\alpha+\beta}} \\
&= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\, w^{\alpha-1}(1-w)^{\beta-1}, \qquad 0 < w < 1.
\end{aligned}$$

5.2-8 **(a)**

$$\begin{aligned}
E(X) &= \int_0^1 x\, \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\, x^{\alpha-1}(1-x)^{\beta-1}\, dx \\
&= \frac{\Gamma(\alpha+\beta)\Gamma(\alpha+1)}{\Gamma(\alpha)\Gamma(\alpha+\beta+1)} \cdot \int_0^1 \frac{\Gamma(\alpha+1+\beta)}{\Gamma(\alpha+1)\Gamma(\beta)}\, x^{\alpha+1-1}(1-x)^{\beta-1}\, dx \\
&= \frac{(\alpha)\Gamma(\alpha)\Gamma(\alpha+\beta)}{(\alpha+\beta)\Gamma(\alpha+\beta)\Gamma(\alpha)} \\
&= \frac{\alpha}{\alpha+\beta};
\end{aligned}$$

$$\begin{aligned}
E(X^2) &= \frac{\Gamma(\alpha+\beta)\Gamma(\alpha+2)}{\Gamma(\alpha)\Gamma(\alpha+2+\beta)} \int_0^1 \frac{\Gamma(\alpha+2+\beta)}{\Gamma(\alpha+2)\Gamma(\beta)}\, x^{\alpha+2-1}(1-x)^{\beta-1}\, dx \\
&= \frac{(\alpha+1)\alpha}{(\alpha+\beta+1)(\alpha+\beta)}.
\end{aligned}$$

Thus

$$\sigma^2 = \frac{\alpha(\alpha+1)}{(\alpha+\beta+1)(\alpha+\beta)} - \frac{\alpha^2}{(\alpha+\beta)^2} = \frac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2}.$$

(b)
$$\begin{aligned} f(x) &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\, x^{\alpha-1}(1-x)^{\beta-1} \\ f'(x) &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\left[(\alpha-1)\,x^{\alpha-2}\,(1-x)^{\beta-1} - (\beta-1)\,x^{\alpha-1}\,(1-x)^{\beta-2}\right]. \end{aligned}$$

Set $f'(x)$ equal to zero and solve for x:

$$\begin{aligned} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\, x^{\alpha-2}\,(1-x)^{\beta-2}\,[(\alpha-1)(1-x)-(\beta-1)\,x] &= 0 \\ \alpha - \alpha x - 1 + x - \beta x + x &= 0 \\ (\alpha+\beta-2)x &= \alpha - 1 \\ x &= \frac{\alpha-1}{\alpha+\beta-2}. \end{aligned}$$

5.2-10 Use integration by parts two times to show

$$\begin{aligned} \int_0^p \frac{6!}{3!2!}\, y^3(1-y)^2\,dy &= \left[\binom{6}{4}y^4(1-y)^2 + \binom{6}{5}y^5(1-y)^1 + \binom{6}{6}y^6(1-y)^0\right]_0^p \\ &= \sum_{y=4}^{6}\binom{n}{y}p^y(1-p)^{6-y}. \end{aligned}$$

5.2-12 **(a)** $w_1 = 2x_1$ and $\dfrac{dw_1}{dx_1} = 2$. Thus

$$f(x_1) = \frac{2}{\pi(1+4x_1^2)}\,, \qquad -\infty < x_1 < \infty.$$

(b) For $x_2 = y_1 - y_2,\ x_1 = y_2,\ |\,J\,| = 1$. Thus

$$g(y_1, y_2) = f(y_2)f(y_1-y_2), \qquad -\infty < y_i < \infty, \quad i = 1, 2.$$

(c) $g_1(y_1) = \displaystyle\int_{-\infty}^{\infty} f(y_2)f(y_1-y_2)\,dy_2.$

(d)
$$\begin{aligned} g_1(y_1) &= \int_{-\infty}^{\infty} \frac{2}{\pi[1+4y_2^2]}\cdot\frac{2}{\pi[1+4(y_1-y_2)^2]}\,dy_2 \;=\; \int_{-\infty}^{\infty} h(y_2)\,dy_2 \\ &= \frac{4}{\pi^2}\int_{-\infty}^{\infty} \frac{1}{[1+2iy_2][1-2iy_2]}\cdot\frac{1}{[1+2i(y_1-y_2)][1-2i(y_1-y_2)]}\,dy_2 \\ &= \frac{4}{\pi^2}\int_{-\infty}^{\infty} \frac{1}{2i}\cdot\frac{1}{y_2-\frac{i}{2}}\cdot\frac{-1}{2i}\cdot\frac{1}{y_2+\frac{i}{2}}\cdot\frac{-1}{2i}\cdot\frac{-1}{y_2-(y_1-\frac{i}{2})}\cdot\frac{1}{2i}\cdot\frac{1}{y_2-(y_1+\frac{i}{2})}dy_2 \\ &= \frac{4(2\pi i)}{\pi^2}\left[\mathrm{Res}\left(h(y_2); y_2 = \frac{i}{2}\right) + \mathrm{Res}\left(h(y_2); y_2 = y_1 + \frac{i}{2}\right)\right] \\ &= \frac{8\pi i}{\pi^2}\,\frac{1}{16}\left[\frac{1}{i}\cdot\frac{1}{i-y_1}\cdot\frac{1}{-y_1} + \frac{1}{y_1}\cdot\frac{1}{y_1+i}\cdot\frac{1}{i}\right] \\ &= \frac{1}{2\pi}\cdot\frac{1}{y_1}\left[\frac{1}{y_1-i}+\frac{1}{y_1+i}\right] = \frac{1}{2\pi}\cdot\frac{1}{y_1}\left[\frac{y_1+i+y_1-i}{(y_1-i)(y_1+i)}\right] \\ &= \frac{1}{\pi(1+y_1^2)}. \end{aligned}$$

A *Maple* solution for Exercise 5.2-12:

```
>f := x-> 2/Pi/(1 + 4*x^2);
```

$$f := x- > 2\,\frac{1}{\pi\,(1+4\,x^2)}$$

```
>simplify(int(f(y[2])*f(y[1]-y[2]),y[2]=-infinity..infinity));
```

$$\frac{1}{\pi\,(1+y_1^2)}$$

A *Mathematica* solution for Exercise 5.2-12:

```
In[1]:=
      f[x_] := 2/(Pi*(1 + 4(x)^2))
      g[y1_,y2_] := f[y2]*f[y1-y2]
In[3]:=
      Integrate[g[y1,y2], {y2, -Infinity,Infinity}]
Out[3]=
             1
        -------------
                  2
         Pi + Pi y1
```

5.2-14 The joint pdf is

$$h(x,y) = \frac{x}{5^3}\,e^{-(x+y)/5}, \qquad 0 < x < \infty, \quad 0 < y < \infty;$$

$$\begin{aligned} z &= \frac{x}{y}, & w &= y \\ x &= zw, & y &= w \end{aligned}$$

The Jacobian is

$$J = \begin{vmatrix} w & z \\ 0 & 1 \end{vmatrix} = w.$$

The joint pdf of Z and W is

$$f(z,w) = \frac{zw}{5^3}\,e^{-(z+1)w/5}\,w, \qquad 0 < z < \infty, \quad 0 < w < \infty;$$

The marginal pdf of Z is

$$\begin{aligned} f_Z(z) &= \int_0^\infty \frac{zw}{5^3}\,e^{-(z+1)w/5}\,w\,dw \\ &= \frac{\Gamma(3)z}{5^3}\left(\frac{5}{z+1}\right)^3 \int_0^\infty \frac{w^{3-1}}{\Gamma(3)(5/[z+1])^3}\,e^{-w/(5/[z+1])}\,dw \\ &= \frac{2z}{(z+1)^3}, \qquad 0 < z < \infty. \end{aligned}$$

5.2-16 The pdf of W is

$$f(w) = \frac{(r_1/r_2)^{r_1/2}\Gamma[(r_1+r_2)/2]w^{r_1/2-1}}{\Gamma(r_1/2)\Gamma(r_2/2)[1+(r_1 w/r_2)]^{(r_1+r_2)/2}}.$$

Let $\alpha = r_1/2$ and $\beta = r_2/2$. Then the pdf becomes

$$f(w) = \frac{(\alpha/\beta)^{\alpha}\Gamma[(\alpha+\beta]w^{\alpha-1}}{\Gamma(\alpha)\Gamma(\beta)[1+\alpha w/\beta]^{(\alpha+\beta)}}.$$

We are given

$$\begin{aligned} z &= \frac{1}{1+\alpha w/\beta} \\ w &= \frac{\beta}{\alpha}\left(\frac{1}{z}-1\right) = \frac{\beta}{\alpha}\left(\frac{1-z}{z}\right) \\ \frac{dw}{dz} &= -\frac{\beta}{\alpha}\frac{1}{z^2} \end{aligned}$$

It follows that the pdf of Z is

$$\begin{aligned} g(z) &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\left(\frac{\alpha}{\beta}\right)^{\alpha}\left[\frac{\beta}{\alpha}\left(\frac{1-z}{z}\right)\right]^{\alpha-1} z^{\alpha+\beta}\,\frac{\beta}{\alpha}\left(\frac{1}{z^2}\right) \\ &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}(1-z)^{\alpha-1}z^{\beta-1}, \qquad 0 < z < 1. \end{aligned}$$

5.3 Several Random Variables

5.3–2 **(a)**

$$\begin{aligned} P(X_1 = 2, X_2 = 4) &= \left[\frac{3!}{2!1!}\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^1\right]\left[\frac{5!}{4!1!}\left(\frac{1}{2}\right)^4\left(\frac{1}{2}\right)^1\right] \\ &= \frac{15}{2^8} = \frac{15}{256}. \end{aligned}$$

(b) $\{X_1 + X_2 = 7\}$ can occur in the two mutually exclusive ways: $\{X_1 = 3, X_2 = 4\}$ and $\{X_1 = 2, X_2 = 5\}$. The sum of the probabilities of the two latter events is

$$\left[\frac{3!}{3!0!}\left(\frac{1}{2}\right)^3\right]\left[\frac{5!}{4!1!}\left(\frac{1}{2}\right)^5\right] + \left[\frac{3!}{2!1!}\left(\frac{1}{2}\right)^3\right]\left[\frac{5!}{5!0!}\left(\frac{1}{2}\right)^5\right] = \frac{5+3}{2^8} = \frac{1}{32}.$$

5.3–4 **(a)**

$$\begin{aligned} \left(\int_{0.5}^{1.0} 2e^{-2x_1}\,dx_1\right)\left(\int_{0.7}^{1.2} 2e^{-2x_2}\,dx_2\right) &= (e^{-1}-e^{-2})(e^{-1.4}-e^{-2.4}) \\ &= (0.368-0.135)(0.247-0.091) \\ &= (0.233)(0.156) = 0.036. \end{aligned}$$

(b) $E(X_1) = E(X_2) = 0.5$,

$E[X_1(X_2-0.5)^2] = E(X_1)\text{Var}(X_2) = (0.5)(0.25) = 0.125.$

5.3–6
$$E(X) = \int_0^1 x6x(1-x)\,dx = \int_0^1 (6x^2-6x^3)\,dx = \left[2x^3 - \left(\frac{3}{2}\right)x^4\right]_0^1 = \frac{1}{2};$$
$$E(X^2) = \int_0^1 (6x^3-6x^4)\,dx = \left[\left(\frac{3}{2}\right)x^4 - \left(\frac{6}{5}\right)x^5\right]_0^1 = \frac{3}{10}.$$

Thus
$$\mu_X = \frac{1}{2};\quad \sigma_X^2 = \frac{3}{10} - \frac{1}{4} = \frac{1}{20},\text{ and}$$
$$\mu_Y = \frac{1}{2} + \frac{1}{2} = 1;\quad \sigma_Y^2 = \frac{1}{20} + \frac{1}{20} = \frac{1}{10}.$$

5.3–8 Let $Y = \max(X_1, X_2)$. Then
$$\begin{aligned}
G(y) &= [P(X \le y)]^2 \\
&= \left[\int_1^y \frac{4}{x^5}\,dx\right]^2 \\
&= \left[1 - \frac{1}{y^4}\right]^2, \qquad 1 < y < \infty \\
g(y) &= G'(y) \\
&= 2\left(1-\frac{1}{y^4}\right)\left(\frac{4}{y^5}\right), \qquad 1 < y < \infty; \\
E(Y) &= \int_1^\infty y \cdot 2\left(1-\frac{1}{y^4}\right)\left(\frac{4}{y^5}\right)dy \\
&= \int_1^\infty 8\,[y^{-4} - y^{-8}]\,dy \\
&= \frac{32}{21} \quad \text{(in 1000s of dollars).}
\end{aligned}$$

5.3–10 **(a)** $P(X_1 = 1)P(X_2 = 3)P(X_3 = 1) = \left(\frac{3}{4}\right)\left[\left(\frac{3}{4}\right)\left(\frac{1}{4}\right)^2\right]\left(\frac{3}{4}\right) = \frac{27}{1024};$

(b) $3P(X_1 = 3, X_2 = 1, X_3 = 1) + 3P(X_1 = 2, X_2 = 2, X_3 = 1) =$
$$3\left(\frac{27}{1024}\right) + 3\left(\frac{27}{1024}\right) = \frac{162}{1024};$$

(c) $P(Y \le 2) = \left(\frac{3}{4} + \frac{3}{4}\cdot\frac{1}{4}\right)^3 = \left(\frac{15}{16}\right)^3.$

5.3–12 $P(1 < \min X_i) = [P(1 < X_i)]^3 = \left(\int_1^\infty e^{-x}\,dx\right)^3 = e^{-3} = 0.05.$

5.3–14
$$\begin{aligned}
P(Y > 1000) &= P(X_1 > 1000)P(X_2 > 1000)P(X_3 > 1000) \\
&= e^{-1}e^{-2/3}e^{-1/2} \\
&= e^{-13/6} = 0.1146.
\end{aligned}$$

5.3–16
$$\begin{aligned} G(y) = P(Y \le y) &= P(X_1 \le y) \cdots P(X_8 \le y) = [P(X \le y)]^8 \\ &= [y^{10}]^8 = y^{80}, \qquad 0 < y < 1; \\ g(y) = G'(y) &= 80y^{79}, \qquad 0 < y < 1; \\ P(0.9999 < Y < 1) &= G(1) - G(0.9999) = 1 - 0.9999^{80} = 0.008. \end{aligned}$$

5.3–18 Denote the three lifetimes by X_1, X_2, X_3 and let $Y = X_1 + X_2 + X_3$.

$$E(Y) = E(X_1 + X_2 + X_3) = E(X_1) + E(X_2) + E(X_3) = 3 \cdot 2 \cdot 2 = 12.$$

$$\text{Var}(X_1 + X_2 + X_3) = \text{Var}(X_1) + \text{Var}(X_2) + \text{Var}(X_3) = 3 \cdot 2 \cdot 2^2 = 24.$$

5.3–20
$$\begin{aligned} \rho &= \frac{\text{Cov}(W, V)}{\sigma_W \sigma_V} \\ &= \frac{E(WV) - \mu_W \mu_V}{\sigma_W \sigma_V} \\ &= \frac{E(X^2)E(Y) - E(X)E(Y)E(X)}{\sigma_{XY}\sigma_X} \\ &= \frac{(\sigma_X^2 + \mu_X^2)(\mu_Y) - \mu_X^2 \mu_Y}{\sqrt{E(X^2Y^2) - [E(X)E(Y)]^2}\,\sigma_X} \\ &= \frac{\mu_Y \sigma_X}{\sqrt{(\sigma_X^2 + \mu_X^2)(\sigma_Y^2 + \mu_Y^2) - \mu_X^2 \mu_Y^2}} \\ &= \frac{\mu_Y \sigma_X}{\sqrt{\sigma_X^2 \sigma_Y^2 + \sigma_X^2 \mu_Y^2 + \sigma_Y^2 \mu_X^2}}. \end{aligned}$$

5.4 The Moment-Generating Function Technique

5.4–2
$$\begin{aligned} M_Y(t) &= E[e^{t(X_1+X_2)}] = E[e^{tX_1}]E[e^{tX_2}] \\ &= (q + pe^t)^{n_1}(q + pe^t)^{n_2} = (q + pe^t)^{n_1+n_2}. \end{aligned}$$

Thus Y is $b(n_1 + n_2,\, p)$.

5.4–4
$$\begin{aligned} E[e^{t(X_1+\cdots+X_n)}] &= \prod_{i=1}^{n} E[e^{tX_i}] = \prod_{i=1}^{n} e^{\mu_i(e^t-1)} \\ &= e^{(\mu_1+\mu_2+\cdots+\mu_n)(e^t-1)}, \end{aligned}$$

the moment generating function of a Poisson random variable with mean $\mu_1 + \mu_2 + \cdots + \mu_n$.

5.4–6 **(a)**
$$\begin{aligned} E[e^{tY}] &= E[e^{t(X_1+X_2+X_3+X_4+X_5)}] \\ &= E[e^{tX_1}e^{tX_2}e^{tX_3}e^{tX_4}e^{tX_5}] \\ &= E[e^{tX_1}]E[e^{tX_2}]E[e^{tX_3}]E[e^{tX_4}]E[e^{tX_5}] \\ &= \frac{(1/3)e^t}{1-(2/3)e^t}\,\frac{(1/3)e^t}{1-(2/3)e^t} \cdots \frac{(1/3)e^t}{1-(2/3)e^t} \\ &= \left[\frac{(1/3)e^t}{1-(2/3)e^t}\right]^5 \\ &= \frac{[(1/3)e^t]^5}{[1-(2/3)e^t]^5}, \qquad t < -\ln(1 - 1/3). \end{aligned}$$

(b) So Y has a negative binomial distribution with $p = 1/3$ and $r = 5$.

5.4–8
$$\begin{aligned} E[e^{tW}] &= E[e^{t(X_1+X_2+\cdots+X_h)}] = E[e^{tX_1}]E[e^{tX_2}]\cdots E[e^{tX_h}] \\ &= [1/(1-\theta t)]^h = 1/(1-\theta t)^h, \qquad t < 1/\theta, \end{aligned}$$

the moment generating function for the gamma distribution with mean $h\theta$.

5.4–10 **(a)** $E[e^{tX}] = (1/4)(e^{0t} + e^{1t} + e^{2t} + e^{3t})$;

(b) $E[e^{tY}] = (1/4)(e^{0t} + e^{4t} + e^{8t} + e^{12t})$;

(c)
$$\begin{aligned} E[e^{tW}] &= E[e^{t(X+Y)}] \\ &= E[e^{tX}]E[e^{tY}] \\ &= (1/16)(e^{0t} + e^{1t} + e^{2t} + e^{3t})(e^{0t} + e^{4t} + e^{8t} + e^{12t}) \\ &= (1/16)(e^{0t} + e^{1t} + e^{2t} + e^{3t} + \cdots e^{15t}); \end{aligned}$$

(d) $P(W = x) = 1/16, \qquad w = 0, 1, 2, \ldots, 15.$

5.4–12 First die has 0 on four faces and 2 on four faces; second die has faces numbered 0, 1, 4, 5, 8, 9, 12, 13.

5.4–14 Let X_1, X_2, X_3 be the number of accidents in weeks 1, 2, and 3, respectively. Then $Y = X_1 + X_2 + X_3$ is Poisson with mean $\lambda = 6$ and

$$P(Y = 7) = 0.744 - 0.606 = 0.138.$$

5.4–16 Let X_1, X_2, X_3, X_4 be the number of sick days for employee i, $i = 1, 2, 3, 4$, respectively. Then $Y = X_1 + X_2 + X_3 + X_4$ is Poisson with mean $\lambda = 8$ and

$$P(Y > 10) = 1 - P(Y \le 10) = 1 - 0.0816 = 0.184.$$

5.4–18 Let X_i equal the number of cracks in mile i, $i = 1, 2, \ldots, 40$. Then

$$Y = \sum_{i=1}^{40} X_i \qquad \text{is Poisson with mean} \qquad \lambda = 20.$$

It follows that

$$P(Y < 15) = P(Y \le 14) = \sum_{y=0}^{14} \frac{20^y e^{-20}}{y!} = 0.1049.$$

5.4–20 $Y = X_1 + X_2 + X_3 + X_4$ has a gamma distribution with $\alpha = 6$ and $\theta = 10$. So

$$P(Y > 90) = \int_{90}^{\infty} \frac{1}{\Gamma(6)10^6} y^{6-1} e^{-y/10}\, dy = 1 - 0.8843 = 0.1157.$$

5.4–22 **(a)**
$$\begin{aligned} E(e^{tY}) &= E(e^{tX_1}e^{tX_2}) \\ &= E(e^{tX_1})E(e^{tX_2}) \\ \frac{1}{(1-2t)^{r/2}} &= \frac{1}{(1-2t)^{r_1/2}}\, E(e^{tX_2}) \\ E(e^{tX_2}) &= \frac{1}{(1-2t)^{(r-r_1)/2}} \end{aligned}$$

(b) X_2 is $\chi^2(r - r_1)$.

5.5 Random Functions Associated with Normal Distributions

5.5–2

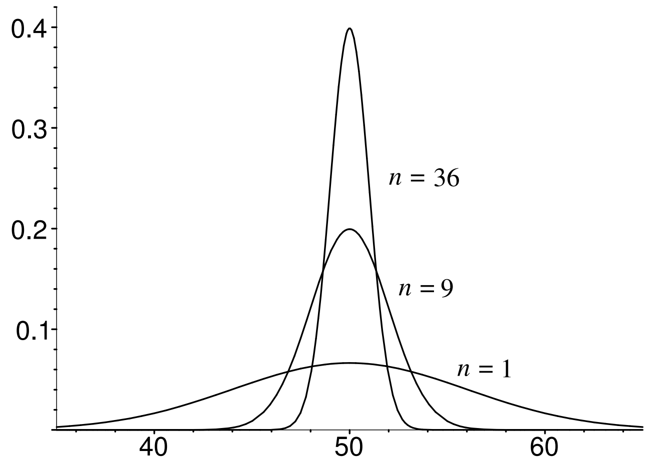


Figure 5.5–2: X is $N(50,\ 36)$, $\overline{X}$ is $N(50,\ 36/n),\ n = 9, 36$

5.5–4 **(a)** $P(X < 6.0171) = P(Z < -1.645) = 0.05$;

(b) Let W equal the number of boxes that weigh less than 6.0171 pounds. Then W is $b(9,\ 0.05)$ and $P(W \le 2) = 0.9916$;

(c)
$$\begin{aligned} P(\overline{X} \le 6.035) &= P\left(Z \le \frac{6.035 - 6.05}{0.02/3}\right) \\ &= P(Z \le -2.25) = 0.0122. \end{aligned}$$

5.5–6 **(a)**

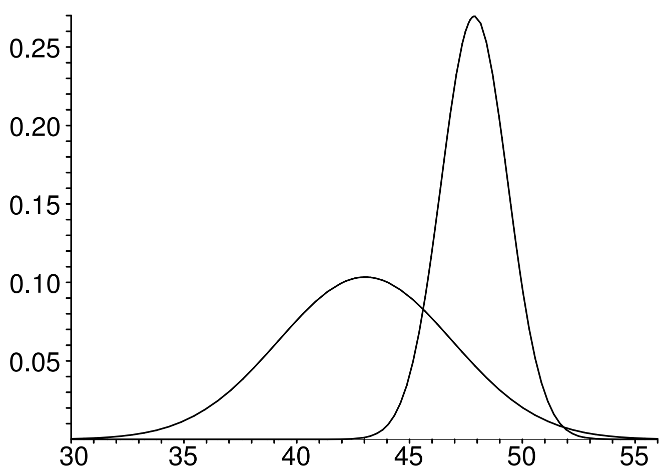


Figure 5.5–6: $N(43.04,\ 14.89)$ and $N(47.88,\ 2.19)$ pdfs

(b) The distribution of $X_1 - X_2$ is $N(4.84,\ 17.08)$. Thus

$$P(X_1 > X_2) = P(X_1 - X_2 > 0) = P\left(Z > \frac{-4.84}{\sqrt{17.08}}\right) = 0.8790.$$

5.5–8 $X - Y$ is $N(184.09 - 171.93, 39.37 + 50.88)$;

$$P(X > Y) = P\left(\frac{X - Y - 12.16}{\sqrt{90.25}} > \frac{0 - 12.16}{9.5}\right) = P(Z > -1.28) = 0.8997.$$

5.5–10 Let $Y = X_1 + X_2 + \cdots + X_n$. Then Y is $N(800n, 100^2 n)$. Thus

$$\begin{aligned}
P(Y \geq 10000) &= 0.90 \\
P\left(\frac{Y - 800n}{100\sqrt{n}} \geq \frac{10000 - 800n}{100\sqrt{n}}\right) &= 0.90 \\
-1.282 &= \frac{10000 - 800n}{100\sqrt{n}} \\
800n - 128.2\sqrt{n} - 10000 &= 0.
\end{aligned}$$

Either use the quadratic formula to solve for $\sqrt{n}$ or use Maple to solve for n. We find that $\sqrt{n} = 3.617$ or $n = 13.08$ so use $n = 14$ bulbs.

5.5–12 **(a)** The joint pdf of X_1 and X_2 is

$$f(x_1, x_2) = \frac{1}{\sqrt{2\pi}} e^{-x_1^2/2} \frac{1}{\Gamma(r/2)2^{r/2}}\, x_2^{r/2-1} e^{-x_2/2}, \qquad -\infty < x_1 < \infty, \quad 0 < x_2 < \infty;$$

$$\begin{aligned}
y_1 &= x_1/\sqrt{x_2/r}, & y_2 &= x_2 \\
x_1 &= y_1\sqrt{y_2/r}, & x_2 &= y_2
\end{aligned}$$

The Jacobian is

$$J = \begin{vmatrix} \sqrt{y_2/r} & y_1(\frac{1}{2})y_2^{-1/2}/\sqrt{r} \\ 0 & 1 \end{vmatrix} = \sqrt{y_2/r}.$$

The joint pdf of Y_1 and Y_2 is

$$g(y_1, y_2) = \frac{1}{\sqrt{2\pi}} e^{-y_1^2 y_2/2r} \frac{1}{\Gamma(r/2)2^{r/2}}\, y_2^{r/2-1} e^{-y_2/2} \frac{\sqrt{y_2}}{\sqrt{r}}, \qquad -\infty < y_1 < \infty, \quad 0 < y_2 < \infty.$$

(b) The marginal pdf of Y_1 is

$$\begin{aligned}
g_1(y_1) &= \int_0^\infty \frac{1}{\sqrt{2\pi}}\, e^{-y_1^2 y_2/2r}\, \frac{1}{\Gamma(r/2)2^{r/2}} y_2^{r/2-1}\, e^{-y_2/2}\, \frac{\sqrt{y_2}}{\sqrt{r}}\, dy_2 \\
&= \frac{\Gamma[(r+1)/2]}{\sqrt{\pi r}\,\Gamma(r/2)} \int_0^\infty \frac{1}{\Gamma[(r+1)/2]2^{(r+1)/2}}\, y_2^{(r+1)/2-1}\, e^{-(y_2/2)(1+y_1^2/r)}.
\end{aligned}$$

Let $u = y_2(1 + y_1^2/r)$. Then $y_2 = \dfrac{u}{1 + y_1^2/r}$ and $\dfrac{dy_2}{du} = \dfrac{1}{1 + y_1^2/r}$. So

$$\begin{aligned}
g_1(y_1) &= \frac{\Gamma[(r+1)/2]}{\sqrt{\pi r}\,\Gamma(r/2)(1 + y_1^2/r)^{(r+1)/2}} \int_0^\infty \frac{1}{\Gamma[(r+1)/2]2^{(r+1)/2}}\, u^{(r+1)/2-1}\, e^{-u/2} \\
&= \frac{\Gamma[(r+1)/2]}{\sqrt{\pi r}\,\Gamma(r/2)(1 + y_1^2/r)^{(r+1)/2}}, \qquad -\infty < y_1 < \infty.
\end{aligned}$$

5.5–14 Because Z is $N(0,1)$, $E(Z)=1$ and $E(Z^2)=1$. U is $\chi^2(r)$ so it follows that

$$\begin{aligned}
E[1/\sqrt{U}] &= \int_0^\infty \frac{1}{\sqrt{u}}\frac{1}{\Gamma(r/2)2^{r/2}}\,u^{r/2-1}\,e^{-u/2}\,du \\
&= \frac{\Gamma[(r-1)/2]}{\sqrt{2}\,\Gamma(r/2)}\int_0^\infty \frac{1}{\Gamma[(r-1)/2]2^{(r-1)/2}}\,u^{(r-1)/2-1}e^{-u/2}\,du \\
&= \frac{\Gamma[(r-1)/2]}{\sqrt{2}\,\Gamma(r/2)}.
\end{aligned}$$

Note that the last integral is equal to one because the integrand is the pdf of a $\chi^2(r-1)$ random variable.

To find $E(1/U)$ we have

$$\begin{aligned}
E[1/U] &= \int_0^\infty \frac{1}{u}\frac{1}{\Gamma(r/2)2^{r/2}}\,u^{r/2-1}\,e^{-u/2}\,du \\
&= \frac{\Gamma[(r-2)/2]}{2\,\Gamma(r/2)}\int_0^\infty \frac{1}{\Gamma[(r-2)/2]2^{(r-2)/2}}\,u^{(r-2)/2-1}\,e^{-u/2}\,du \\
&= \frac{\Gamma(r/2-1)}{2\Gamma(r/2-1)\,(r/2-1)} \;=\; \frac{1}{r-2}.
\end{aligned}$$

Note that the last integral is equal to one because the integrand is the pdf of a $\chi^2(r-2)$ random variable.

$$\begin{aligned}
E(T) &= E\left[\frac{Z}{\sqrt{U/r}}\right] \\
&= E[Z]E[1/\sqrt{U/r}] \\
&= 0\left[\frac{\sqrt{r}\Gamma[(r-1)/2]}{\sqrt{2}\,\Gamma(r/2)}\right] \;=\; 0 \qquad \text{provided } r \ge 2; \\
\mathrm{Var}(T) &= E(T^2) - 0^2 \\
&= E[Z^2]E[r/U] \\
&= \frac{r}{r-2}, \qquad \text{provided } r \ge 3.
\end{aligned}$$

5.5–16 $T = \dfrac{\overline{X}-\mu}{S/\sqrt{9}}$ is t with $r = 9-1 = 8$ degrees of freedom.

(a) $t_{0.025}(8) = 2.306$;

(b)

$$\begin{aligned}
-t_{0.025} \;&\le\; \frac{\overline{X}-\mu}{S/\sqrt{n}} \;\le\; t_{0.025} \\
-t_{0.025}\frac{S}{\sqrt{n}} \;&\le\; \overline{X}-\mu \;\le\; t_{0.025}\frac{S}{\sqrt{n}} \\
-\overline{X}-t_{0.025}\frac{S}{\sqrt{n}} \;&\le\; -\mu \;\le\; -\overline{X}+t_{0.025}\frac{S}{\sqrt{n}} \\
\overline{X}-t_{0.025}\frac{S}{\sqrt{n}} \;&\le\; \mu \;\le\; \overline{X}+t_{0.025}\frac{S}{\sqrt{n}}
\end{aligned}$$

5.6 The Central Limit Theorem

5.6–2 If

$$\begin{aligned} f(x) &= (3/2)x^2, \qquad -1 < x < 1, \\ E(X) &= \int_{-1}^{1} x(3/2)x^2\,dx \;=\; 0; \\ \mathrm{Var}(X) &= \int_{-1}^{1} (3/2)x^4\,dx \;=\; \left[\frac{3}{10}x^5\right]_{-1}^{1} \;=\; \frac{3}{5}. \end{aligned}$$

Thus

$$\begin{aligned} P(-0.3 \le Y \le 1.5) &= P\left(\frac{-0.3-0}{\sqrt{15(3/5)}} \le \frac{Y-0}{\sqrt{15(3/5)}} \le \frac{1.5-0}{\sqrt{15(3/5)}}\right) \\ &\approx P(-0.10 \le Z \le 0.50) \;=\; 0.2313. \end{aligned}$$

5.6–4

$$\begin{aligned} P(39.75 \le \overline{X} \le 41.25) &= P\left(\frac{39.75-40}{\sqrt{(8/32)}} \le \frac{\overline{X}-40}{\sqrt{(8/32)}} \le \frac{41.25-40}{\sqrt{(8/32)}}\right) \\ &\approx P(-0.50 \le Z \le 2.50) \;=\; 0.6853. \end{aligned}$$

5.6–6 **(a)**

$$\begin{aligned} \mu &= \int_0^2 x(1-x/2)\;dx = \left[\frac{x^2}{2} - \frac{x^3}{6}\right]_0^2 \;=\; 2 - \frac{4}{3} \;=\; \frac{2}{3}; \\ \sigma^2 &= \int_0^2 x^2(1-x/2)\,dx - \left(\frac{2}{3}\right)^2 \\ &= \left[\frac{x^3}{3} - \frac{x^4}{8}\right]_0^2 - \frac{4}{9} \;=\; \frac{2}{9}. \end{aligned}$$

(b)

$$\begin{aligned} P\left(\frac{2}{3} \le \overline{X} \le \frac{5}{6}\right) &= P\left(\frac{\frac{2}{3}-\frac{2}{3}}{\sqrt{\frac{2}{9}/18}} \le \frac{\overline{X}-\frac{2}{3}}{\sqrt{\frac{2}{9}/18}} \le \frac{\frac{5}{6}-\frac{2}{3}}{\sqrt{\frac{2}{9}/18}}\right) \\ &\approx P(0 \le Z \le 1.5) \;=\; 0.4332. \end{aligned}$$

5.6–8 **(a)** $E(\overline{X}) = \mu = 24.43;$

(b) $\mathrm{Var}(\overline{X}) = \dfrac{\sigma^2}{n} = \dfrac{2.20}{30} = 0.0733;$

(c)

$$\begin{aligned} P(24.17 \le \overline{X} \le 24.82) &\approx P\left(\frac{24.17-24.43}{\sqrt{0.0733}} \le Z \le \frac{24.82-24.43}{\sqrt{0.0733}}\right) \\ &= P(-0.96 \le Z < 1.44) \;=\; 0.7566. \end{aligned}$$

5.6–10

$$\begin{aligned} E(X+Y) &= 30 + 50 \;=\; 80; \\ \mathrm{Var}(X+Y) &= \sigma_X^2 + \sigma_Y^2 + 2\rho\sigma_X\sigma_Y \\ &= 52 + 64 + 28 \;=\; 144; \\ Z &= \sum_{i=1}^{25}(X_i + Y_i) \;\;\text{is approximately}\;\; N(25\cdot 80,\, 25\cdot 144). \end{aligned}$$

Thus

$$\begin{aligned} P(1970 < Z < 2090) &= P\left(\frac{1970-2000}{60} < \frac{Z-2000}{60} < \frac{2090-2000}{60}\right) \\ &\approx \Phi(1.5) - \Phi(-0.5) \\ &= 0.9332 - 0.3085 \;=\; 0.6247. \end{aligned}$$

5.6–12 Let X_i equal the time between sales of ticket $i-1$ and i, for $i = 1, 2, \ldots, 10$. Each X_i has a gamma distribution with $\alpha = 3$, $\theta = 2$. $Y = \sum_{i=1}^{10} X_i$ has a gamma distribution with parameters $\alpha_Y = 30, \theta_Y = 2$.

(a) Thus

$$P(Y \le 60) = \int_0^{60} \frac{1}{\Gamma(30)2^{30}} y^{30-1} e^{-y/2}\, dy \;=\; 0.52428 \;\text{ using Maple.}$$

(b) The normal approximation is given by

$$P\left(\frac{Y-60}{\sqrt{120}} \le \frac{60-60}{\sqrt{120}}\right) \;\approx\; \Phi(0) \;=\; 0.5000.$$

5.6–14 We are given that $Y = \sum_{i=1}^{20} X_i$ has mean 200 and variance 80. We want to find y so that

$$P(Y > y) = P\left(\frac{Y-200}{\sqrt{80}} > \frac{y-200}{\sqrt{80}}\right) < 0.20.$$

We have that

$$\begin{aligned} \frac{y-200}{\sqrt{80}} &= 0.842 \\ y &= 207.5 \uparrow 208 \text{ days.} \end{aligned}$$

5.7 Approximations for Discrete Distributions

5.7–2 **(a)** $P(2 < X < 9) = 0.9532 - 0.0982 = 0.8550$;

(b)
$$\begin{aligned} P(2 < X < 9) &= P\left(\frac{2.5-5}{2} \le \frac{X - 25(0.2)}{\sqrt{25(0.2)(0.8)}} \le \frac{8.5-5}{2}\right) \\ &\approx P(-1.25 \le Z \le 1.75) \\ &= 0.8543. \end{aligned}$$

5.7–4
$$\begin{aligned} P(35 \le X \le 40) &\approx P\left(\frac{34.5-36}{3} \le Z \le \frac{40.5-36}{3}\right) \\ &= P(-0.50 \le Z \le 1.50) \;=\; 0.6247. \end{aligned}$$

5.7–6 $\mu_X = 84(0.7) = 58.8$, $\mathrm{Var}(X) = 84(0.7)(0.3) = 17.64$,

$$P(X \le 52.5) \approx \Phi\left(\frac{52.5-58.8}{4.2}\right) = \Phi(-1.5) = 0.0668.$$

5.7–8 **(a)**
$$\begin{aligned} P(X < 20.857) &= P\left(\frac{X-21.37}{0.4} < \frac{20.857-21.37}{0.4}\right) \\ &= P(Z < -1.282) \;=\; 0.10. \end{aligned}$$

(b) The distribution of Y is $b(100, 0.10)$. Thus

$$P(Y \le 5) = P\left(\frac{Y - 100(0.10)}{\sqrt{100(0.10)(0.90)}} \le \frac{5.5-10}{3}\right) \approx P(Z \le -1.50) = 0.0668.$$

(c) $$\begin{aligned} P(21.31 \le \overline{X} \le 21.39) &\approx P\left(\frac{21.31-21.37}{0.4/10} \le Z \le \frac{21.39-21.37}{0.4/10}\right) \\ &= P(-1.50 \le Z \le 0.50) \;=\; 0.6247. \end{aligned}$$

5.7–10 The distribution of Y is $b(1000, 18/38)$. Thus

$$P(Y > 500) \approx P\left(Z \ge \frac{500.5 - 1000(18/38)}{\sqrt{1000(18/38)(20/38)}}\right) = P(Z \ge 1.698) = 0.0448.$$

5.7–12 **(a)** $E(X) = 100(0.1) = 10,\ \mathrm{Var}(X) = 9,$

$$\begin{aligned} P(11.5 < X < 14.5) &\approx \Phi\left(\frac{14.5-10}{3}\right) - \Phi\left(\frac{11.5-10}{3}\right) \\ &= \Phi(1.5) - \Phi(0.5) \;=\; 0.9332 - 0.6915 \;=\; 0.2417. \end{aligned}$$

(b) $P(X \le 14) - P(X \le 11) = 0.917 - 0.697 = 0.220;$

(c) $\displaystyle\sum_{x=12}^{14} \binom{100}{x} (0.1)^x (0.9)^{100-x} = 0.2244.$

5.7–14 **(a)** $E(Y) = 24(3.5) = 84,\ \mathrm{Var}(Y) = 24(35/12) = 70,$

$$P(Y \ge 85.5) \approx 1 - \Phi\left(\frac{85.5-84}{\sqrt{70}}\right) = 1 - \Phi(0.18) = 0.4286;$$

(b) $P(Y < 85.5) \approx 1 - 0.4286 = 0.5714;$

(c) $P(70.5 < Y < 86.5) \approx \Phi(0.30) - \Phi(-1.61) = 0.6179 - 0.0537 = 0.5642.$

5.7–16 **(a)** See Figure 5.7-16.

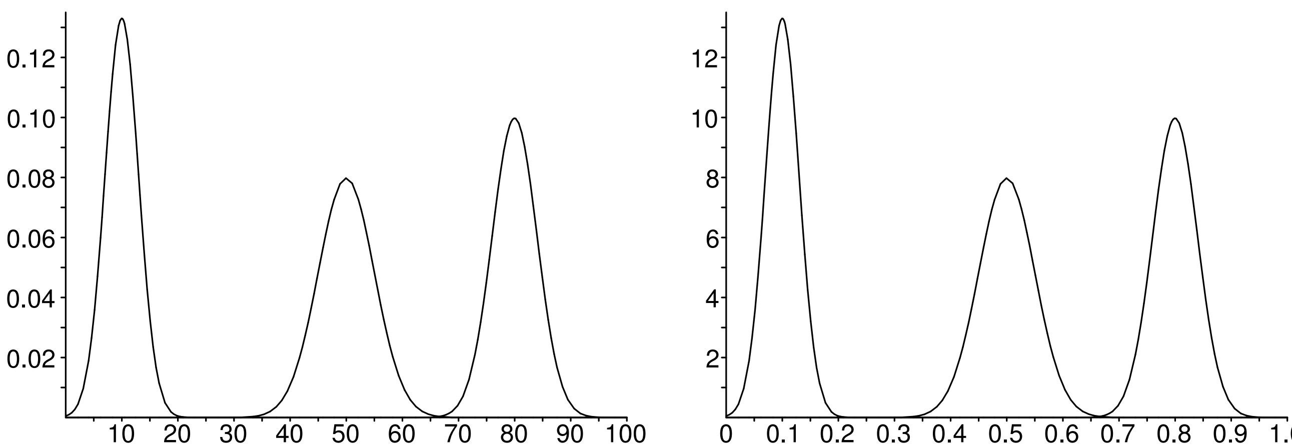


Figure 5.7–16: Normal approximations of the pdfs of Y and $Y/100$, $p = 0.1, 0.5, 0.8$

(b) **(i)** When $p = 0.1$,

$$P(-1.5 < Y - 10 < 1.5) \approx \Phi\left(\frac{1.5}{3}\right) - \Phi\left(\frac{-1.5}{3}\right) = 0.6915 - 0.3085 = 0.3830;$$

(ii) When $p = 0.5$,

$$P(-1.5 < Y - 50 < 1.5) \approx \Phi\left(\frac{1.5}{5}\right) - \Phi\left(\frac{-1.5}{5}\right) = 0.6179 - 0.3821 = 0.2358;$$

(iii) When $p = 0.8$,

$$P(-1.5 < Y - 80 < 1.5) \approx \Phi\left(\frac{1.5}{4}\right) - \Phi\left(\frac{-1.5}{4}\right) = 0.6462 - 0.3538 = 0.2924.$$

5.7–18 X is $N(0, 0.5^2)$. The probability that one item exceeds 0.98 in absolute value is

$$\begin{aligned} P(|X| > 0.98) &= 1 - P(-0.98 \le X \le 0.98) \\ &= 1 - P\left(\frac{-0.98 - 0}{0.5} \le \frac{X - 0}{0.5} \le \frac{0.98 - 0}{0.5}\right) \\ &= 1 - P(-1.96 \le Z \le 1.96) = 1 - 0.95 = 0.05 \end{aligned}$$

If we let Y equal the number out of 100 that exceed 0.98 in absolute value, Y is $b(100, 0.05)$.

(a) Let $\lambda = 100(0.05) = 5$.

$$P(Y \ge 7) = 1 - P(Y \le 6) = 1 - 0.762 = 0.238.$$

(b)
$$\begin{aligned} P(Y \ge 7) &= P\left(\frac{Y - 5}{\sqrt{100(0.05)(0.95)}} \ge \frac{6.5 - 5}{2.179}\right) \\ &\approx P(Z \ge 0.688) \\ &= 1 - 0.7543 = 0.2447. \end{aligned}$$

(c) $P(Y \ge 7) = 1 - P(Y \le 6) = 1 - 0.7660 = 0.2340.$

5.8 Chebyshev's Inequality and Convergence in Probability

5.8–2 $\text{Var}(X) = 298 - 17^2 = 9$.

(a)
$$\begin{aligned} P(10 < X < 24) &= P(10 - 17 < X - 17 < 24 - 17) \\ &= P(\,|X - 17| < 7) \ge 1 - \frac{9}{49} = \frac{40}{49}, \end{aligned}$$

because $k = 7/3$;

(b) $P(\,|X - 17| \ge 16) \le \dfrac{9}{16^2} = 0.035$, because $k = 16/3$.

5.8–4 **(a)** $P\left(\left|\dfrac{Y}{100} - 0.5\right| < 0.08\right) \ge 1 - \dfrac{(0.5)(0.5)}{100(0.08)^2} = 0.609;$

because $k = 0.08/\sqrt{(0.5)(0.5)/100}$;

(b) $P\left(\left|\dfrac{Y}{500} - 0.5\right| < 0.08\right) \ge 1 - \dfrac{(0.5)(0.5)}{500(0.08)^2} = 0.922;$

because $k = 0.08/\sqrt{(0.5)(0.5)/500}$;

(c) $P\left(\left|\dfrac{Y}{1000} - 0.5\right| < 0.08\right) \ge 1 - \dfrac{(0.5)(0.5)}{1000(0.08)^2} = 0.961,$

because $k = 0.08/\sqrt{(0.5)(0.5)/1000}$.

5.8–6
$$\begin{aligned} P(75 < \overline{X} < 85) &= P(75 - 80 < \overline{X} - 80 < 85 - 80) \\ &= P(\,|\overline{X} - 80| < 5) \ge 1 - \frac{60/15}{25} = 0.84, \end{aligned}$$

because $k = 5/\sqrt{60/15} = 5/2$.

5.9 Limiting Moment-Generating Functions

5.9–2 Using Table III with $\lambda = np = 400(0.005) = 2$, $P(X \le 2) = 0.677$.

5.9–4 Let $Y = \sum_{i=1}^{n} X_i$, where $X_1, X_2, \ldots, X_n$ are mutually independent $\chi^2(1)$ random variables. Then $\mu = E(X_i) = 1$ and $\sigma^2 = \mathrm{Var}(X_i) = 2$, $i = 1, 2, \ldots, n$. Hence

$$\frac{Y - n\mu}{\sqrt{n\sigma^2}} = \frac{Y - n}{\sqrt{2n}}$$

has a limiting distribution that is $N(0, 1)$.

Chapter 6

Point Estimation

6.1 Descriptive Statistics

6.1–2 $\overline{x} = 3.58$; $s = 0.5116$.

6.1–4 **(a)**

Class Interval	Class Limits	Frequency f_i	Class Mark, u_i
(303.5, 307.5)	(304, 307)	1	305.5
(307.5, 311.5)	(308, 311)	5	309.5
(311.5, 315.5)	(312, 315)	6	313.5
(315.5, 319.5)	(316, 319)	10	317.5
(319.5, 323.5)	(320, 323)	11	321.5
(323.5, 327.5)	(324, 327)	9	325.5
(327.5, 331.5)	(328, 331)	7	329.5
(331.5, 335.5)	(332, 335)	1	333.5

(b) $\overline{x} = 320.1$, $s = 6.7499$;

(c)

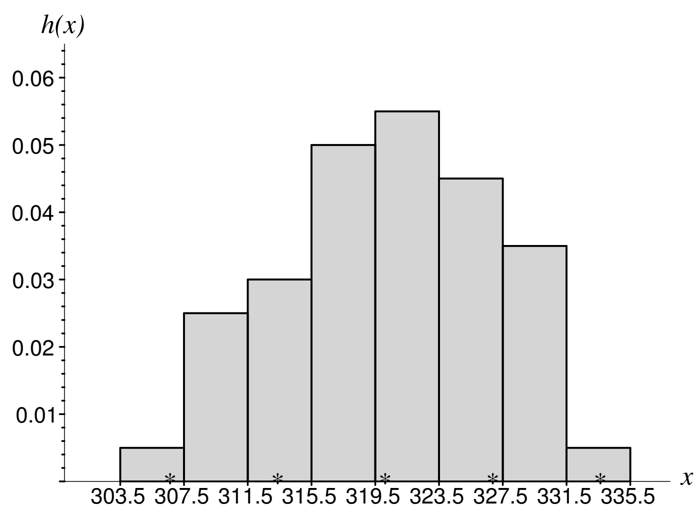


Figure 6.1–4: Melting points of metal alloys

There are 31 observations within one standard deviation of the mean (62%) and 48 observations within two standard deviations of the mean (96%).

6.1–6 **(a)** With the class boundaries $0.5, 5.5, 17.5, 38.5, 163.5, 549.5,$ the respective frequencies are $11, 9, 10, 10, 10$.

(b)

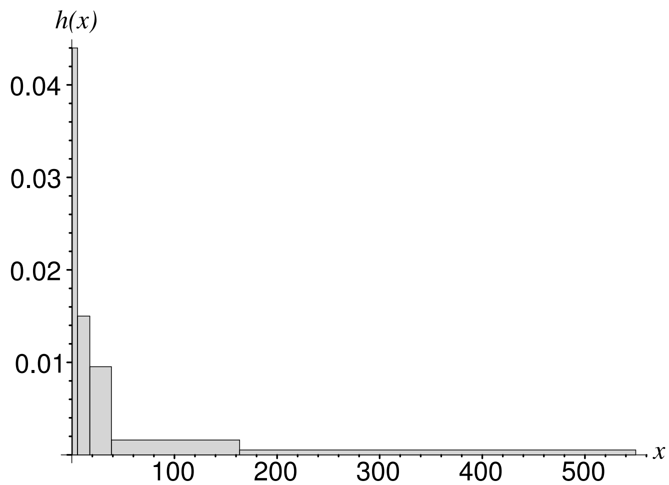


Figure 6.1–6: Mobil home losses

(c) This is a skewed to the right distribution.

6.1–8 **(a)** With the class boundaries $3.5005, 3.5505, 3.6005, \ldots, 4.1005$, the respective class frequencies are $4, 7, 24, 23, 7, 4, 3, 9, 15, 23, 18, 2$.

(b)

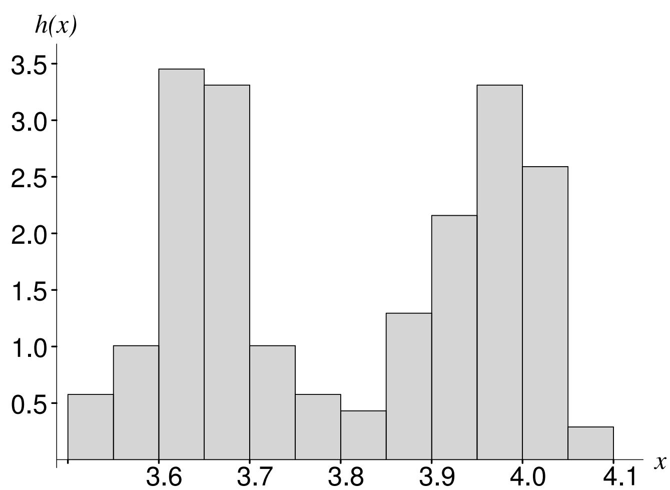


Figure 6.1–8: Weights of mirror parts

(c) This is a bimodal histogram.

6.1–10 **(a)** $50/204 = 0.245$; $93/329 = 0.283$;

(b) $124/355 = 0.349$; $21/58 = 0.362$;

(c) $174/559 = 0.311$; $114/387 = 0.295$.

(d) Although James' batting average is higher that Hrbek's on both grass and artificial turf, Hrbek's is higher over all. Note the different numbers of at bats on grass and artificial turf and how this affects the batting averages.

6.2 Exploratory Data Analysis

6.2–2 **(a)** Baby carrots: 1.02, 1.03, 1.04, 1.04, 1.06;

Regular-size carrots: 1.00, 1.15, 1.21, 1.26, 1.43;

(b)

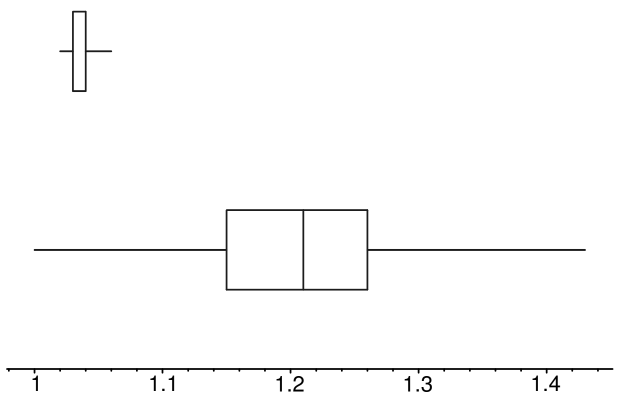


Figure 6.2–2: **(b)** Box-and-whisker diagrams of small and regular-size carrots

(c) Regular-size packages tend to be heavier.

6.2–4 **(a)** The five-number summary is: min $= 1$, $\widetilde{q}_1 = 6.75$, $\widetilde{m} = 32$, $\widetilde{q}_3 = 90.75$, max $= 527$.

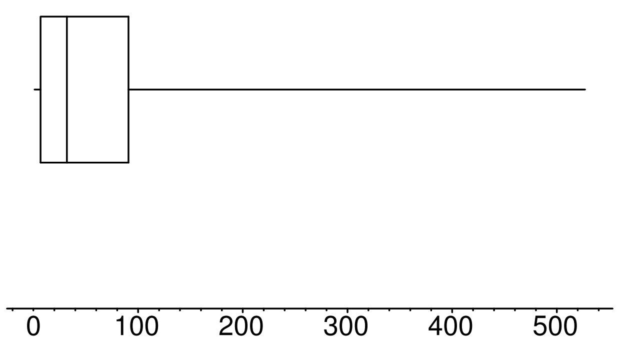


Figure 6.2–4: **(a)** Box-and-whisker diagram of mobile home losses

(b) IQR $= 90.75 - 6.75 = 84$. The inner fence is at 216.75 and the outer fence is at 342.75.

(c)

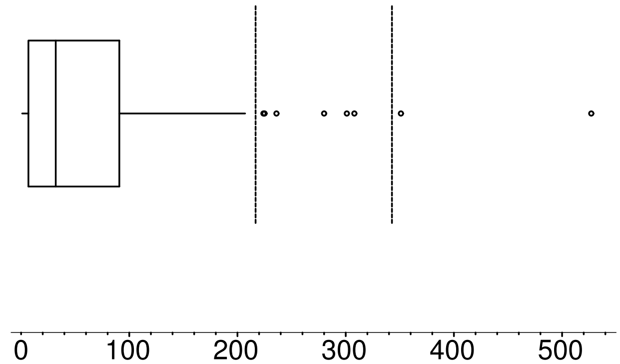


Figure 6.2–4: **(c)** Box-and-whisker diagram of losses with fences and outliers

6.2–6 **(a)**

Stems	Leaves	Freq
0•	5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 6 6 6 6 6 6 6 6 6 6 6 6 6 6 7 7 7 7 7 7 7 8 8 8 8 8 8 8 9 9 9 9 9 9	53
1∗	0 0 0 0 0 0 1 1 1 1 1 1 1 2 2 2 3 3 4	19
1•	5 5 5 5 6 6 6 6 7 7 8 8 9	13
2∗	0 1 1 1 1 3 3 4 4 4	10
2•	5	1
3∗	4	1
3•	5	1
4∗		0
4•	5	1
5∗		0
5•	5	1

(b) The five-number summary is: min $= 5$, $\widetilde{q}_1 = 6$, $\widetilde{m} = 9$, $\widetilde{q}_3 = 15$, max $= 55$.

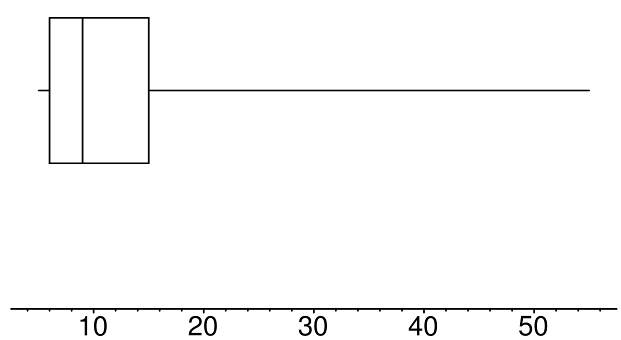


Figure 6.2–6: **(b)** Box-and-whisker diagram of maximum capital

(c) IQR $= 15 - 6 = 9$. The inner fence is at 28.5 and the outer fence is at 42.

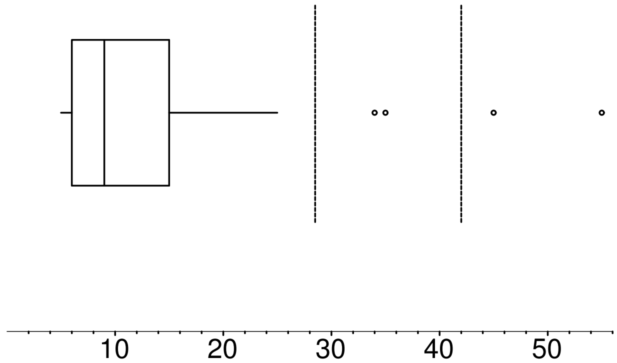


Figure 6.2–6: **(d)** Box-and-whisker diagram of maximum capital with outliers and fences

(e) The 90th percentile is 22.8.

6.2–8 **(a)**

Stems	Leaves	Frequency
101	7	1
102	0 0 0	3
103		0
104		0
105	8 9	2
106	1 3 3 6 6 7 7 8 8	9
107	3 7 9	3
108	8	1
109	1 3 9	3
110	0 2 2	3

(Multiply numbers by 10^{-1}.)

Table 6.2–8: Ordered stem-and-leaf diagram of weights of indicator housings

(b) min $= 101.7$, $\widetilde{q}_1 = 106.0$, $\widetilde{m} = 106.7$, $\widetilde{q}_3 = 108.95$, max $= 110.2$;

(c) The interquartile range in IQR $= 108.95 - 106.0 = 2.95$. The inner fence is located at $106.7 - 1.5(2.95) = 102.275$ so there are four suspected outliers.

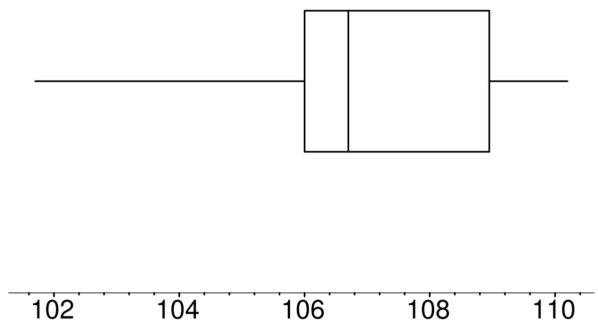


Figure 6.2–8: **(b)** Weights of indicator housings

6.2–10 **(a)** With the class boundaries $2.85, 3.85, \ldots, 16.85$ the respective frequencies are 1, 0, 2, 4, 1, 14, 20, 11, 4, 5, 0, 1, 0, 1.

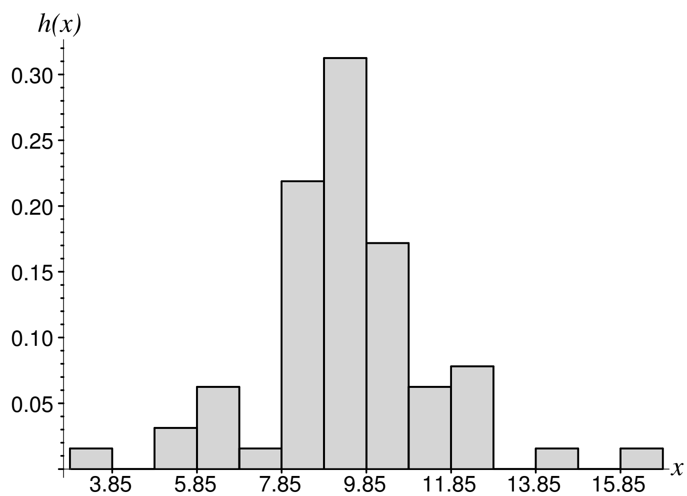


Figure 6.2–10: **(a)** Lead concentrations

(b) $\overline{x} = 9.422$, $s = 2.082$.

(c)

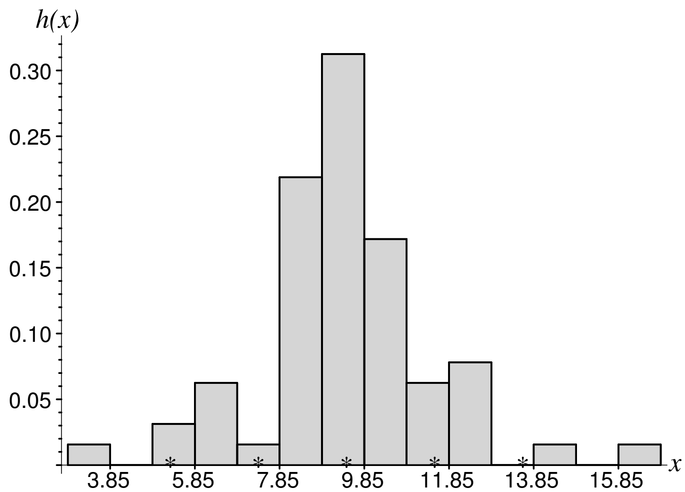


Figure 6.2–10: **(c)** Lead concentrations showing $\overline{x}, \overline{x} \pm s, \overline{x} \pm 2s$

There are 44 ($44/64 = 68.75\%$) within one standard deviation of the mean and 56 ($56/64 = 87.5\%$) within two standard deviations of the mean.

(d)

1976 Leaves	Stems	1977 Leaves
1	2	9
9 2	3	
9	4	
9 9 4 3 2 0 0	5	0 7
9 8 8 7 5 5 4 4 4 4 3 2 2 1 1 0 0 0 0 0	6	3 5 6 8
9 8 6 6 3 2 2 1 0	7	3
7 6 6 5 5 4 3 3 1 1 0 0	8	0 1 1 2 2 2 3 6 7 7 7 8 8 8 9 9
9 7 5 3 2 0	9	1 1 2 3 3 3 3 4 4 4 4 5 5 6 7 8 8 8 9 9 9 9
9 6 1	10	2 2 3 4 5 5 7 9
2	11	0 4 6 9
	12	0 3 4 6
	13	
1	14	8
	15	
	17	7

Multiply numbers by 10^{-1}

Table 6.2–10: **(d)** Back-to-Back Stem-and-Leaf Diagram of Lead Concentrations

(e)

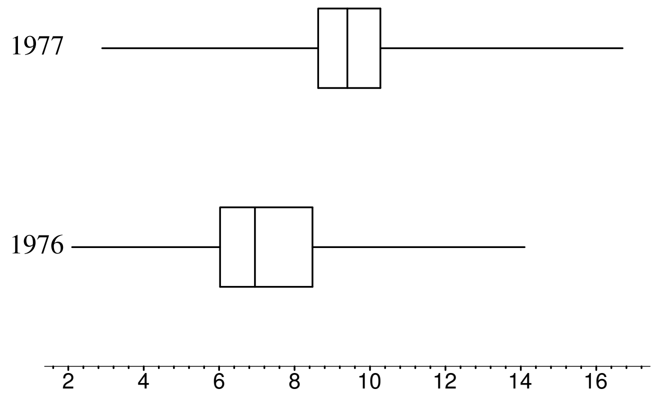


Figure 6.2–10: **(e)** Box-and-whisker diagrams of 1976 and 1977 lead concentrations

6.3 Order Statistics

6.3–2 **(a)** The location of the median is $(0.5)(17+1) = 9$, thus the median is

$$\widetilde{m} = 5.2.$$

The location of the first quartile is $(0.25)(17+1) = 4.5$. Thus the first quartile is

$$\widetilde{q}_1 = (0.5)(4.3) + (0.5)(4.7) = 4.5.$$

The location of the third quartile is $(0.75)(17+1) = 13.5$. Thus the third quartile is

$$\widetilde{q}_3 = (0.5)(5.6) + (0.5)(5.7) = 5.65.$$

(b) The location of the 35^{th} percentile is $(0.35)(18) = 6.3$. Thus

$$\widetilde{\pi}_{0.35} = (0.7)(4.8) + (0.3)(4.9) = 4.83.$$

The location of the 65^{th} percentile is $(0.65)(18) = 11.7$. Thus

$$\widetilde{\pi}_{0.65} = (0.3)(5.6) + (0.7)(5.6) = 5.6.$$

6.3–4

$$\begin{aligned}
g(y) &= \sum_{k=3}^{5} \left\{ \frac{6!}{k!(6-k)!}\,(k)[F(y)]^{k-1} f(y)[1-F(y)]^{6-k} \right. \\
&\quad \left. + \frac{6!}{k!(6-k)!}\,[F(y)]^{k}(6-k)[1-F(y)]^{6-k-1}[-f(y)] \right\} + 6[F(y)]^5 f(y) \\
&= \frac{6!}{2!3!}\,[F(y)]^2 f(y)[1-F(y)]^3 - \frac{6!}{3!2!}\,[F(y)]^3[1-F(y)]^2 f(y) \\
&\quad + \frac{6!}{3!2!}\,[F(y)]^3 f(y)[1-F(y)]^2 - \frac{6!}{4!1!}\,[F(y)]^4[1-F(y)]^1 f(y) \\
&\quad + \frac{6!}{4!1!}\,[F(y)]^4 f(y)[1-F(y)]^1 - \frac{6!}{5!0!}\,[F(y)]^5[1-F(y)]^0 f(y) + 6[F(y)]^5 f(y) \\
&= \frac{6!}{2!3!}\,[F(y)]^2[1-F(y)]^3 f(y), \qquad a < y < b.
\end{aligned}$$

6.3–6 **(a)**
$$\begin{aligned} F(x) &= x, \qquad 0 < x < 1. \text{ Thus} \\ g_1(w) &= n[1-w]^{n-1}(1), \qquad 0 < w < 1; \\ g_n(w) &= n[w]^{n-1}(1), \qquad 0 < w < 1. \end{aligned}$$

(b)
$$\begin{aligned} E(W_1) &= \int_0^1 (w)(n)(1-w)^{n-1}\,dw \\ &= \left[-w(1-w)^n - \frac{1}{n+1}\,(1-w)^{n+1}\right]_0^1 = \frac{1}{n+1}. \\ E(W_n) &= \int_0^1 (w)(n)w^{n-1}dw = \left[\frac{n}{n+1}w^{n+1}\right]_0^1 = \frac{n}{n+1}. \end{aligned}$$

(c) Let $w = w_r$. The pdf of W_r is
$$\begin{aligned} g_r(w) &= \frac{n!}{(r-1)!(n-r)!}[w]^{r-1}[1-w]^{n-r}\cdot 1 \\ &= \frac{\Gamma(r+n-r+1)}{\Gamma(r)\Gamma(n-r+1)}w^{r-1}(1-w)^{n-r+1-1}. \end{aligned}$$

Thus W_r has a beta distribution with $\alpha = r$, $\beta = n - r$.

6.3–8 **(a)**
$$\begin{aligned} E(W_r^2) &= \int_0^1 w^2\,\frac{n!}{(r-1)!(n-r)!}\,w^{r-1}(1-w)^{n-r}dw \\ &= \frac{r(r+1)}{(n+2)(n+1)}\int_0^1 \frac{(n+2)!}{(r+1)!(n-r)!}\,w^{r+1}(1-w)^{n-r}dw \\ &= \frac{r(r+1)}{(n+2)(n+1)} \end{aligned}$$

since the integrand is like that of a pdf of the $(r+2)$th order statistic of a sample of size $n+2$ and hence the integral must equal one.

(b) $\mathrm{Var}(W_r) = \dfrac{r(r+1)}{(n+2)(n+1)} - \dfrac{r^2}{(n+1)^2} = \dfrac{r(n-r+1)}{(n+2)(n+1)^2}\,.$

6.3–12

k	Strengths	$p = k/10$	z_{1-p}	k	Strengths	$p = k/10$	z_{1-p}
1	7.2	0.10	−1.282	6	11.7	0.60	0.253
2	8.9	0.20	−0.842	7	12.9	0.70	0.524
3	9.7	0.30	−0.524	8	13.9	0.80	0.842
4	10.5	0.40	−0.253	9	15.3	0.90	1.282
5	10.9	0.50	0.000				

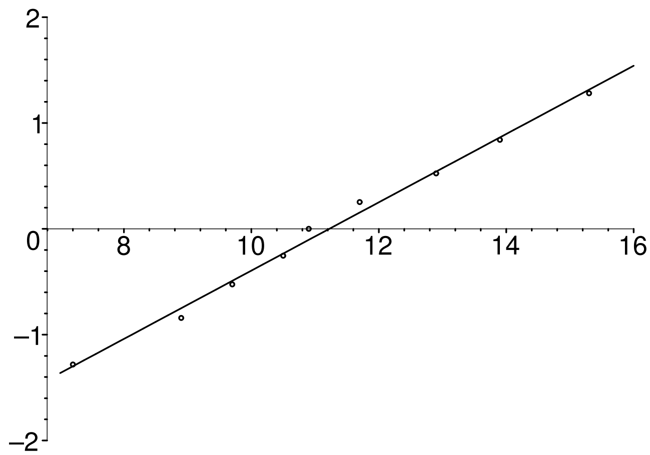


Figure 6.3–12: q–q plot of $N(0,1)$ quantiles versus data quantiles

It seems to be an excellent fit.

6.3–14 **(a)**

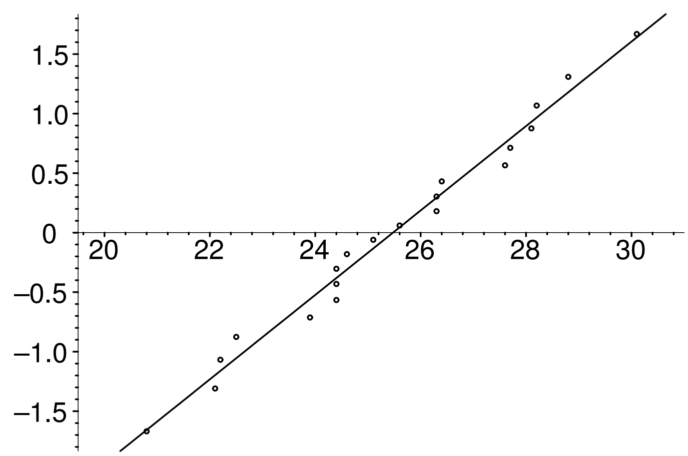


Figure 6.3–14: q–q plot of $N(0,1)$ quantiles versus data quantiles

(b) It looks like an excellent fit.

6.4 Maximum Likelihood Estimation

6.4–2 The likelihood function is

$$L(\theta) = \left[\frac{1}{2\pi\theta}\right]^{n/2} \exp\left[-\sum_{i=1}^{n}(x_i-\mu)^2/2\theta\right], \qquad 0<\theta<\infty.$$

The logarithm of the likelihood function is

$$\ln L(\theta) = -\frac{n}{2}(\ln 2\pi) - \frac{n}{2}(\ln\theta) - \frac{1}{2\theta}\sum_{i=1}^{n}(x_i-\mu)^2.$$

Setting the first derivative equal to zero and solving for θ yields

$$\begin{aligned}\frac{d\ln L(\theta)}{d\theta} &= -\frac{n}{2\theta} + \frac{1}{2\theta^2}\sum_{i=1}^{n}(x_i-\mu)^2 = 0\\ \theta &= \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2.\end{aligned}$$

Thus

$$\widehat{\theta} = \frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2.$$

To see that $\widehat{\theta}$ is an unbiased estimator of θ, note that

$$E(\widehat{\theta}) = E\left(\frac{\sigma^2}{n}\sum_{i=1}^{n}\frac{(X_i-\mu)^2}{\sigma^2}\right) = \frac{\sigma^2}{n}\cdot n = \sigma^2,$$

since $(X_i-\mu)^2/\sigma^2$ is $\chi^2(1)$ and hence the expected value of each of the n summands is equal to 1.

6.4–4 **(a)** $\overline{x} = 394/7 = 56.2857$; $s^2 = 5452/97 = 56.2062$;

(b) $\widehat{\lambda} = \overline{x} = 394/7 = 56.2857$;

(c) Yes;

(d) $\overline{x}$ is better than s^2 because

$$\mathrm{Var}(\overline{X}) \approx \frac{56.2857}{98} = 0.5743 < 65.8956 = \frac{56.2857[2(56.2857*98)+97]}{98(97)} \approx \mathrm{Var}(S^2).$$

6.4–6 $\widehat{\theta}_1 = \widehat{\mu} = 33.4267$; $\widehat{\theta}_2 = \widehat{\sigma^2} = 5.0980$.

6.4–8 **(a)**

$$\begin{aligned}L(\theta) &= \left(\frac{1}{\theta^n}\right)\left(\prod_{i=1}^{n}x_i\right)^{1/\theta-1}, \qquad 0<\theta<\infty\\ \ln L(\theta) &= -n\ln\theta + \left(\frac{1}{\theta}-1\right)\ln\prod_{i=1}^{n}x_i\\ \frac{d\ln L(\theta)}{d\theta} &= \frac{-n}{\theta} - \frac{1}{\theta^2}\ln\prod_{i=1}^{n}x_i = 0\\ \widehat{\theta} &= -\frac{1}{n}\ln\prod_{i=1}^{n}x_i\\ &= -\frac{1}{n}\sum_{i=1}^{n}\ln x_i.\end{aligned}$$

(b) We first find $E(\ln X)$:

$$E(\ln X) = \int_0^1 \ln x (1/\theta) x^{1/\theta - 1}\, dx.$$

Using integration by parts, with $u = \ln x$ and $dv = (1/\theta)x^{1/\theta-1}dx$,

$$E(\ln X) = \lim_{a \to 0} \Big[x^{1/\theta} \ln x - \theta x^{1/\theta} \Big]_a^1 = -\theta.$$

Thus

$$E(\,\widehat{\theta}\,) = -\frac{1}{n}\sum_{i=1}^{n}(-\theta) = \theta.$$

6.4–10 **(a)** $\overline{x} = 1/p$ so $\widetilde{p} = 1/\overline{X} = n/\sum_{i=1}^{n} X_i$;

(b) $\widetilde{p}$ equals the number of successes, n, divided by the number of Bernoulli trials, $\sum_{i=1}^{n} X_i$;

(c) $20/252 = 0.0794$.

6.4–12 **(a)** $E(\,\overline{X}\,) = E(Y)/n = np/n = p$;

(b) $\mathrm{Var}(\,\overline{X}\,) = \mathrm{Var}(Y)/n^2 = np(1-p)/n^2 = p(1-p)/n$;

(c)
$$\begin{aligned} E[\,\overline{X}(1-\overline{X}\,)/n] &= [E(\,\overline{X}\,) - E(\,\overline{X}^{\,2})]/n \\ &= \{p - [p^2 + p(1-p)/n]\}/n = [p(1-1/n) - p^2(1-1/n)]/n \\ &= (1-1/n)p(1-p)/n = (n-1)p(1-p)/n^2; \end{aligned}$$

(d) From part (c), the constant $c = 1/(n-1)$.

6.4–14 **(a)**
$$\begin{aligned} E(cS) &= E\left\{ \frac{c\sigma}{\sqrt{n-1}} \left[\frac{(n-1)S^2}{\sigma^2} \right]^{1/2} \right\} \\ &= \frac{c\sigma}{\sqrt{n-1}} \int_0^\infty \frac{v^{1/2} v^{(n-1)/2-1} e^{-v/2}}{\Gamma\left(\dfrac{n-1}{2}\right) 2^{(n-1)/2}}\, dv \\ &= \frac{c\sigma}{\sqrt{n-1}}\, \frac{\sqrt{2}\,\Gamma(n/2)}{\Gamma[(n-1)/2)]}, \end{aligned}$$

so $c = \dfrac{\sqrt{n-1}\,\Gamma[(n-1)/2]}{\sqrt{2}\,\Gamma(n/2)}$;

(b) When $n = 5$, $c = 8/(3\sqrt{2\pi}\,)$ and when $n = 6$, $c = 3\sqrt{5\pi}/(8\sqrt{2}\,)$.

(c)

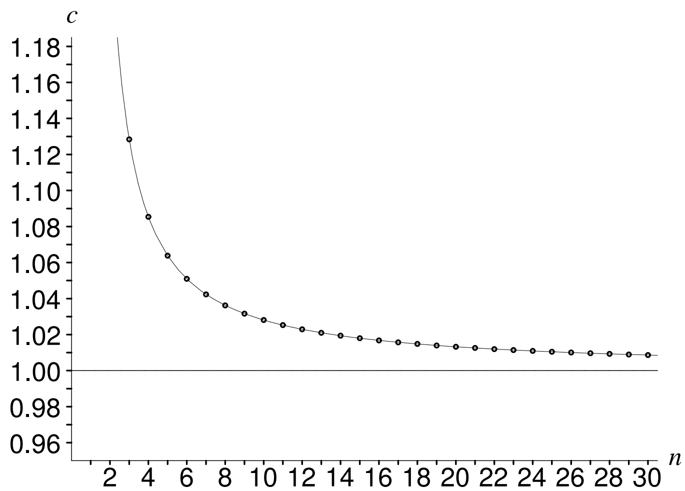


Figure 6.4–14: c as a function of n

We see that

$$\lim_{n\to\infty} c = 1.$$

6.4–16 The experiment has a hypergeometric distribution with $n = 8$ and $N = 64$. From the sample, $\overline{x} = 1.4667$. Using this as an estimate for μ we have

$$1.4667 = 8\left(\frac{N_1}{64}\right) \quad \text{implies that} \quad \widetilde{N_1} = 11.73.$$

A guess for the value of N_1 is therefore 12.

6.4–18

$$L = \left(\frac{1}{2\pi\sigma^2}\right)^{kn/2} \exp\left\{-\sum_{i=1}^{n}\sum_{j=1}^{k}\left[x_{ij} - c - d\left(j - \frac{k+1}{2}\right)\right]^2 / (2\sigma^2)\right\}.$$

To find $\widehat{c}$ and $\widehat{d}$ try to minimize

$$S = \sum_{i=1}^{n}\sum_{j=1}^{k}\left[x_{ij} - c - d\left(j - \frac{k+1}{2}\right)\right]^2.$$

We have

$$\frac{\partial S}{\partial c} = \sum_{i=1}^{n}\sum_{j=1}^{k} 2\left[x_{ij} - c - d\left(j - \frac{k+1}{2}\right)\right](-1) = 0.$$

Since

$$\sum_{j=1}^{k} d\left(-\frac{k+1}{2}\right) = 0, \qquad \widehat{c} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{k} x_{ij}}{kn} = \overline{x}, \quad \text{say}.$$

$$\frac{\partial S}{\partial d} = \sum_{i=1}^{n}\sum_{j=1}^{k} 2\left[x_{ij} - c - d\left(j - \frac{k+1}{2}\right)\right]\left[-\left(j - \frac{k+1}{2}\right)\right] = 0.$$

Thus

$$\widehat{d} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{k}(x_{ij} - \overline{x})(j - \{k+1\}/2)}{\sum_{i=1}^{n}\sum_{j=1}^{k}(j - \{k+1\}/2)^2} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{k} x_{ij}(j - \{k+1\}/2)}{\sum_{j=1}^{k} n(j - \{k+1\}/2)^2}.$$

6.5 A Simple Regression Problem

6.5–2 **(a)**
$$\begin{aligned} L(\beta, \sigma^2) &= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(y_i - \beta x_i)^2}{2\sigma^2}\right] \\ &= \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} \exp\left[-\frac{\sum_{i=1}^{n}(y_i - \beta x_i)^2}{2\sigma^2}\right]. \end{aligned}$$

Maximizing L or equivalently maximizing

$$\ln L(\beta, \sigma^2) = \frac{n}{2}\ln(2\pi\sigma^2) + \frac{\sum_{i=1}^{n}(y_i - \beta x_i)^2}{2\sigma^2}$$

requires selecting β to minimize $H(\beta) = \sum_{i=1}^{n}(y_i - \beta x_i)^2$.

$$\begin{aligned} H'(\beta) &= 2\sum_{i=1}^{n}(y_i - \beta x_i)(-x_i) \;=\; 0 \\ \widehat{\beta} &= \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2} \end{aligned}$$

which is a point of minimum since $H''(\widehat{\beta}) = 2\sum_{i=1}^{n} x_i^2 > 0$.

$$\begin{aligned} \frac{\partial \ln L(\beta, \sigma^2)}{\partial(\sigma^2)} &= \frac{n}{2\sigma^2} - \frac{\sum_{i=1}^{n}(y_i - \beta x_i)^2}{2(\sigma^2)^2} \;=\; 0 \\ \widehat{\sigma^2} &= \frac{1}{n}\sum_{i=1}^{n}(y_i - \widehat{\beta}x_i)^2. \end{aligned}$$

(b) $\widehat{\beta} = \left(\sum_{i=1}^{n} x_i^2\right)^{-1} \sum_{i=1}^{n} x_i Y_i$. Since $\widehat{\beta}$ is a linear combination of independent normal random variables, it is normal with mean

$$\begin{aligned} E(\widehat{\beta}) &= \left(\sum_{i=1}^{n} x_i^2\right)^{-1} \sum_{i=1}^{n} x_i E(Y_i) \\ &= \left(\sum_{i=1}^{n} x_i^2\right)^{-1} \sum_{i=1}^{n} x_i(\beta x_i) \\ &= \left(\sum_{i=1}^{n} x_i^2\right)^{-1} \beta \left(\sum_{i=1}^{n} x_i^2\right) \\ &= \beta \end{aligned}$$

and variance

$$\begin{aligned} \mathrm{Var}(\widehat{\beta}) &= \left(\sum_{i=1}^{n} x_i^2\right)^{-2} \sum_{i=1}^{n} x_i^2\,\mathrm{Var}(Y_i) \\ &= \left(\sum_{i=1}^{n} x_i^2\right)^{-2} \sum_{i=1}^{n} x_i^2 \sigma^2 \\ &= \sigma^2 \Big/ \sum_{i=1}^{n} x_i^2. \end{aligned}$$

Also

$$\sum_{i=1}^{n}(Y_i - \beta x_i)^2 = \sum_{i=1}^{n}[Y_i - \widehat{\beta}x_i + (\widehat{\beta}-\beta)x_i]^2$$
$$= \sum_{i=1}^{n}(Y_i - \widehat{\beta}x_i)^2 + (\widehat{\beta}-\beta)^2\sum_{i=1}^{n}x_i^2.$$

Since $Y_i \sim$ independent $N(\beta x_i, \sigma^2)$ and $\widehat{\beta} \sim N(\beta, \sigma^2 \sum_{i=1}^{n} x_i^2)$,

$$\frac{\sum_{i=1}^{n}(Y_i - \beta x_i)^2}{\sigma^2} \sim \chi^2(n) \qquad \text{and} \qquad \frac{(\widehat{\beta}-\beta)^2\sum_{i=1}^{n}x_i^2}{\sigma^2} \sim \chi^2(1).$$

By Theorem 9.3-1,

$$\sum_{i=1}^{n}(Y_i - \widehat{\beta}x_i)^2 \sim \chi^2(n-1).$$

6.5–4 (a)

x	y	x^2	xy	y^2	$(y-\widehat{y})^2$
2.0	1.3	4.00	2.60	1.69	0.361716
3.3	3.3	10.89	10.89	10.89	0.040701
3.7	3.3	13.69	12.21	10.89	0.027725
2.0	2.0	4.00	4.00	4.00	0.009716
2.3	1.7	5.29	3.91	2.89	0.228120
2.7	3.0	7.29	8.10	9.00	0.206231
4.0	4.0	16.00	16.00	16.00	0.006204
3.7	3.0	13.69	11.10	9.00	0.217630
3.0	2.7	9.00	8.10	7.29	0.014900
2.3	3.0	5.29	6.90	9.00	0.676310
29.0	27.3	89.14	83.81	80.65	1.849254

$$\widehat{\alpha} = \overline{y} = 27.3/10 = 2.73;$$
$$\widehat{\beta} = \frac{83.81 - (29.0)(27.3)/10}{89.14 - (29.0)(29.0)/10} = \frac{4.64}{5.04} = 0.9206;$$
$$\widehat{y} = 2.73 + (4.64/5.04)(x - 2.90)$$

(b)

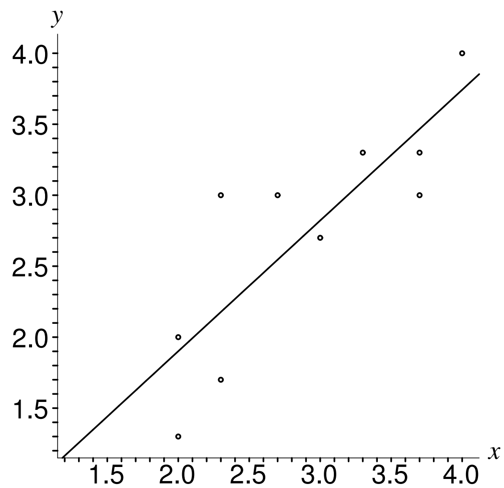


Figure 6.5–4: Earned grade (y) versus predicted grade (x)

(c) $\widehat{\sigma^2} = \dfrac{1.849254}{10} = 0.184925.$

6.5–6 **(a)**

$$\begin{aligned}
\widehat{\alpha} &= \frac{395}{15} = 26.333, \\
\widehat{\beta} &= \frac{9292 - (346)(395)/15}{8338 - (346)^2/15} = \frac{180.667}{356.933} = 0.506, \\
\widehat{y} &= 26.333 + \frac{180.667}{356.933}\left(x - \frac{346}{15}\right) \\
&= 0.506x + 14.657;
\end{aligned}$$

(b)

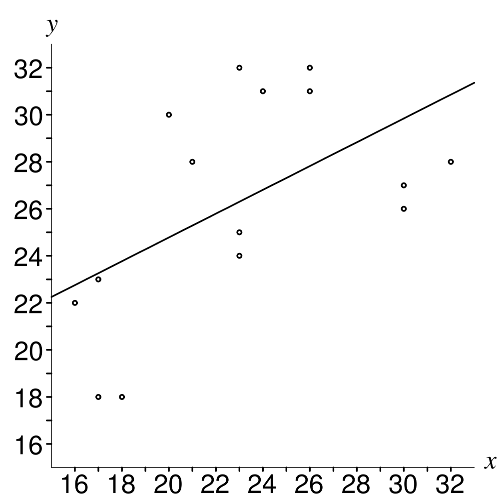


Figure 6.5–6: ACT natural science (y) versus ACT social science (x) scores

(c)
$$\begin{aligned}\widehat{\alpha} &= 26.33,\ \widehat{\beta} = 0.506,\\ n\widehat{\sigma^2} &= 10,705 - \frac{395^2}{15} - 0.5061636(9292) + 0.5061636(346)(395)/15\\ &= 211.8861,\\ \widehat{\sigma^2} &= \frac{211.8861}{15} = 14.126.\end{aligned}$$

6.5–8 **(a)** $y = 3.575 + 1.225x$.

(b)

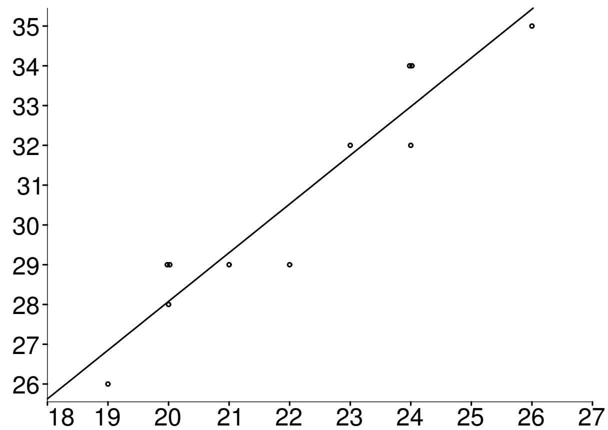


Figure 6.5–8: **(b)** Highway mpg (y) versus City mpg (x)

(c) $y = 54.6018 - 0.0072x$.

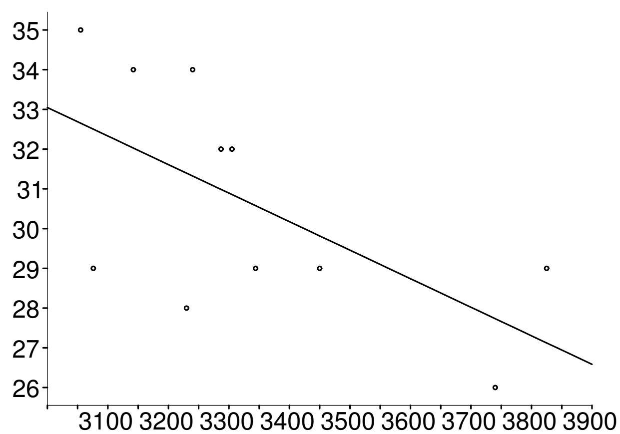
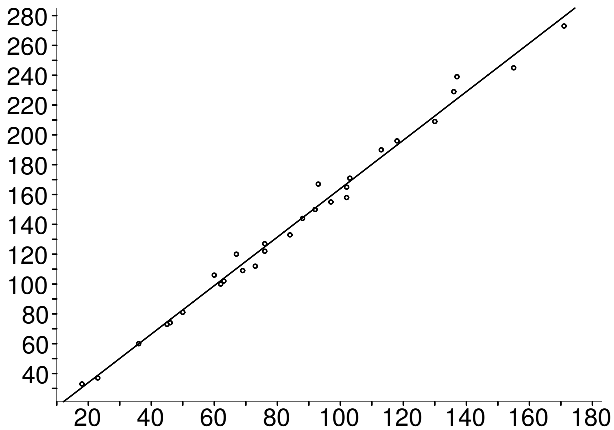


Figure 6.5–8: **(c)** Highway mpg (y) versus Curb Weight (x)

6.5–10 **(b)** The least squares regression line for $y = a + b$ versus b is $\widehat{y} = 1.360 + 1.626b$;

(c) $y = \phi x = 1.618x$ is added on the right figure below.



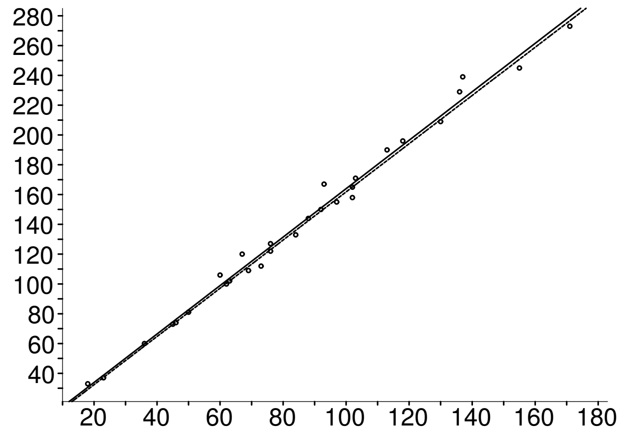


Figure 6.5–10: Scatter plot of $a + b$ versus b with least squares regression line and with $y = \phi x$

(d) The sample mean of the points $(a+b)/b$ is 1.647 which is close to the value of $\phi = 1.618$.

6.6 Asymptotic Distributions of Maximum Likelihood Estimators

6.6–2 **(a)**

$$\begin{aligned}
f(x;p) &= p^x(1-p)^{1-x}, \qquad x = 0, 1 \\
\ln f(x;p) &= x\ln p + (1-x)\ln(1-p) \\
\frac{\partial \ln f(x;p)}{\partial p} &= \frac{x}{p} + \frac{x-1}{1-p} \\
\frac{\partial^2 \ln f(x;p)}{\partial p^2} &= -\frac{x}{p^2} + \frac{x-1}{(1-p)^2} \\
E\left[\frac{X}{p^2} - \frac{X-1}{(1-p)^2}\right] &= \frac{p}{p^2} - \frac{p-1}{(1-p)^2} = \frac{1}{p(1-p)}.
\end{aligned}$$

Rao-Cramér lower bound $= \dfrac{p(1-p)}{n}$.

(b) $\dfrac{p(1-p)/n}{p(1-p)/n} = 1.$

6.6–4 **(a)**

$$\begin{aligned}
\ln f(x;\theta) &= -2\ln\theta + \ln x - x/\theta \\
\frac{\partial \ln f(x;\theta)}{\partial \theta} &= -\frac{2}{\theta} + \frac{x}{\theta^2} \\
\frac{\partial^2 \ln f(x;\theta)}{\partial \theta^2} &= \frac{2}{\theta^2} - \frac{2x}{\theta^3} \\
E\left[-\frac{2}{\theta^2} + \frac{2X}{\theta^3}\right] &= -\frac{2}{\theta^2} + \frac{2(2\theta)}{\theta^3} = \frac{2}{\theta^2}
\end{aligned}$$

Rao-Cramér lower bound $= \dfrac{\theta^2}{2n}$.

(b) Very similar to (a); answer $= \dfrac{\theta^2}{3n}$.

(c)

$$\begin{aligned}
\ln f(x;\theta) &= -\ln\theta + \left(\frac{1-\theta}{\theta}\right)\ln x \\
\frac{\partial \ln f(x;\theta)}{\partial\theta} &= -\frac{1}{\theta} - \frac{1}{\theta^2}\ln x \\
\frac{\partial^2 \ln f(x;\theta)}{\partial\theta^2} &= \frac{1}{\theta^2} + \frac{2}{\theta^2}\ln x \\
E[\ln X\,] &= \int_0^1 \frac{\ln x}{\theta}\, x^{(1-\theta)/\theta} dx. \quad \text{Let } y = \ln x,\ dy = \frac{1}{x}\,dx. \\
&= -\int_0^\infty \frac{y}{\theta}\, e^{-y(1-\theta)/\theta} e^{-y} dy = -\theta\,\Gamma(2) = -\theta.
\end{aligned}$$

$$\text{Rao-Cramér lower bound} = \frac{1}{n\left(-\dfrac{1}{\theta^2} + \dfrac{2}{\theta^2}\right)} = \frac{\theta^2}{n}.$$

6.7 Sufficient Statistics

6.7–2 The distribution of Y is Poisson with mean $n\lambda$. Thus, since $y = \Sigma\, x_i$,

$$\begin{aligned}
P(X_1 = x_1,\ X_2 = x_2,\ \cdots,\ X_n = x_n \,|\, Y = y) &= \frac{(\lambda^{\Sigma\, x_i} e^{-n\lambda})/(x_1!\,x_2!\cdots x_n!)}{(n\lambda)^y\, e^{-n\lambda}/y!} \\
&= \frac{y!}{x_1!\ x_2!\cdots x_n!\, n^y}\,,
\end{aligned}$$

which does not depend on λ.

6.7–4 **(a)** $f(x;\theta) = e^{(\theta-1)\ln x + \ln\theta}, \qquad 0 < x < 1, \qquad 0 < \theta < \infty;$

so $K(x) = \ln x$ and thus

$$Y = \sum_{i=1}^n \ln X_i = \ln(X_1 X_2 \cdots X_n)$$

is a sufficient statistic for θ.

(b)

$$\begin{aligned}
L(\theta) &= \theta^n (x_1 x_2 \cdots x_n)^{\theta-1} \\
\ln L(\theta) &= n\ln\theta + (\theta-1)\ln(x_1 x_2 \cdots x_n) \\
\frac{d\ln L(\theta)}{d\theta} &= \frac{n}{\theta} + \ln(x_1 x_2 \cdots x_n) = 0.
\end{aligned}$$

Hence

$$\widehat{\theta} = -n/\ln(X_1 X_2 \cdots X_n),$$

which is a function of Y.

(c) Since $\widehat{\theta}$ is a single valued function of Y with a single valued inverse, knowing the value of $\widehat{\theta}$ is equivalent to knowing the value of Y, and hence it is sufficient.

6.7–6 **(a)**

$$\begin{aligned}
f(x_1, x_2, \ldots, x_n) &= \frac{(x_1 x_2 \cdots x_n)^{\alpha-1} e^{-\Sigma\, x_i/\theta}}{[\Gamma(\alpha)]^n \theta^{\alpha n}} \\
&= \left(\frac{e^{-\Sigma\, x_i/\theta}}{\theta^{\alpha n}}\right)\left(\frac{(x_1 x_2 \cdots x_n)^{\alpha-1}}{[\Gamma(\alpha)]^n}\right).
\end{aligned}$$

The second factor is free of θ. The first factor is a function of the $x_i s$ through $\sum_{i=1}^n x_i$ only, so $\sum_{i=1}^n x_i$ is a sufficient statistic for θ.

(b)
$$\begin{aligned}
\ln L(\theta) &= \ln(x_1x_2\cdots x_n)^{\alpha-1} - \textstyle\sum_{i=1}^n x_i/\theta - \ln[\Gamma(\alpha)]^n - \alpha n \ln\theta \\
\frac{d\ln L(\theta)}{d\theta} &= \textstyle\sum_{i=1}^n x_i/\theta^2 - \alpha n/\theta \;=\; 0 \\
\alpha n\theta &= \textstyle\sum_{i=1}^n x_i \\
\widehat{\theta} &= \frac{1}{\alpha n}\sum_{i=1}^n X_i.
\end{aligned}$$

$Y = \displaystyle\sum_{i=1}^n X_i$ has a gamma distribution with parameters αn and θ. Hence

$$E(\widehat{\theta}) = \frac{1}{\alpha n}(\alpha n\theta) = \theta.$$

6.7–8

$$E(e^{tZ}) = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\left(\frac{1}{\sqrt{2\pi\theta}}\right)^n \exp\left[-\sum_{i=1}^n \frac{x_i^2}{2\theta}\right]\exp\left[\frac{t\sum_{i=1}^n a_ix_i}{\sum_{i=1}^n x_i}\right]dx_1dx_2\ldots dx_n.$$

Let $x_i/\sqrt{\theta} = y_i,\ i = 1, 2, \ldots, n$. The Jacobian is $(\sqrt{\theta}\,)^n$. Hence

$$\begin{aligned}
E(e^{tZ}) &= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}(\sqrt{\theta})^n\left(\frac{1}{\sqrt{2\pi\theta}}\right)^n \exp\left[-\sum_{i=1}^n \frac{y_i^2}{2}\right]\exp\left[\frac{t\sum_{i=1}^n a_iy_i}{\sum_{i=1}^n y_i}\right]dy_1dy_2\ldots dy_n. \\
&= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\left(\frac{1}{\sqrt{2\pi}}\right)^n \exp\left[-\sum_{i=1}^n \frac{y_i^2}{2}\right]\exp\left[\frac{t\sum_{i=1}^n a_iy_i}{\sum_{i=1}^n y_i}\right]dy_1dy_2\ldots dy_n
\end{aligned}$$

which is free of θ. Since the distribution of Z is free of θ, Z and $Y = \sum_{i=1}^n X_i^2$, the sufficient statistics, are independent.

6.7–10 The joint pdf can be written

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

so by definition

$$\prod_{i=1}^n X_i(1-X_i)$$

is a sufficient statistics for θ.

6.7–12 The mean Y is a complete sufficient statistic for μ in $N(\mu,\sigma^2)$. The difference $Y - Z$ has a distribution free of μ; so Y and $Y - Z$ are independent. Now $Z = Y + (Z - Y)$ and from that independence

$$\mathrm{Var}(Z) = \mathrm{Var}(Y) + \mathrm{Var}(Z - Y) = \frac{\sigma^2}{n} + \mathrm{Var}(Z - Y).$$

Since the $\mathrm{Var}(Z - Y)$ is much smaller than $\mathrm{Var}(Z)$, in order to get an estimate of $\mathrm{Var}(Z)$ by Monte Carlo, it would be much easier to estimate $\mathrm{Var}(Z-Y)$ by Monte Carlo and thus get an estimate of $\mathrm{Var}(Z)$. This trick can be used often in practice and is called the Monte Carlo Swindle.

6.8 Bayesian Estimation

6.8–2 **(a)**

$$\begin{aligned} g(\tau \mid x_1, x_2, \ldots, x_n) &\propto \frac{(x_1 x_2 \cdots x_n)^{\alpha - 1}}{[\Gamma(\alpha)]^n} \frac{\tau^{n\alpha} \tau^{\alpha_0 - 1} e^{-\tau/\theta_0} e^{-\Sigma x_i/(1/\tau)}}{\Gamma(\alpha_0)\theta_0^{\alpha_0}} \\ &\propto \tau^{\alpha n + \alpha_0 - 1} e^{-(1/\theta_0 + \Sigma x_i)\tau} \end{aligned}$$

which is $\Gamma\left(n\alpha + \alpha_0, \dfrac{\theta_0}{1 + \theta_0 \Sigma x_i}\right)$.

(b)

$$\begin{aligned} E(\tau \mid x_1, x_2, \ldots, x_n) &= (n\alpha + \alpha_0)\frac{\theta_0}{1 + \theta_0 \overline{X} n} \\ &= \frac{\alpha_0 \theta_0}{1 + \theta_0 n \overline{X}} + \frac{\alpha n \theta_0}{1 + n\theta_0 \overline{X}} \\ &= \frac{n\alpha + \alpha_0}{1/\theta_0 + n\overline{X}}. \end{aligned}$$

(c) The posterior distribution is $\Gamma(30 + 10, 1/[1/2 + 10\,\overline{x}\,])$. Select a and b so that $P(a < \tau < b) = 0.95$ with equal tail probabilities. Then

$$\int_a^b \frac{(1/2 + 10\,\overline{x}\,)^{40}}{\Gamma(40)} w^{40-1} e^{-w(1/2 + 10\,\overline{x}\,)} dw = \int_{a(1/2+10\overline{x})}^{b(1/2+10\overline{x})} \frac{1}{\Gamma(40)} z^{39} e^{-z} dz,$$

making the change of variables $w(1/2 + 10\,\overline{x}\,) = z$. Let $v_{0.025}$ and $v_{0.975}$ be the quantiles for the $\Gamma(40, 1)$ distribution. Then

$$a = \frac{v_{0.025}}{1/2 + 10\,\overline{x}};$$

$$b = \frac{v_{0.975}}{1/2 + 10\,\overline{x}}.$$

It follows that

$$P(a < \tau < b) = 0.95.$$

6.8–4

$$(3\theta)^n (x_1 x_2 \cdots x_n)^2 e^{-\theta \Sigma x_i^3} \cdot \theta^{4-1} e^{-4\theta} \propto \theta^{n+3} e^{-(4 + \Sigma x_i^3)\theta}$$

which is $\Gamma\left(n + 4, \dfrac{1}{4 + \Sigma\, x_i^3}\right)$. Thus

$$E(\theta \mid x_1, x_2, \ldots, x_n) = \frac{n + 4}{4 + \Sigma\, x_i^3}.$$

6.8–6 **(a)**

$$\begin{aligned} g(y\,|\,\theta) &= \frac{ny^{n-1}}{\theta^n}, \qquad 0 < y < \theta \\ g(y\,|\,\theta)h(\theta) &\propto \frac{1}{\theta^{n+\beta+1}}, \qquad \max(\alpha, y) < \theta < \infty \\ k(\theta\,|\,y) &= \frac{c}{\theta^{n+\beta+1}}, \qquad \max(\alpha, y) < \theta < \infty \\ c &= [\max(\alpha, y)]^{n+\beta} \;\text{ by }\; \int_{\max(\alpha,y}^{\infty} k(\theta\,|\,y)\,d\theta = 1 \\ E(\theta\,|\,y) &= \int_{\max(\alpha,y)}^{\infty} (\theta)\frac{[\max(\alpha,y)]^{n+\beta}(n+\beta)}{\theta^{n+\beta+1}}\,d\theta \\ &= \left(\frac{n+\beta}{n+\beta-1}\right)\max(\alpha, y) \;=\; w(y) \\ w(Y) &= \left(\frac{n+\beta}{n+\beta-1}\right)\max(\alpha, Y). \end{aligned}$$

(b) With $n = 4$, $\alpha = 1$, $\beta = 2$,

$$k(\theta\,|\,y) = \frac{[\max(1,y)]^6(6)}{\theta^7}, \qquad \max(1, y) < \theta.$$

Since

$$\text{Loss } = |\theta - w(Y)|,$$

we want

$$\begin{aligned} w(y) &= \text{ median} \\ \frac{1}{2} &= \int_{\max(1,y)}^{w(y)} \frac{[\max(1,y)]^6(6)}{\theta^7}\,dy \\ &= 1 - \frac{[\max(1,y)]^6}{[w(y)]^6} \end{aligned}$$

so

$$w(Y) = \sqrt[6]{2}\,\max(1, Y).$$

6.8–8

$$L = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left\{-\sum_{i=1}^{n}[y_i - \alpha - \beta(x_i - \overline{x})]^2/2\sigma^2\right\}.$$

The summation in the exponent can be written as follows:

$$\sum_{i=1}^{n}[y_i - \alpha - \beta(x_i - \overline{x})] = n(\widehat{\alpha} - \alpha)^2 + (\widehat{\beta} - \beta)^2\sum_{i=1}^{n}(x_i - \overline{x})^2 + \sum_{i=1}^{n}[y_i - \widehat{\alpha} - \widehat{\beta}(x_i - \overline{x})]^2.$$

Note: Let's say σ^2 is known in the exercise. Then α has prior $N(\alpha_1, \sigma_1^2)$; so posterior of α (eliminating factors <u>not</u> involving α) is

$$\begin{aligned}
\text{posterior of } \alpha \quad &\propto \left(\frac{1}{\sqrt{2\pi\sigma_1^2}}\right)\exp\left[\frac{-(\alpha-\alpha_1)^2}{2\sigma_1^2}\right]\exp\left[\frac{-n(\alpha-\widehat{\alpha})^2}{2\sigma^2}\right] \\
&\propto \exp\left[-\frac{(\alpha^2-2\alpha\alpha_1)}{2\sigma_1^2}-\frac{(n\alpha^2-2n\widehat{\alpha}\alpha)}{2\sigma^2}\right]+\text{terms not involving } \alpha \\
&= \exp\left[-\left\{\alpha^2\left(\frac{1}{2\sigma_1^2}+\frac{n}{2\sigma^2}\right)-2\alpha\left(\frac{\alpha_1}{2\sigma_1^2}+\frac{n\widehat{\alpha}}{2\sigma^2}\right)\right\}\right]+ \\
&\qquad\qquad\qquad\qquad \text{terms not involving } \alpha \\
&= \exp\left[-\left(\frac{1}{2\sigma_1^2}+\frac{n}{2\sigma^2}\right)\left\{\alpha^2-2\alpha\frac{\alpha_1/(2\sigma_1^2)+n\widehat{\alpha}/(2\sigma^2)}{1/(2\sigma_1^2)+n/(2\sigma^2)}\right\}\right]+ \\
&\qquad\qquad\qquad\qquad \text{terms not involving } \alpha \\
&= \exp\left[-\left(\frac{1}{2\sigma_1^2}+\frac{n}{2\sigma^2}\right)\left\{\alpha-\frac{\alpha_1/(2\sigma_1^2)+n\widehat{\alpha}/(2\sigma^2)}{1/(2\sigma_1^2)+n/(2\sigma^2)}\right\}^2\right]+ \\
&\qquad\qquad\qquad\qquad \text{terms not involving } \alpha.
\end{aligned}$$

This is normal with posterior mean equal to

$$\frac{\dfrac{\alpha_1}{\sigma_1^2}+\dfrac{n\widehat{\alpha}}{\sigma^2}}{\dfrac{1}{\sigma_1^2}+\dfrac{n}{\sigma^2}}.$$

Now β has prior $N(\beta_0, \sigma_0^2)$; so posterior of β is

$$\begin{aligned}
\text{posterior of } \beta \quad &\propto \exp\left[-\frac{(\beta-\widehat{\beta})^2\sum_{i=1}^n(x_i-\overline{x})^2}{2\sigma^2}-\frac{(\beta-\beta_0)^2}{2\sigma_0^2}\right] \\
&\propto \exp\left[-\left(\frac{\sum_{i=1}^n(x_i-\overline{x})^2}{2\sigma^2}+\frac{1}{2\sigma_0^2}\right)\beta^2+2\left(\frac{\widehat{\beta}\sum_{i=1}^n(x_i-\overline{x})^2}{2\sigma^2}+\frac{\beta_0}{2\sigma_0^2}\right)\beta\right] \\
&\propto \exp\left[-\left(\frac{\sum_{i=1}^n(x_i-\overline{x})^2}{2\sigma^2}+\frac{1}{2\sigma_0^2}\right)\left\{\beta-\frac{\widehat{\beta}\sum_{i=1}^n(x_i-\overline{x})^2/\sigma^2+\beta_0/\sigma_0^2}{\sum_{i=1}^n(x_i-\overline{x})^2/\sigma^2+1/\sigma_0^2}\right\}^2\right].
\end{aligned}$$

This is normal with posterior mean

$$\frac{\widehat{\beta}\sum_{i=1}^n(x_i-\overline{x})^2/\sigma^2+\beta_0/\sigma_0^2}{\sum_{i=1}^n(x_i-\overline{x})^2/\sigma^2+1/\sigma_0^2}.$$

Hence the posterior mean of $\alpha+\beta(x-\overline{x})$ is

$$\frac{\dfrac{\alpha_1}{\sigma_1^2}+\dfrac{n\widehat{\alpha}}{\sigma^2}}{\dfrac{1}{\sigma_1^2}+\dfrac{n}{\sigma^2}}+\frac{\dfrac{\widehat{\beta}\sum_{i=1}^n(x_i-\overline{x})^2}{\sigma^2}+\dfrac{\beta_0}{\sigma_0^2}}{\dfrac{\sum_{i=1}^n(x_i-\overline{x})^2}{\sigma^2}+\dfrac{1}{\sigma_0^2}}(x-\overline{x}).$$

6.9 More Bayesian Concepts

6.9–2 $k(x,\theta)=\dbinom{n}{x}\theta^x(1-\theta)^{n-x}\cdot\dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1}, \qquad x=0,1,\ldots,n,\quad 0<\theta<1.$

$$\begin{aligned} k_1(x) &= \int_0^1 \binom{n}{x} \theta^x (1-\theta)^{n-x} \cdot \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\, \theta^{\alpha-1}(1-\theta)^{\beta-1} d\theta \\ &= \frac{n!\,\Gamma(\alpha+\beta)\,\Gamma(x+\alpha)\,\Gamma(n-x+\beta)}{x!\,(n-x)!\,\Gamma(\alpha)\,\Gamma(\beta)\,\Gamma(n+\alpha+\beta)}, \qquad x = 0, 1, 2, \ldots, n. \end{aligned}$$

6.9–4

$$\begin{aligned} k(x,\theta) &= \int_0^\infty \theta\tau x^{\tau-1} e^{-\theta x^\tau} \cdot \frac{1}{\Gamma(\alpha)\beta^\alpha}\, \theta^{\alpha-1} e^{-\theta/\beta} d\theta, \qquad 0 < x < \infty \\ &= \frac{\tau x^{\tau-1}}{\Gamma(\alpha)\beta^\alpha} \int_0^\infty \theta^{\alpha+1-1} e^{-(x^\tau+1/\beta)\theta} d\theta \\ &= \frac{\tau x^{\tau-1}}{\Gamma(\alpha)\beta^\alpha} \frac{\Gamma(\alpha+1)}{(x^\tau+1/\beta)^{\alpha+1}}, \qquad 0 < x < \infty \\ &= \frac{\alpha\beta\tau x^{\tau-1}}{(\beta x^\tau+1)^{\alpha+1}}, \qquad 0 < x < \infty. \end{aligned}$$

6.9–6

$$g(\theta_1, \theta_2 \,|\, x_1 = 3, x_2 = 7) \propto \left(\frac{1}{\pi}\right)^2 \frac{\theta_2^2}{[\theta_2^2 + (3-\theta_1)^2][\theta_2^2 + (7-\theta_1)^2]}.$$

The figures show the graph of

$$h(\theta_1, \theta_2) = \frac{\theta_2^2}{[\theta_2^2 + (3-\theta_1)^2][\theta_2^2 + (7-\theta_1)^2]}.$$

The second figure shows a contour plot.

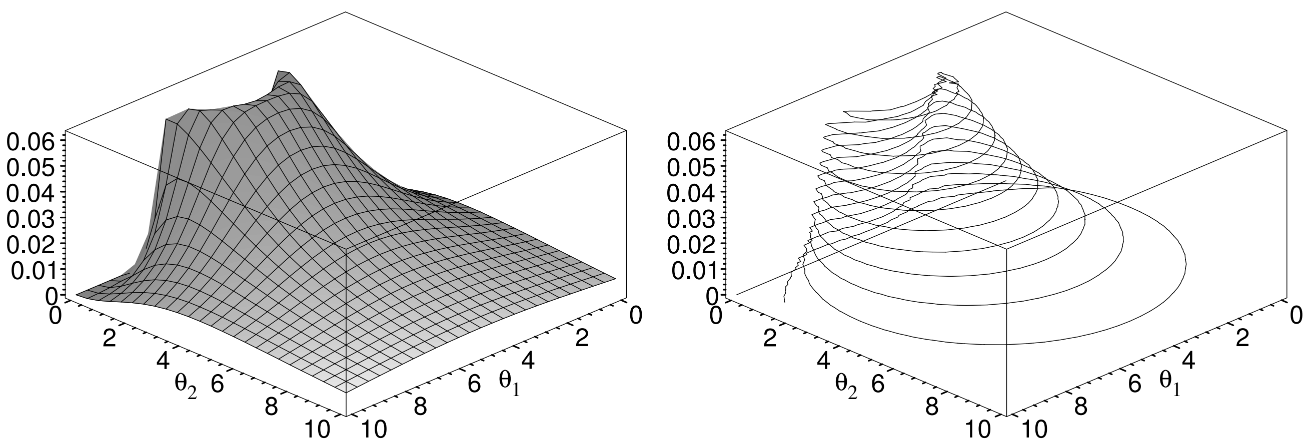


Figure 6.9–6: Graphs to help to see where θ_1 and θ_2 maximize the posterior pdf

Using *Maple*, a solution is $\theta_1 = 5$ and $\theta_2 = 2$. Other solutions satisfy

$$\theta_2 = \sqrt{-\theta_1^2 + 10\theta_1 - 21}.$$

Chapter 7

Interval Estimation

7.1 Confidence Intervals for Means

7.1–2 **(a)** $[77.272,\ 92.728]$; **(b)** $[79.12,\ 90.88]$;
(c) $[80.065,\ 89.935]$; **(d)** $[81.154,\ 88.846]$.

7.1–4 **(a)** $\overline{x} = 56.8$;

(b) $\left[56.8 - 1.96(2/\sqrt{10}\,),\ \ 56.8 + 1.96(2/\sqrt{10}\,\right] = [55.56,\ 58.04]\,$;

(c) $P(X < 52) = P\left(Z < \dfrac{52 - 56.8}{2}\right) = P(Z < -2.4) = 0.0082.$

7.1–6 $\left[11.95 - 1.96\left(\dfrac{11.80}{\sqrt{37}}\right),\ \ 11.95 + 1.96\left(\dfrac{11.80}{\sqrt{37}}\right)\right] = [8.15,\ 15.75].$

If more extensive t-tables are available or if a computer program is used, we have

$$\left[11.95 - 2.028\left(\frac{11.80}{\sqrt{37}}\right),\ \ 11.95 + 2.028\left(\frac{11.80}{\sqrt{37}}\right)\right] = [8.016,\ 15.884].$$

7.1–8 **(a)** $\overline{x} = 46.42$;

(b) $46.72 \pm 2.132s/\sqrt{5}$ or $[40.26,\ 52.58]$.

7.1–10 **(a)** $\overline{x} = 3.580$;

(b) $s = 0.512$;

(c) $[0,\ 3.580 + 1.833(0.512/\sqrt{10}\,] = [0,\ 3.877]$.

7.1–12 **(a)** $\overline{x} = 245.80$, $s = 23.64$, so a 95% confidence interval for μ is

$$[245.80 - 2.145(23.64)/\sqrt{15}\,,\ 245.80 + 2.145(23.64)/\sqrt{15}\,] = [232.707,\ 258.893];$$

(b)

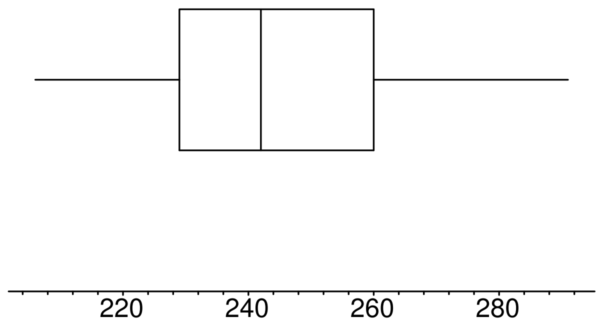


Figure 7.1–12: Box-and-whisker diagram of signals from detectors

(c) The standard deviation is quite large.

7.1–14 **(a)** $(\,\overline{x}+1.96\sigma/\sqrt{5}\,)-(\,\overline{x}-1.96\sigma/\sqrt{5}\,)=3.92\sigma/\sqrt{5}=1.753\sigma$;

(b) $(\,\overline{x}+2.776s/\sqrt{5}\,)-(\,\overline{x}-2.776s/\sqrt{5}\,)=5.552s/\sqrt{5}$.

From Exercise 6.4–14 with $n=5$, $E(S)=\dfrac{\sqrt{2}\,\Gamma(5/2)\sigma}{\sqrt{4}\,\Gamma(4/2)}=\dfrac{3\sqrt{\pi}\sigma}{2^{5/2}}=0.94\sigma$, so that $E[5.552S/\sqrt{5}\,]=2.334\sigma$.

7.1–16

$$\begin{aligned} P\left[a\le\frac{(n-1)S^2}{\sigma^2}\le b\right] &= 1-\alpha \\ P\left[\frac{(n-1)S^2}{b}\le\sigma^2\le\frac{(n-1)S^2}{a}\right] &= 1-\alpha. \end{aligned}$$

Letting $a=\chi^2_{1-\alpha/2}(n-1)$ and $b=\chi^2_{\alpha/2}(n-1)$, a $100(1-\alpha)\%$ confidence interval for σ^2 is

$$\left[\frac{(n-1)s^2}{\chi^2_{\alpha/2}(n-1)},\,\frac{(n-1)s^2}{\chi^2_{1-\alpha/2}(n-1)}\right].$$

Thus a 90% confidence interval for σ^2 is

$$\left[\frac{128.41}{21.03},\,\frac{128.41}{5.226}\right]=[6.11, 24.57].$$

It follows that a 90% confidence interval for σ is

$$[\sqrt{6.11},\sqrt{24.57}\,]=[2.47, 4.96].$$

7.2 Confidence Intervals for the Difference of Two Means

7.2–2 $\overline{x}=539.2$, $s_x^2=4{,}948.7$, $\overline{y}=544.625$, $s_y^2=4{,}327.982$, $s_p=67.481$, $t_{0.05}(11)=1.796$, so the confidence interval is $[-74.517,\ 63.667]$.

7.2–4 **(a)** $\overline{x}-\overline{y}=1511.714-1118.400=393.314$;

(b) $s_x^2=49,669.905$, $s_y^2=15,297.600$, $r=[8.599]=8$, $t_{0.025}(8)=2.306$, so the confidence interval is $[179.148,\ 607.480]$.

7.2–6 **(a)** $\overline{x}=2.584$, $\overline{y}=1.564$, $s_x^2=0.1042$, $s_y^2=0.0428$, $s_p=0.2711$, $t_{0.025}(18)=2.101$. Thus a 95% confidence interval for $\mu_X-\mu_Y$ is $[0.7653,\ 1.2747]$.

(b)

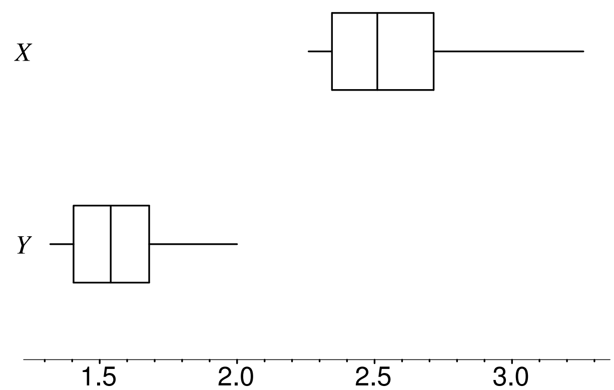


Figure 7.2–6: Box-and-whisker diagrams, wedge on (X) and wedge off (Y)

(c) Yes.

7.2–8 From (a), (b), and (c), we know

(d) $$\frac{\overline{X}-\overline{Y}-(\mu_X-\mu_Y)}{\sqrt{\dfrac{d\sigma_Y^2}{n}+\dfrac{\sigma_Y^2}{m}}} \div \sqrt{\left[\frac{(n-1)S_X^2}{d\sigma_Y^2}+\frac{(m-1)S_Y^2}{\sigma_Y^2}\right]\Big/ (n+m-2)}$$

has a $t(n+m-2)$ distribution. Clearly, this ratio does not depend upon σ_Y^2; so

$$\overline{x}-\overline{y}\pm t_{\alpha/2}(n+m-2)\sqrt{\frac{(n-1)s_x^2/d+(m-1)s_y^2}{n+m-2}\left(\frac{d}{n}+\frac{1}{m}\right)}$$

provides a $100(1-\alpha)\%$ confidence interval for $\mu_X-\mu_Y$.

7.2–10 **(a)** $\overline{d}=0.07875$;

(b) $[\,\overline{d}-1.7140.25492/\sqrt{24},\,\infty)=[-0.0104,\infty)$;

(c) not necessarily.

7.2–12 **(a)** $\overline{x}=136.61$, $\overline{y}=134.87$, $s_x^2=3.2972$, $s_y^2=1.0957$;

(b) Using Welch with $r=18$ degrees of freedom, the 95% confidence interval is [0.436, 3.041]. Assuming equal variances with $r=20$ degrees of freedom, the 95% confidence interval is [0.382, 3.095].

(c) The five-number summary for the X observations is 133.30, 135.625, 136.95, 137.80, 139.40. The five-number summary for the Y observations is 132.70, 134.15, 134.95, 135.825, 136.00.

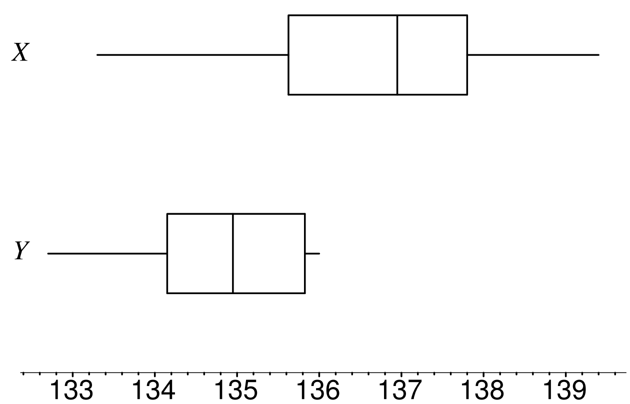


Figure 7.2–12: Hardness of hot (X) and cold (Y) water

(d) The mean for hot seems to be larger than the mean for cold.

7.2–14

$$\begin{aligned} P\left(c \le \frac{S_Y^2/\sigma_Y^2}{S_X^2/\sigma_X^2} \le d\right) &= 1-\alpha \\ P\left(c\frac{S_X^2}{S_Y^2} \le \frac{\sigma_X^2}{\sigma_Y^2} \le d\frac{S_X^2}{S_Y^2}\right) &= 1-\alpha, \end{aligned}$$

where $c = F_{1-\alpha/2}(r_1, r_2) = 1/F_{\alpha/2}(r_2, r_1)$ and $d = F_{\alpha/2}(r_1, r_2)$. With $r_1 = 9-1 = 8$, $r_2 = 13-1 = 12$, $12s_x^2 = 128.41$, and $8s_y^2 = 36.72$, a 98% confidence interval for σ_X^2/σ_Y^2 is

$$\left[\frac{1}{5.67}\frac{10.7008}{4.59}, 4.50\frac{10.7008}{4.59}\right] = [0.41, 10.49].$$

A 98% confidence interval for σ_X/σ_Y is

$$[\sqrt{0.41}, \sqrt{10.49}] = [0.64, 3.24].$$

7.3 Confidence Intervals For Proportions

7.3–2 **(a)** $\widehat{p} = \dfrac{142}{200} = 0.71;$

(b) $\left[0.71 - 1.645\sqrt{\dfrac{(0.71)(0.29)}{200}},\ 0.71 + 1.645\sqrt{\dfrac{(0.71)(0.29)}{200}}\ \right] = [0.657,\ 0.763];$

(c) $\dfrac{0.71 + 1.645^2/400 \pm 1.645\sqrt{0.71(0.29)/200 + 1.645^2/(4\cdot 200^2)}}{1 + 1.645^2/200} = [0.655,\ 0.760];$

(d) $\widetilde{p} = \dfrac{142+2}{200+4} = \dfrac{12}{17} = 0.7059;$

$$\left[\frac{12}{17} - 1.645\sqrt{\frac{(12/17)(5/17)}{204}},\ \frac{12}{17} - 1.645\sqrt{\frac{(12/17)(5/17)}{204}}\right] = [0.653, 0.758];$$

(e) $\left[0.71 - 1.282\sqrt{\dfrac{(0.71)(0.29)}{200}},\ 1\right] = [0.669,\ 1].$

7.3–4 $\left[0.70 - 1.96\sqrt{\dfrac{(0.70)(0.30)}{1234}},\ 0.70 + 1.96\sqrt{\dfrac{(0.70)(0.30)}{1234}}\ \right] = [0.674,\ 0.726].$

7.3–6 $\left[0.26 - 2.326\sqrt{\dfrac{(0.26)(0.74)}{5757}},\ 0.26 + 2.326\sqrt{\dfrac{(0.26)(0.74)}{5757}}\ \right] = [0.247,\ 0.273].$

7.3–8 **(a)** $\widehat{p} = \dfrac{388}{1022} = 0.3796;$

(b) $0.3796 \pm 1.645\sqrt{\dfrac{(0.3796)(0.6204)}{1022}}$ or $[0.3546,\ 0.4046].$

7.3–10 **(a)** $\widehat{p}_1 = 28/194 = 0.144;$
(b) $0.144 \pm 1.96\sqrt{(0.144)(0.856)/194}$ or $[0.095, 0.193];$
(c) $\widehat{p}_1 - \widehat{p}_2 = 28/194 - 11/162 = 0.076;$
(d) $\left[0.076 - 1.645\sqrt{\dfrac{(0.144)(0.856)}{194} + \dfrac{(0.068)(0.932)}{162}},\ 1\right]$ or $[0.044,\ 1].$

7.3–12 **(a)** $\widehat{p}_A = 170/460 = 0.37,\quad \widehat{p}_B = 141/440 = 0.32,$

$0.37 - 0.32 \pm 1.96\sqrt{\dfrac{(0.37)(0.63)}{460} + \dfrac{(0.32)(0.68)}{440}}$ or $[-0.012,\ 0.112];$

(b) yes, the interval includes zero.

7.4 Sample Size

7.4–2 $n = \dfrac{(1.96)^2(169)}{(1.5)^2} = 288.5$ so the sample size needed is 289.

7.4–4 $n = \dfrac{(1.96)^2(34.9)}{(0.5)^2} = 537$, rounded up to the nearest integer.

7.4–6 $n = \dfrac{(1.96)^2(33.7)^2}{5^2} = 175$, rounded up to the nearest integer.

7.4–8 $n = \dfrac{(1.645)^2(0.394)(0.606)}{(0.04)^2} = 404$, rounded up to the nearest integer.

7.4–10 $n = \dfrac{(1.645)^2(0.80)(0.20)}{(0.03)^2} = 482$, rounded up to the nearest integer.

7.4–12 $m = \dfrac{(1.96)^2(0.5)(0.5)}{(0.04)^2} = 601$, rounded up to the nearest integer.

(a) $n = \dfrac{601}{1 + 600/1500} = 430;$

(b) $n = \dfrac{601}{1 + 600/15{,}000} = 578;$

(c) $n = \dfrac{601}{1 + 600/25{,}000} = 587.$

7.4–14 For the difference of two proportions with equal sample sizes

$$\varepsilon = z_{\alpha/2}\sqrt{\frac{p_1^*(1-p_1^*)}{n} + \frac{p_2^*(1-p_2^*)}{n}}$$

or

$$n = \frac{z_{\alpha/2}^2[p_1^*(1-p_1^*) + p_2^*(1-p_2^*)]}{\varepsilon^2}.$$

For unknown p^*,

$$n = \frac{z_{\alpha/2}^2[0.25+0.25]}{\varepsilon^2} = \frac{z_{\alpha/2}^2}{2\varepsilon^2}.$$

So $n = \dfrac{1.282^2}{2(0.05)^2} = 329$, rounded up.

7.5 Distribution-Free Confidence Intervals for Percentiles

7.5–2 **(a)** $(y_3 = 5.4,\ y_{10} = 6.0)$ is a 96.14% confidence interval for the median, m.

(b) $(y_1 = 4.8,\ y_7 = 5.8)$;

$$\begin{aligned} P(Y_1 < \pi_{0.3} < Y_7) &= \sum_{k=1}^{6}\binom{12}{k}(0.3)^k(0.7)^{12-k} \\ &= 0.9614 - 0.0138 = 0.9476, \end{aligned}$$

using Table II with $n = 12$ and $p = 0.30$.

7.5–4 **(a)** $(y_4 = 80.28,\ y_{11} = 80.51)$ is a 94.26% confidence interval for m.

(b) $(y_6 = 80.32,\ y_{12} = 80.53)$;

$$\begin{aligned} \sum_{k=6}^{11}\binom{14}{k}(0.6)^k(0.4)^{14-k} &= \sum_{k=3}^{8}\binom{14}{k}(0.4)^k(0.6)^{14-k} \\ &= 0.9417 - 0.0398 = 0.9019. \end{aligned}$$

The interval is $(y_6 = 80.32, y_{12} = 80.53)$.

7.5–6 **(a)**

Stems	Leaves	Frequency
$20f$	5	1
$20s$	6 6 7 7	4
$20\bullet$	8 8 9 9 9	5
$21*$	0 0 0 0 0 1 1	7
$21t$	2 2 2 2 2 3 3 3 3 3 3 3 3	13
$21f$	4 4 4 4 4 4 4 5 5 5 5 5 5 5 5	15
$21s$	6 6 6 6 6 6 6 6 6 6 6 6 6 7 7 7 7 7 7 7 7 7 7	23
$21\bullet$	8 8 8 8 8 8 9 9 9 9	10
$22*$	0 0 0	3

(Multiply numbers by 10^{-1}.)

(b) **(i)**

$$\begin{aligned} q_1 &= \frac{y_{20}+y_{21}}{2} = \frac{21.2+21.2}{2} = 21.2 \\ q_2 &= m = y_{41} = 21.5 \\ q_3 &= \frac{y_{61}+7_{62}}{2} = \frac{21.7+21.7}{2} = 21.7. \end{aligned}$$

(ii) $\pi_{0.60} = 0.8y_{49} + 0.2y_{50} = 21.6.$

(iii) $\pi_{0.15} = 0.7y_{12} + 0.3y_{13} = 21.0.$

(c) **(i)** We first find i and j so that $P(Y_i < \pi_{0.25} < Y_j) \approx 0.95$. Let the distribution of W be $b(81,\ 0.25)$. Then

$$\begin{aligned} P(Y_i < \pi_{0.25} < Y_j) &= P(i \le W \le j-1) \\ &\approx P\left(\frac{i-0.5-20.25}{\sqrt{15.1875}} \le Z \le \frac{j-1+0.5-20.25}{\sqrt{15.1875}}\right). \end{aligned}$$

If we let

$$\frac{i-20.75}{\sqrt{15.1875}} = -1.96 \quad \text{and} \quad \frac{j-20.75}{\sqrt{15.1875}} = 1.96$$

we find that $i \approx 13$ and $j \approx 28$. Furthermore $P(13 \le W \le 28-1) \approx 0.9453$. Also note that the point estimate of $\pi_{0.25}$,

$$\widetilde{\pi}_{0.25} = (y_{20} + y_{21})/2$$

falls near the center of this interval. So a 94.53% confidence interval for $\pi_{0.25}$ is $(y_{13} = 21.0,\ y_{28} = 21.3)$.

(ii) Let the distribution of W be $b(81,\ 0.5)$. Then

$$\begin{aligned} P(Y_i < \pi_{0.5} < Y_{82-i}) &= P(i \le W \le 81-i) \\ &\approx P\left(\frac{i-0.5-40.5}{\sqrt{20.25}} \le Z \le \frac{81-i+0.5-40.5}{\sqrt{20.25}}\right). \end{aligned}$$

If

$$\frac{i-41}{4.5} = -1.96,$$

then $i = 32.18$ so let $i = 32$. Also

$$\frac{81-i-40}{4.5} = 1.96$$

implies that $i = 32$. Furthermore

$$P(Y_{32} < \pi_{0.5} < Y_{50}) = P(32 \le W \le 49) \approx 0.9544.$$

So an approximate 95.44% confidence interval for $\pi_{0.5}$ is $(y_{32} = 21.4,\ y_{50} = 21.6)$.

(iii) Similar to part (a), $P(Y_{54} < \pi_{0.75} < Y_{69}) \approx 0.9453$. Thus a 94.53% confidence interval for $\pi_{0.75}$ is $(y_{54} = 21.6,\ y_{69} = 21.8)$.

7.5–8 A 95.86% confidence interval for m is $(y_6 = 14.60,\ y_{15} = 16.20)$.

7.5–10 **(a)**

Stems	Leaves	Frequency
3	80	1
4	74	1
5	20 51 73 73 92	5
6	01 31 32 52 57 58 71 74 84 92 95	11
7	08 22 36 42 46 57 70 80	8
8	03 11 49 51 57 71 82 92 93 93	10
9	33 40 61	3
10	07 09 10 30 31 40 58 75	8
11	16 38 41 43 51 55 66	7
12	10 22 78	3
13	34 44 50	3

(b) A point estimate for the median is $\widetilde{m} = (y_{30} + y_{31})/2 = (8.51 + 8.57)/2 = 8.54$.

(c) Let the distribution of W be $b(60,\ 0.5)$. Then

$$\begin{aligned} P(Y_i < \pi_{0.5} < Y_{61-i}) &= P(i \le W \le 60-i) \\ &\approx P\left(\frac{i-0.5-30}{\sqrt{15}} \le Z \le \frac{60-i+0.5-30}{\sqrt{15}}\right). \end{aligned}$$

If

$$\frac{i-30.5}{\sqrt{15}} = -1.96$$

then $i \approx 23$. So

$$P(Y_{23} < \pi_{0.5} < Y_{38}) = P(23 \le W \le 37) \approx 0.9472.$$

So an approximate 94.72% confidence interval for $\pi_{0.5}$ is

$$(y_{23} = 7.46,\ y_{38} = 9.40).$$

(d) $\widetilde{\pi}_{0.40} = y_{24} + 0.4(y_{25} - y_{24}) = 7.57 + 0.4(7.70 - 7.57) = 7.622.$

(e) Let the distribution of W be $b(60,\ 0.40)$ then

$$\begin{aligned} P(Y_i < \pi_{0.40} < Y_j) &= P(i \le W \le j-1) \\ &\approx P\left(\frac{i-0.5-24}{\sqrt{14.4}} \le Z \le \frac{j-1+0.5-24}{\sqrt{14.4}}\right). \end{aligned}$$

If we let $\dfrac{i-24.5}{\sqrt{14.4}} = -1.645$ and $\dfrac{j-24.5}{\sqrt{14.4}} = 1.645$ then $i \approx 18$ and $j \approx 31$. Also $P(18 \le W \le 31-1) = 0.9133$. So an approximate 91.33% confidence interval for $\pi_{0.4}$ is $(y_{18} = 6.95,\ y_{31} = 8.57)$.

7.5–12 **(a)** $P(Y_7 < \pi_{0.70}) = \displaystyle\sum_{k=7}^{8} \binom{8}{k} (0.7)^k (0.3)^{8-k} = 0.2553;$

(b) $P(Y_5 < \pi_{0.70} < Y_8) = \displaystyle\sum_{k=5}^{7} \binom{8}{k} (0.7)^k (0.3)^{8-k} = 0.7483.$

7.6 More Regression

7.6–2 The mean of the observations is

$$\overline{Y}^{*} = \frac{1}{m}\sum_{i=1}^{m} Y_i^{*},$$

where $Y_1^*, Y_2^*, \ldots Y_m^*$ are the future observations at $X = x^*$. Since $\overline{Y}^{*}$ is a linear combination of independent and normally distributed random variables, it is normally distributed with mean

$$\begin{aligned} E\left(\frac{1}{m}\sum_{i=1}^{m} Y_i^*\right) &= \frac{1}{m}\sum_{i=1}^{m} E(Y_i^*) \\ &= \frac{1}{m}\sum_{i=1}^{m} [\alpha + \beta(x^* - \overline{x})] \\ &= \alpha + \beta(x^* - \overline{x}) \end{aligned}$$

and variance

$$\begin{aligned}\operatorname{Var}\left(\frac{1}{m}\sum_{i=1}^{m} Y_i^*\right) &= \frac{1}{m^2}\sum_{i=1}^{m}\operatorname{Var}(Y_i^*)\\ &= \frac{1}{m^2}\sum_{i=1}^{m}\sigma^2 \;=\; \frac{\sigma^2}{m}.\end{aligned}$$

Furthermore,

$$\overline{Y}^* - \widehat{\alpha} - \widehat{\beta}(x^* - \overline{x})$$

is a linear combination of independent and normally distributed random variables, so it is normally distributed with mean

$$E[\overline{Y}^* - \widehat{\alpha} - \widehat{\beta}(x^* - \overline{x})] = \alpha + \beta(x^* - \overline{x}) - \alpha - \beta(x^* - \overline{x}) = 0$$

and variance

$$\frac{\sigma^2}{m} + \frac{\sigma^2}{n} + \frac{(x^* - \overline{x})^2\sigma^2}{\sum(x_i - \overline{x})^2}.$$

Since $\overline{Y}^*$, $\widehat{\alpha}$, and $\widehat{\beta}$ are independent of $\widehat{\sigma^2}$, we can form the t-statistic

$$T = \frac{\dfrac{\overline{Y}^* - \widehat{\alpha} - \widehat{\beta}(x^* - \overline{x})}{\sigma\sqrt{\dfrac{1}{m} + \dfrac{1}{n} + \dfrac{(x^* - \overline{x})^2}{\sum(x_i - \overline{x})^2}}}}{\sqrt{\dfrac{n\widehat{\sigma^2}}{n-2}}} \sim t(n-2).$$

We can then obtain the following expression for the endpoints of the desired prediction interval:

$$\widehat{\alpha} + \widehat{\beta}(x^* - \overline{x}) \pm h t_{\alpha/2}(n-2)$$

where

$$h = \widehat{\sigma}\sqrt{\frac{n}{n-2}}\sqrt{\frac{1}{m} + \frac{1}{n} + \frac{(x^* - \overline{x})^2}{\sum(x_i - \overline{x})^2}}.$$

7.6–4 **(a)** In Exercise 6.5–4 we found that

$$\widehat{\beta} = 4.64/5.04, \qquad n\widehat{\sigma^2} = 1.84924, \qquad \sum_{i=1}^{10}(x_i - \overline{x})^2 = 5.04.$$

So the endpoints for the confidence interval are given by

$$2.73 + \frac{4.64}{5.04}(x - 2.90) \;\pm\; 2.306\sqrt{\frac{1.8493}{8}}\sqrt{\frac{1}{10} + \frac{(x - 2.90)^2}{5.04}},$$

$x = 2:$ $[1.335,\ 2.468]$,

$x = 3:$ $[2.468,\ 3.176]$,

$x = 4:$ $[3.096,\ 4.389]$.

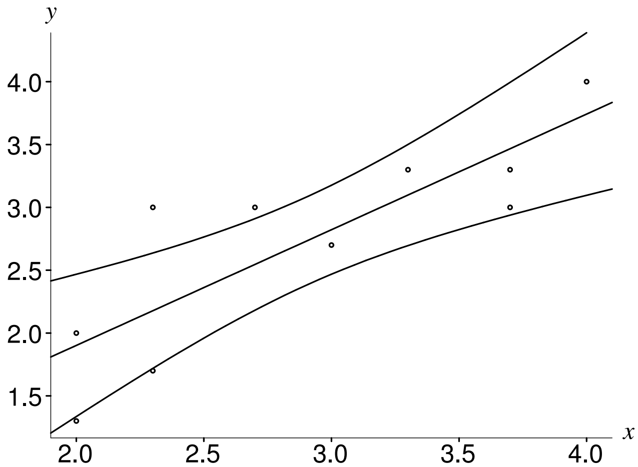
(b) The endpoints for the prediction interval are given by

$$2.73 + \frac{4.64}{5.04}(x - 2.90) \;\pm\; 2.306\sqrt{\frac{1.8493}{8}}\sqrt{1 + \frac{1}{10} + \frac{(x - 2.90)^2}{5.04}},$$

$x = 2:$ $[0.657,\ 3.146]$,

$x = 3:$ $[1.658,\ 3.986]$,

$x = 4:$ $[2.459,\ 5.026]$.



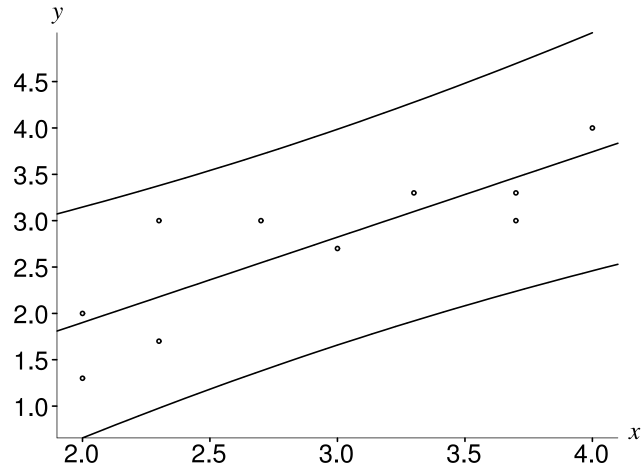


Figure 7.6–4: A 95% confidence interval for $\mu(x)$ and a 95% prediction band for Y

7.6–6 **(a)** For these data,

$$\sum_{i=1}^{10} x_i = 55, \quad \sum_{i=1}^{10} y_i = 9811, \quad \sum_{i=1}^{10} x_i^2 = 385,$$

$$\sum_{i=1}^{10} x_i y_i = 65{,}550, \quad \sum_{i=1}^{10} y_i^2 = 11{,}280{,}031.$$

Thus $\widehat{\alpha} = 9811/10 = 981.1$ and

$$\widehat{\beta} = \frac{65{,}550 - (55)(9811)/10}{385 - (55)^2/10} = \frac{11589.5}{82.5} = 140.4788.$$

The least squares regression line is

$$\widehat{y} = 981.1 + 140.4788(x - 5.5) = 208.467 + 140.479x.$$

(b)

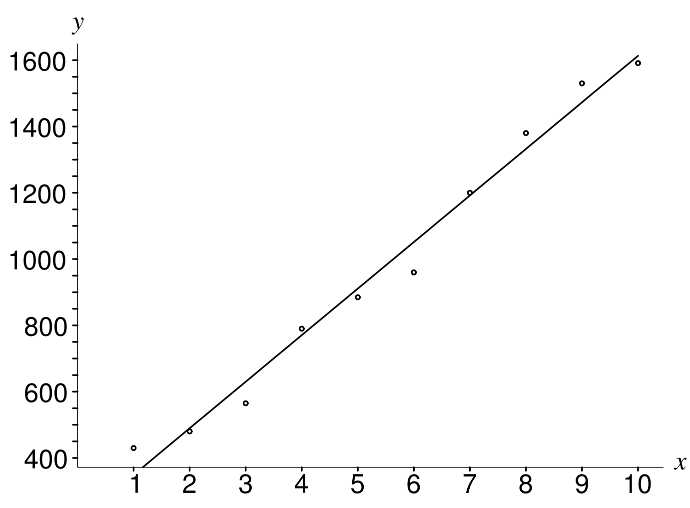


Figure 7.6–6: Number of programs (y) vs. year (x)

(c) 1753.733 ± 160.368 or $[1593.365,\ 1914.101]$.

7.6–8 Let $K(\beta_1, \beta_2, \beta_3) = \sum_{i=1}^{n} (y_i - \beta_1 - \beta_2 x_{1i} - \beta_3 x_{2i})^2$. Then

$$\begin{aligned}
\frac{\partial K}{\partial \beta_1} &= \sum_{i=1}^{n} 2(y_i - \beta_1 - \beta_2 x_{1i} - \beta_3 x_{2i})(-1) = 0;\\
\frac{\partial K}{\partial \beta_2} &= \sum_{i=1}^{n} 2(y_i - \beta_1 - \beta_2 x_{1i} - \beta_3 x_{2i})(-x_{1i}) = 0;\\
\frac{\partial K}{\partial \beta_3} &= \sum_{i=1}^{n} 2(y_i - \beta_1 - \beta_2 x_{1i} - \beta_3 x_{2i})(-x_{2i}) = 0.
\end{aligned}$$

Thus, we must solve simultaneously the three equations

$$n\beta_1 + \left(\sum_{i=1}^{n} x_{1i}\right)\beta_2 + \left(\sum_{i=1}^{n} x_{2i}\right)\beta_3 = \sum_{i=1}^{n} y_i$$

$$\left(\sum_{i=1}^{n} x_{1i}\right)\beta_1 + \left(\sum_{i=1}^{n} x_{1i}^2\right)\beta_2 + \left(\sum_{i=1}^{n} x_{1i}x_{2i}\right)\beta_3 = \sum_{i=1}^{n} x_{1i}y_i$$

$$\left(\sum_{i=1}^{n} x_{2i}\right)\beta_1 + \left(\sum_{i=1}^{n} x_{1i}x_{2i}\right)\beta_2 + \left(\sum_{i=1}^{n} x_{2i}^2\right)\beta_3 = \sum_{i=1}^{n} x_{2i}y_i.$$

We have

$$\begin{aligned}
12\beta_1 + 4\beta_2 + 4\beta_3 &= 23\\
4\beta_1 + 26\beta_2 + 5\beta_3 &= 75\\
4\beta_1 + 5\beta_2 + 22\beta_3 &= 37
\end{aligned}$$

so that

$$\widehat{\beta}_1 = \frac{4373}{5956} = 0.734, \qquad \widehat{\beta}_2 = \frac{3852}{1489} = 2.587, \qquad \widehat{\beta}_3 = \frac{1430}{1489} = 0.960.$$

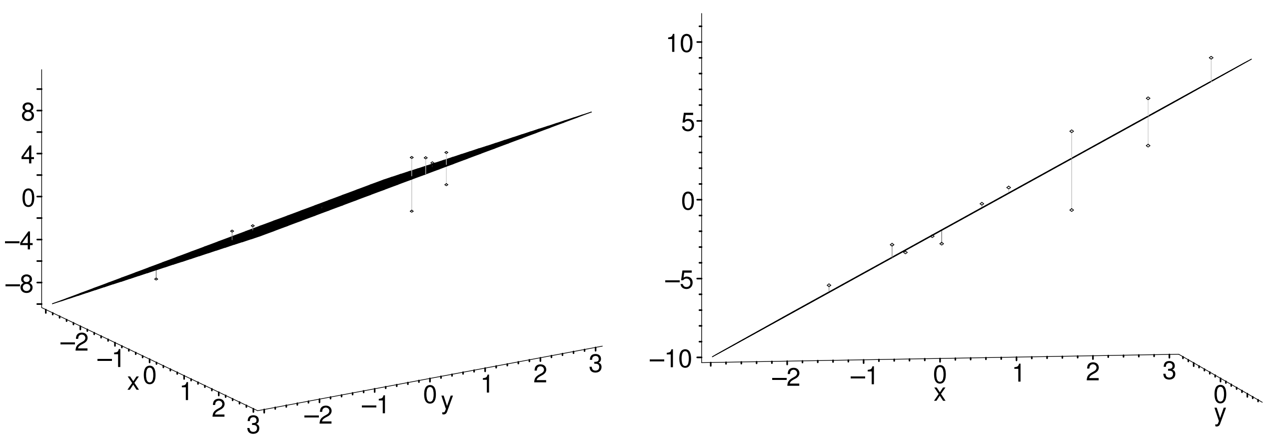


Figure 7.6–8: Two views of the points and the regression plane

7.6–10 **(c)** and **(d)**

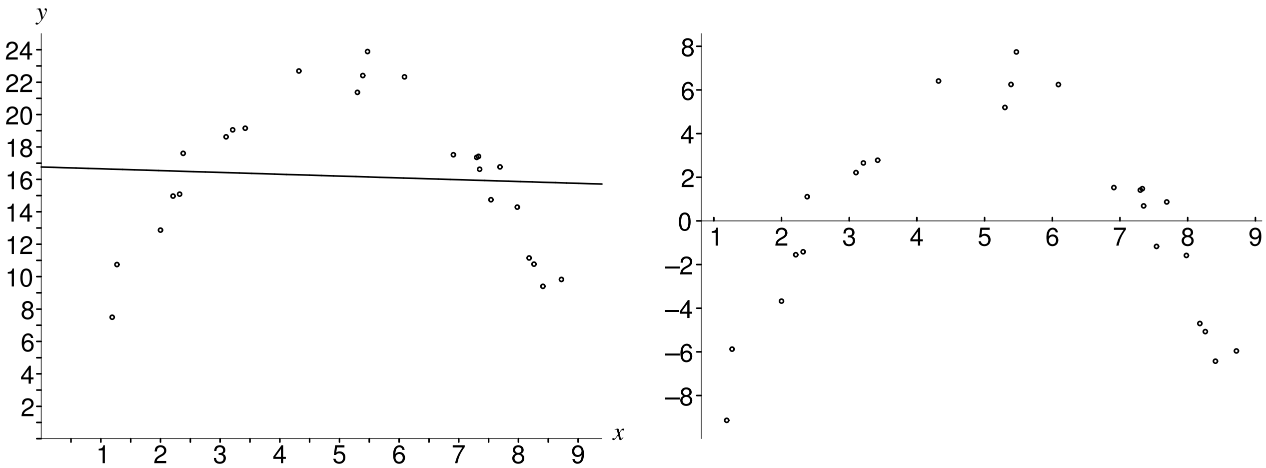


Figure 7.6–10: (y) versus (x) with linear regression line and residual plot

(e) Linear regression is not appropriate. Finding the least-squares quadratic regression line using the raw data yields $\widehat{y} = -1.88 + 9.86x - 0.995x^2$.

(f) and **(g)**

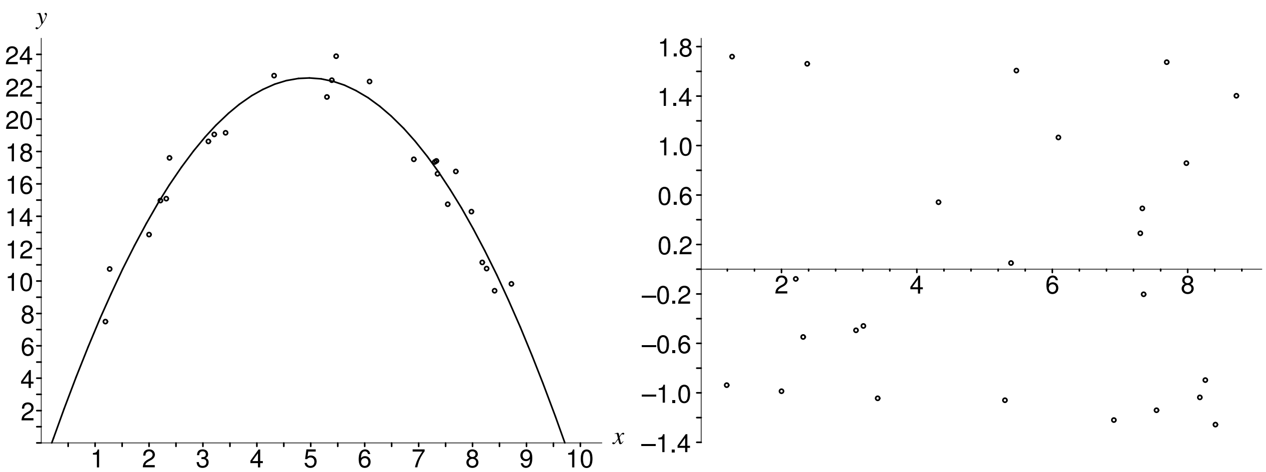


Figure 7.6–10: (y) versus (x) with quadratic regression curve and residual plot

7.6–12 The x is part of the exponent of e. However

$$\ln \mu(x) = \ln \beta_1 + \beta_2 x$$

is linear in x.

7.6–14 Recall that

$$T = T_2 = \frac{\dfrac{\sqrt{n}(\widehat{\alpha} - \alpha)}{\sigma}}{\sqrt{\dfrac{n\widehat{\sigma^2}}{\sigma^2(n-2)}}} = \frac{\widehat{\alpha} - \alpha}{\sqrt{\dfrac{\widehat{\sigma^2}}{n-2}}}$$

has a t distribution with $n-2$ degrees of freedom. Thus

$$
\begin{aligned}
P\left[-t_{\gamma/2}(n-2) \le \frac{\widehat{\alpha}-\alpha}{\sqrt{\widehat{\sigma^2}/(n-2)}} \le t_{\gamma/2}(n-2)\right] &= 1-\gamma \\
P\left[-\widehat{\alpha}-t_{\gamma/2}\sqrt{\frac{\widehat{\sigma^2}}{n-2}} \le -\alpha \le -\widehat{\alpha}+t_{\gamma/2}\sqrt{\frac{\widehat{\sigma^2}}{n-2}}\,\right] &= 1-\gamma \\
P\left[\widehat{\alpha}-t_{\gamma/2}\sqrt{\frac{\widehat{\sigma^2}}{n-2}} \le \alpha \le \widehat{\alpha}+t_{\gamma/2}\sqrt{\frac{\widehat{\sigma^2}}{n-2}}\,\right] &= 1-\gamma.
\end{aligned}
$$

It follows that the endpoints for a $100(1-\gamma)\%$ confidence interval for α are

$$\widehat{\alpha} \pm t_{\gamma/2}(n-2)\sqrt{\frac{\widehat{\sigma^2}}{n-2}}\,.$$

7.6–16 Recall that $\widehat{\alpha} = 2.73,\ \widehat{\beta} = 4.64/5.04,\ \widehat{\sigma^2} = 0.184925,\ n = 10$. The endpoints for the 95% confidence interval are

$$2.73 \pm 2.306\sqrt{\frac{0.184925}{8}} \quad \text{or} \quad [2.379,\ 3.081] \quad \text{for} \quad \alpha;$$

$$4.64/5.04 \pm 2.306\sqrt{\frac{1.84925}{8(5.04)}} \quad \text{or} \quad [0.4268,\ 1.4145] \quad \text{for} \quad \beta;$$

$$\left[\frac{1.84925}{17.54},\quad \frac{1.84925}{2.180}\right] = [0.105,\ 0.848] \quad \text{for} \quad \sigma^2.$$

7.6–18

$$
\begin{aligned}
\widehat{\beta} &= \frac{(1294)-(110)(121)/12}{(1234)-(110)^2/12} = \frac{184.833}{225.667} = 0.819; \\
\widehat{\alpha} &= \frac{121}{12} = 10.083; \\
\widehat{\sigma^2} &= \frac{39.5289}{12} = 3.294.
\end{aligned}
$$

The endpoints for 95% confidence intervals are

$$10.083 \pm 2.228\sqrt{\frac{3.294}{10}} \quad \text{or} \quad [8.804,\ 11.362] \quad \text{for} \quad \alpha;$$

$$0.819 \pm 2.228\sqrt{\frac{39.5289}{10(225.667)}} \quad \text{or} \quad [0.524,\ 1.114] \quad \text{for} \quad \beta;$$

$$\left[\frac{39.5289}{20.48},\quad \frac{39.5289}{3.247}\right] = [1.930,\ 12.174] \quad \text{for} \quad \sigma^2.$$

7.6–20 Recall that

$$\begin{aligned}\widehat{\alpha} &= \frac{395}{15} = 26.333,\\ \widehat{\beta} &= \frac{9292 - (346)(395)/15}{8338 - (346)^2/15} = \frac{180.667}{356.933} = 0.506,\\ \widehat{\sigma^2} &= \frac{211.8861}{15} = 14.126.\end{aligned}$$

The endpoints for 95% confidence intervals are

$$26.333 \pm 2.160\sqrt{\frac{14.126}{13}} \quad \text{or} \quad [24.081,\ 28.585] \quad \text{for} \quad \alpha;$$

$$0.506 \pm 2.160\sqrt{\frac{211.8861}{13(356.933)}} \quad \text{or} \quad [0.044,\ 0.968] \quad \text{for} \quad \beta;$$

$$\left[\frac{211.8861}{24.74},\ \ \frac{211.8861}{5.009}\right] = [8.566,\ 42.301] \quad \text{for} \quad \sigma^2.$$

7.7 Resampling Methods

7.7–2 **(a)**

```
> read ‘C:\\HTZ-CD\\Maple Examples\\stat.m‘:
  with(plots):
  read ‘C:\\HTZ-CD\\Maple Examples\\HistogramFill.txt‘:
  read ‘C:\\HTZ-CD\\Maple Examples\\ScatPlotCirc.txt‘:
  read ‘C:\\HTZ-CD\\Maple Examples\\Chapter_07.txt‘:
  XX := Exercise_7_7_2;
```

$$XX := [12.0, 9.4, 10.0, 13.5, 9.3, 10.1, 9.6, 9.3, 9.1, 9.2, 11.0, 9.1, 10.4, 9.1, 13.3, 10.6]$$

```
> Probs := [seq(1/16, k = 1 .. 16)]:
  XXPDF := zip((XX,Probs)-> (XX,Probs), XX, Probs):
> for k from 1 to 200 do
    X := DiscreteS(XXPDF, 16):
    Svar[k] := Variance(X):
  od:
  Svars := [seq(Svar[k], k = 1 .. 200)]:
> Mean(Svars);
```

$$1.972629584$$

```
> xtics := [seq(0.4*k, k = 1 .. 12)]:
  ytics := [seq(0.05*k, k = 1 .. 11)]:
  P1 := plot([[0,0],[0,0]], x = 0 .. 4.45, y = 0 .. 0.57,
  xtickmarks=xtics, ytickmarks=ytics, labels=[‘‘,‘‘]):
  P2 := HistogramFill(Svars,0 .. 4.4, 11):
  display({P1, P2});
```

The histogram is shown in Figure 7.7-2**(ab)**.

(b)
```
> theta := Mean(XX) - 9;
  for k from 1 to 200 do
    Y := ExponentialS(theta,21):
    Svary[k] := Variance(Y):
  od:
  Svarys := [seq(Svary[k], k = 1 .. 200)]:
```

$$\theta := 1.31250000$$

```
> Mean(Svarys);
```

$$1.747515570$$

```
> xtics := [seq(0.4*k, k = 1 .. 14)]:
  ytics := [seq(0.05*k, k = 1 .. 15)]:
  P3 := plot([[0,0],[0,0]], x = 0 .. 5.65, y = 0 .. 0.62,
  xtickmarks=xtics, ytickmarks=ytics, labels=[``,``]):
  P4 := HistogramFill(Svarys,0 .. 5.6, 14):
  display({P3, P4});
```

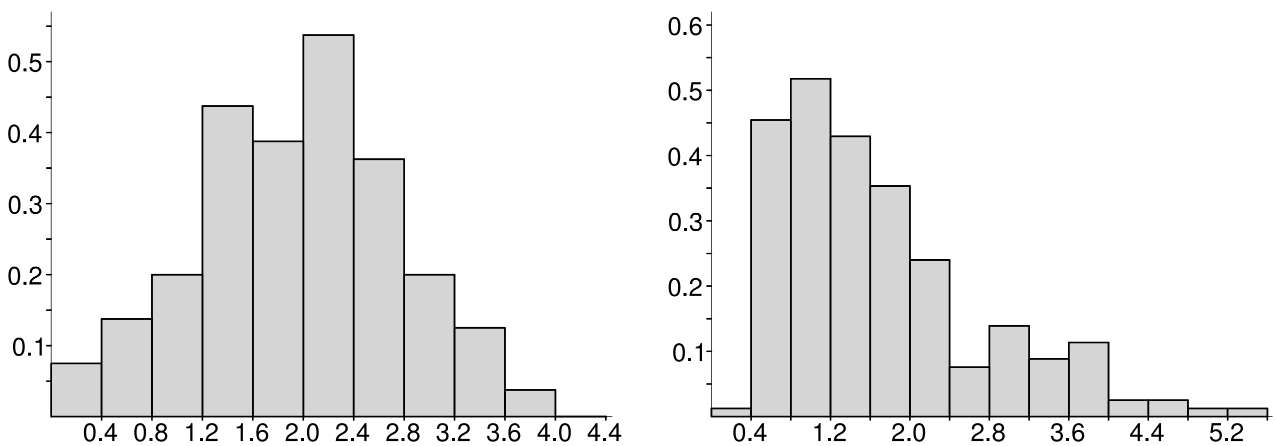


Figure 7.7–2: **(ab)** Histogram of S^2s: Resampling on Left, From Exponential on Right

```
> Svars := sort(Svars):
  Svarys := sort(Svarys):
> xtics := [seq(k*0.5, k = 1 .. 18)]:
  ytics := [seq(k*0.5, k = 1 .. 18)]:
  P5 := plot([[0,0],[5.5,5.5]], x = 0 .. 5.4, y = 0 .. 7.4, color=black,
  thickness=2, xtickmarks=xtics, ytickmarks=ytics, labels=[``,``]):
  P6 := ScatPlotCirc(Svars,Svarys):
  display({P5, P6});
```

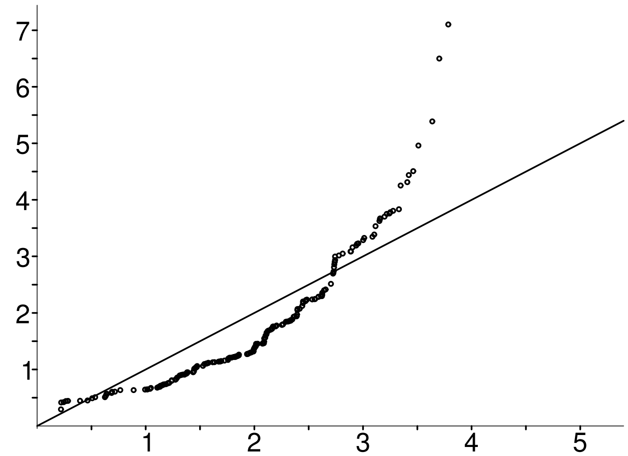


Figure 7.7–2: **(c)** q–q Plot of Exponential S^2s Versus Resampling S^2s

Note that the variance of the sample variances from the exponential distribution is greater than the variance of the sample variances from the resampling distribution.

7.7–4 **(a)**
```
> with(plots):
  read `C:\\HTZ-CD\\Maple Examples\\stat.m`:
  read `C:\\HTZ-CD\\Maple Examples\\HistogramFill.txt`:
  read `C:\\HTZ-CD\\Maple Examples\\Chapter_07.txt`:
  XX := Example_7_7_4;
  XX := [261, 269, 271, 274, 279, 280, 283, 284, 286, 287, 292, 293,
  296, 300, 304, 305, 313, 321, 322, 329, 341, 343, 356, 364, 391,
  417, 476];
  Probs := [seq(1/27, k = 1 .. 27)]:
  XXPDF := zip((XX,Probs)-> (XX,Probs), XX, Probs);
  XXPDF := [261, 1/27, 269, 1/27, 271, 1/27,
  274, 1/27, 279, 1/27, 280, 1/27, 283, 1/27, 284, 1/27, 286, 1/27,
  287, 1/27, 292, 1/27, 293, 1/27, 296, 1/27, 300, 1/27, 304, 1/27,
  305, 1/27, 313, 1/27, 321, 1/27, 322, 1/27, 329, 1/27, 341, 1/27,
  343, 1/27, 356, 1/27, 364, 1/27, 391, 1/27, 417, 1/27, 476, 1/27];

  for j from 1 to 500 do
    X := DiscreteS(XXPDF, 27):
    Y := sort(X):
    YY[j] := Y[7]:
  od:
  Y7 := [seq(YY[k], k = 1 .. 500)]:
  Y7sorted := sort(Y7):
  Y7sorted[1];
                                   271
  Y7sorted[500];
                                   304

  >xtics := [seq(273 + 5*k, k = 0 .. 7)]:
   ytics := [seq(0.01*k, k = 1 .. 7)]:
   P1 := plot([[265,0],[265,0]], x = 268 .. 306, y = 0 .. 0.0771,
   xtickmarks=xtics, ytickmarks=ytics, labels=[``,``]):
   P2 := HistogramFill(Y7,270.5 .. 304.5, 7):
   display({P1, P2});
```

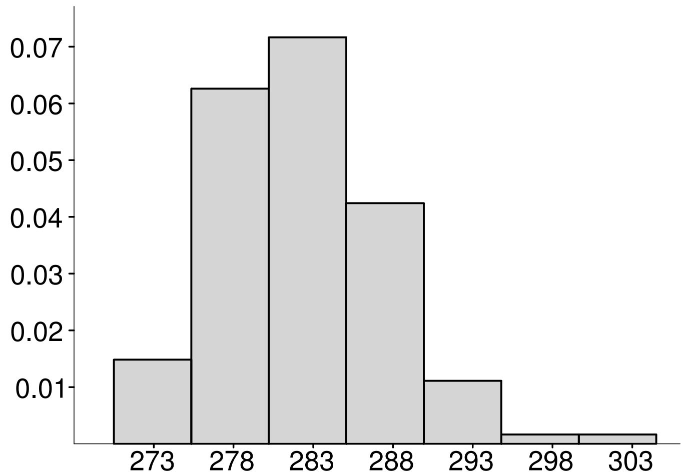


Figure 7.7–4: **(a)** A histogram of $N = 500$ seventh order statistics

(b)
```
>Here is an estimate of the first quartile.
 Mean(Y7);
```

$$\frac{141443}{500} = 282.886$$

(c)
```
>Here is an 82% confidence interval for the first quartile using, in
 an ordered Y7, the 44th number and the 206th number.
```

$$[279, 283]$$

(d) In Exercise 7.5-2, the 82% confidence interval for $\pi_{0.25}$ was $[274, 287]$.

7.7–6 **(a)**
```
> with(plots):
  read `C:\\HTZ-CD\\Maple Examples\\stat.m`:
  read `C:\\HTZ-CD\\Maple Examples\\ScatPlotPoint.txt`:
  read `C:\\HTZ-CD\\Maple Examples\\EmpCDF.txt`:
  read `C:\\HTZ-CD\\Maple Examples\\HistogramFill.txt`:
  read `C:\\HTZ-CD\\Maple Examples\\ScatPlotCirc.txt`:
  read `C:\\HTZ-CD\\Maple Examples\\Chapter_07.txt`:
  Pairs := Exercise_7_7_6;
  Pairs := [[2.500, 72], [4.467, 88], [2.333, 62], [5.000, 87],
  [1.683, 57], [4.500, 94], [4.500, 91], [2.083, 51], [4.367, 98],
  [1.583, 59], [4.500, 93], [4.550, 86], [1.733, 70], [2.150, 63],
  [4.400, 91], [3.983, 82], [1.767, 58], [4.317, 97], [1.917, 59],
  [4.583, 90], [1.833, 58], [4.767, 98], [1.917, 55], [4.433, 107],
  [1.750, 61], [4.583, 82], [3.767, 91], [1.833, 65], [4.817, 97],
  [1.900, 52], [4.517, 94], [2.000, 60], [4.650, 84], [1.817, 63],
  [4.917, 91], [4.000, 83], [4.317, 84], [2.133, 71], [4.783, 83],
  [4.217, 70], [4.733, 81], [2.000, 60], [4.717, 91], [1.917, 51],
  [4.233, 85], [1.567, 55], [4.567, 98], [2.133, 49], [4.500, 85],
  [1.717, 65], [4.783, 102], [1.850, 56], [4.583, 86],[1.733, 62]]:
> r := Correlation(Pairs);
```

$$r := .9087434803$$

```
> xtics := [seq(1.4 + 0.1*k, k = 0 .. 37)]:
  ytics := [seq(48 + 2*k, k = 0 .. 31)]:
```

```
P1 := plot([[1.35,47],[1.35,47]], x = 1.35 .. 5.15, y = 47 .. 109,
xtickmarks = xtics, ytickmarks=ytics, labels=[``,``]):
P2 := ScatPlotCirc(Pairs):
display({P1, P2});
```

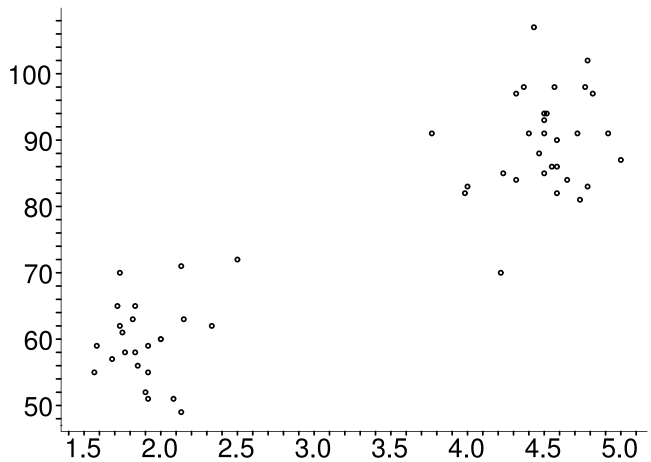


Figure 7.7–6: **(a)** Scatterplot of the 50 Pairs of Old Faithful Data

(b)
```
> Probs := [seq(1/54, k = 1 .. 54)]:
  EmpDist := zip((Pairs,Probs)-> (Pairs,Probs), Pairs, Probs):
> for k from 1 to 500 do
    Samp := DiscreteS(EmpDist, 54);
    RR[k] := Correlation(Samp):
  od:
  R := [seq(RR[k], k = 1 .. 500)]:
  rbar := Mean(R);
```

$$rbar := .9079354926$$

(c)
```
> xtics := [seq(0.8 + 0.01*k, k = 0 .. 20)]:
  ytics := [seq(k, k = 1 .. 25)]:
> P3 := plot([[0.79, 0],[0.79,0]], x = 0.79 .. 1.005,
  y = 0 .. 23.5, xtickmarks=xtics, ytickmarks=ytics, labels=[``,``]):
  P4 := HistogramFill(R, 0.8 .. 1, 20):
  display({P3, P4});
```

The histogram is plotted in Figure 7.7-6 **ce**.

(d) Now simulate a random sample of 500 correlation coefficients, each calculated from a sample of size 54 from a bivariate normal distribution with correlation coefficient $r = 0.9087434803$.

```
> for k from 1 to 500 do
    Samp := BivariateNormalS(0,1,0,1,r,54):
    RR[k] := Correlation(Samp):
  od:
  RBivNorm := [seq(RR[k], k = 1 .. 500)]:
  AverageR := Mean(RBivNorm);
```

$$AverageR := .9073168034$$

```
> P5 := plot([[0.79, 0],[0.79,0]], x = 0.79 .. 1.005,
  y = 0 .. 18.5, xtickmarks=xtics, ytickmarks=ytics, labels=[‘‘,‘‘]):
  P6 := HistogramFill(RBivNorm, 0.8 .. 1, 20):
  display({P5, P6});
```

(e)

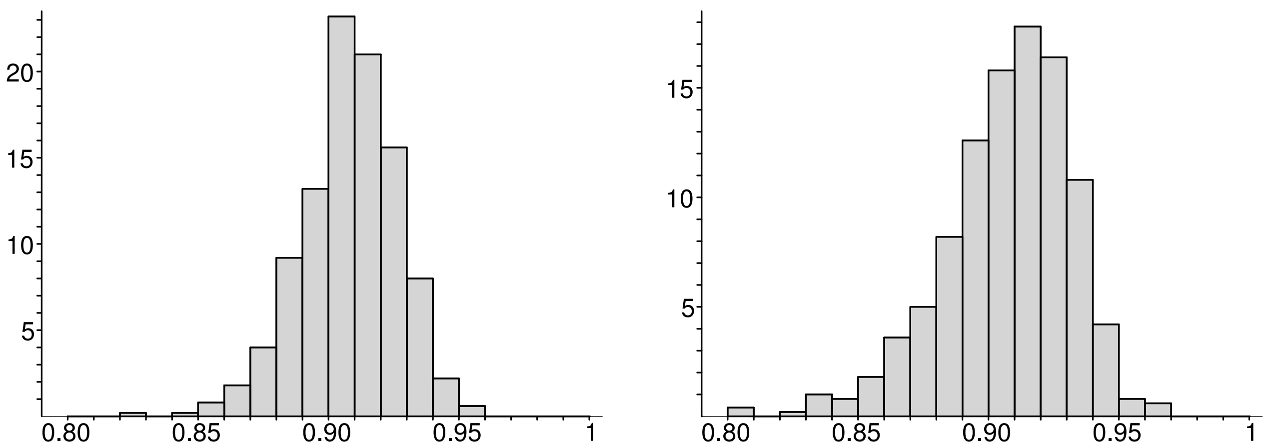


Figure 7.7–6: **(ce)** Histograms of Rs: From Resampling on Left, From Bivariate Normal on Right

(f)
```
> R := sort(R):
  RBivNorm := sort(RBivNorm):
  xtics := [seq(0.8 + 0.01*k, k = 0 .. 20)]:
  ytics := [seq(0.8 + 0.01*k, k = 0 .. 20)]:
  P7 := plot([[0.8, 0.8],[1,1]], x = 0.8 .. 0.97, y = 0.8 .. 0.97,
  color=black, thickness=2, labels=[``,``], xtickmarks=xtics,
  ytickmarks=ytics):
  P8 := ScatPlotCirc(R, RBivNorm):
  display({P7, P8});
```

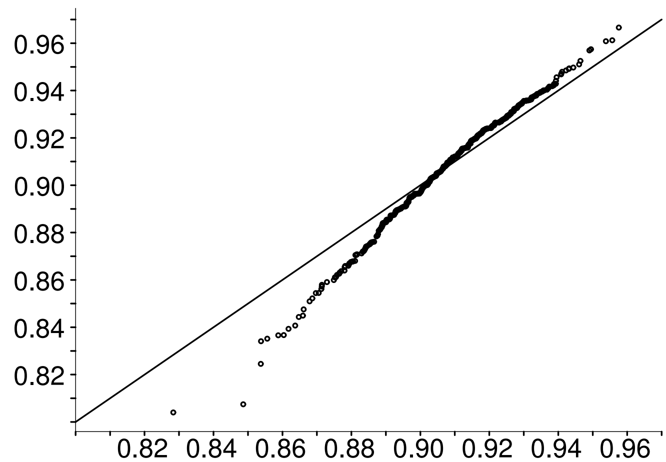


Figure 7.7–6: **(f)** q–q Plot of the Values of R from Bivariate Normal Versus from Resampling

```
> StDev(R);
  StDev(RBivNorm);
```

.01852854051

.02461716901

The means are about equal but the standard deviation of the values of R from the bivariate normal distribution is larger than that of the resampling distribution.

Chapter 8

Tests of Statistical Hypotheses

8.1 Tests About One Mean

8.1–2 **(a)** The critical region is

$$z = \frac{\overline{x} - 13.0}{0.2/\sqrt{n}} = \frac{\overline{x} - 13.0}{0.2/\sqrt{25}} \le -1.96;$$

(b) The observed value of z,

$$z = \frac{12.9 - 13.0}{0.04} = -2.5,$$

is less that -1.96 so we reject H_0.

(c) The p-value of this test is $P(Z \le -2.50) = 0.0062$.

8.1–4 **(a)** $|t| = \dfrac{|\,\overline{x} - 7.5\,|}{s/\sqrt{10}} \ge t_{0.025}(9) = 2.262$.

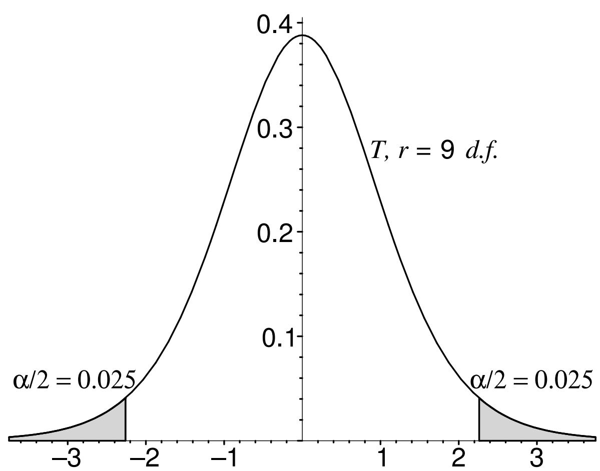


Figure 8.1–4: The critical region is $|\,t\,| \ge 2.262$

(b) $|t| = \dfrac{|\,7.55 - 7.5\,|}{0.1027/\sqrt{10}} = 1.54 < 2.262$, do not reject H_0.

(c) A 95% confidence interval for μ is

$$\left[7.55 - 2.262\left(\frac{0.1027}{\sqrt{10}}\right),\ 7.55 + 2.262\left(\frac{0.1027}{\sqrt{10}}\right)\right] = [7.48,\ 7.62].$$

Hence, $\mu = 7.50$ is contained in this interval. We could have obtained the same conclusion from our answer to part (b).

8.1–6 **(a)** H_0: $\mu = 3.4$;

(b) H_1: $\mu > 3.4$;

(c) $t = (\overline{x} - 3.4)/(s/3)$;

(d) $t \geq 1.860$;

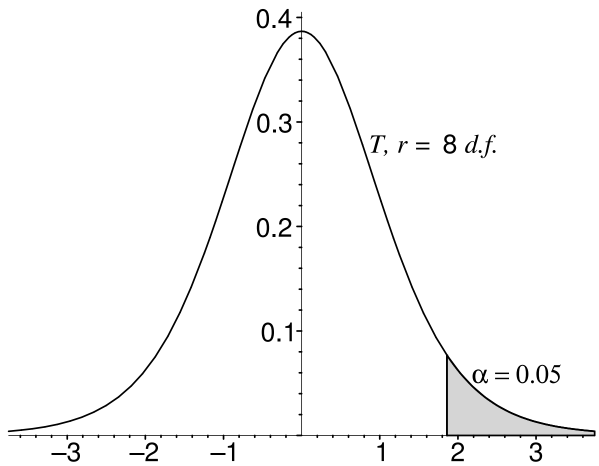


Figure 8.1–6: The critical region is $t \geq 1.860$

(e) $t = \dfrac{3.556 - 3.4}{0.167/3} = 2.802$;

(f) $2.802 > 1.860$, reject H_0;

(g) $0.01 < p\text{-value} < 0.025$, $p\text{-value} = 0.0116$.

8.1–8 **(a)** $|\,t\,| = \dfrac{|\,\overline{x} - 125\,|}{s/\sqrt{15}} \geq t_{0.025}(14) = 2.145$.

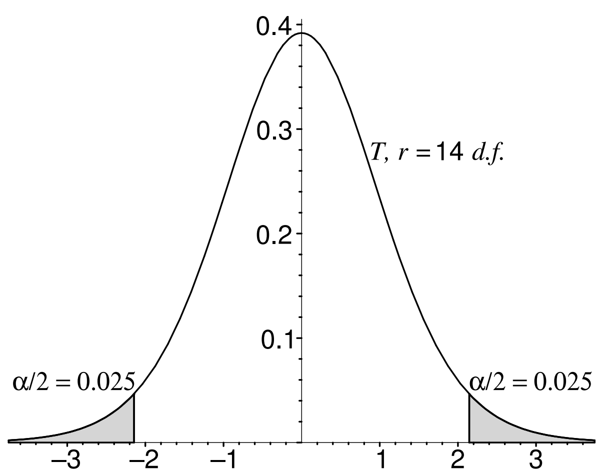


Figure 8.1–8: The critical region is $|\,t\,| \geq 2.145$

(b) $|\,t\,| = \dfrac{|\,127.667 - 125\,|}{9.597/\sqrt{15}} = 1.076 < 2.145$, do not reject H_0.

8.1–10 **(a)** The test statistic and critical region are given by

$$t = \frac{\overline{x} - 5.70}{s/\sqrt{8}} \geq 1.895.$$

(b) The observed value of the test statistic is

$$t = \frac{5.869 - 5.70}{0.19737/\sqrt{8}} = 2.42.$$

(c) The p-value is a little less than 0.025. Using Minitab, the p-value = 0.023.

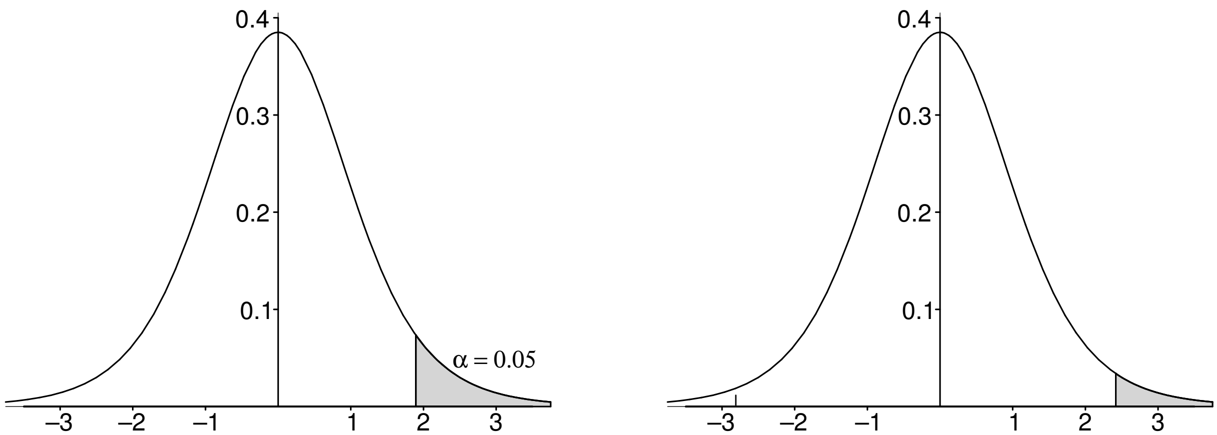


Figure 8.1–10: A $T(7)$ pdf showing the critical region on the left, p-value on the right

8.1–12 The critical region is

$$t = \frac{\overline{d} - 0}{s_d/\sqrt{17}} \geq 1.746 = t_{0.06}(16).$$

Since $\overline{d} = 4.765$ and $s_d = 9.087$, $t = 2.162 > 1.746$ and we reject H_0.

8.1–14 **(a)** $\dfrac{(23-1)s^2}{100} = \dfrac{715.44}{100} = 7.1544 < 10.98 = \chi^2_{0.975}(22)$, so she would reject H_0.

(b) Reject H_0 if $\dfrac{(23-1)s^2}{100} = \dfrac{715.44}{100} < \chi^2_{0.975}(22) = 10.98$ or if $\dfrac{(23-1)s^2}{100} > \chi^2_{0.025}(22) = 36.78$. So again she would reject H_0.

8.2 Tests of the Equality of Two Means

8.2–2 **(a)** $$t = \frac{\overline{x} - \overline{y}}{\sqrt{\dfrac{15s_x^2 + 12s_y^2}{27}\left(\dfrac{1}{16} + \dfrac{1}{13}\right)}} \geq t_{0.01}(27) = 2.473;$$

(b) $$t = \frac{415.16 - 347.40}{\sqrt{\dfrac{15(1356.75) + 12(692.21)}{27}\left(\dfrac{1}{16} + \dfrac{1}{13}\right)}} = 5.570 > 2.473, \text{ reject } H_0.$$

(c)
$$\begin{aligned} c &= \frac{1356.75}{1356.75 + 692.21} = 0.662, \\ \frac{1}{r} &= \frac{0.662^2}{15} + \frac{0.338^2}{12} = 0.0387, \\ r &= 25. \end{aligned}$$

The critical region is therefore $t \ge t_{0.01}(25) = 2.485$. Since $t = 5.570 > 2.485$, we again reject H_0.

8.2–4 **(a)** $t = \dfrac{\overline{x} - \overline{y}}{\sqrt{\dfrac{12s_x^2 + 15s_y^2}{27}\left(\dfrac{1}{13} + \dfrac{1}{16}\right)}} \le -t_{0.05}(27) = -1.703;$

(b) $t = \dfrac{72.9 - 81.7}{\sqrt{\dfrac{(12)(25.6)^2 + (15)(28.3)^2}{27}\left(\dfrac{1}{13} + \dfrac{1}{16}\right)}} = -0.869 > -1.703$, do not reject H_0;

(c) $0.10 < p\text{-value} < 0.25$;

8.2–6 **(a)** Assuming $\sigma_X^2 = \sigma_Y^2$,

$$|t| = \frac{|\,\overline{x} - \overline{y}\,|}{\sqrt{\dfrac{9s_x^2 + 9s_y^2}{18}\left(\dfrac{1}{10} + \dfrac{1}{10}\right)}} \ge t_{0.025}(18) = 2.101;$$

(b) $|-2.151\,| > 2.101$, reject H_0;

(c) $0.01 < p\text{-value} < 0.05$;

(d) Yes.

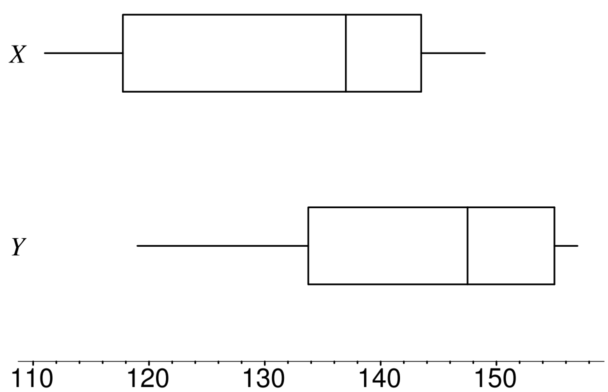


Figure 8.2–6: Box-and-whisker diagram for stud 3 (X) and stud 4 (Y) forces

8.2–8 $t = \dfrac{\overline{x} - \overline{y}}{\sqrt{\dfrac{24s_x^2 + 28s_y^2}{52}\left(\dfrac{1}{25} + \dfrac{1}{29}\right)}} = 3.402 > 2.326 = z_{0.01},$

reject $\mu_X = \mu_Y$.

8.2–10 $t = \dfrac{4.1633 - 5.1050}{\sqrt{\dfrac{11(0.91426) + 7(2.59149)}{18}}\sqrt{\dfrac{1}{12} + \dfrac{1}{8}}} = -1.648$. Since $-1.330 < -1.648 < -1.734$,

$0.05 < p\text{-value} < 0.10$. In fact, $p\text{-value} = 0.058$. We would reject H_0 at an $\alpha = 0.05$ significance level and fail to reject H_0 at an $\alpha = 0.10$ significance level.

8.2–12 **(a)** $\dfrac{\overline{y}-\overline{x}}{\sqrt{\dfrac{s_y^2}{30}+\dfrac{s_x^2}{30}}} > 1.96;$

(b) $8.98 > 1.96$, reject $\mu_X = \mu_Y$.

(c) Yes.

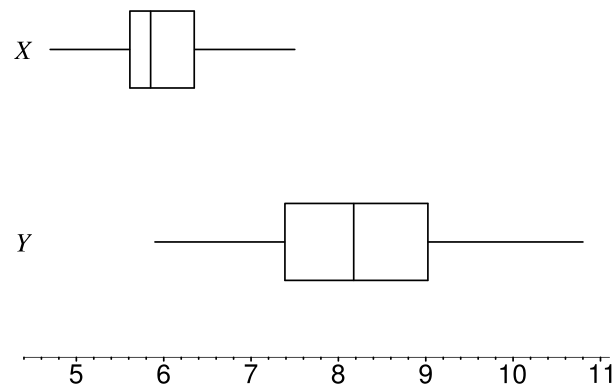


Figure 8.2–12: Lengths of male (X) and female (Y) green lynx spiders

8.2–14 **(a)** For these data, $\overline{x} = 5.9947$, $\overline{y} = 4.3921$, $s_x^2 = 6.0191$, $s_y^2 = 1.9776$. Using the number of degrees of freedom given by Equation 7.2-1 (Welch) we have that $r = \lfloor 28.68 \rfloor = 28$. We have

$$t = \frac{5.9947 - 4.3921}{\sqrt{6.0191/19 + 1.9776/19}} = 2.47 > 2.467 = t_{0.01}(28)$$

so we reject H_0.

(b)

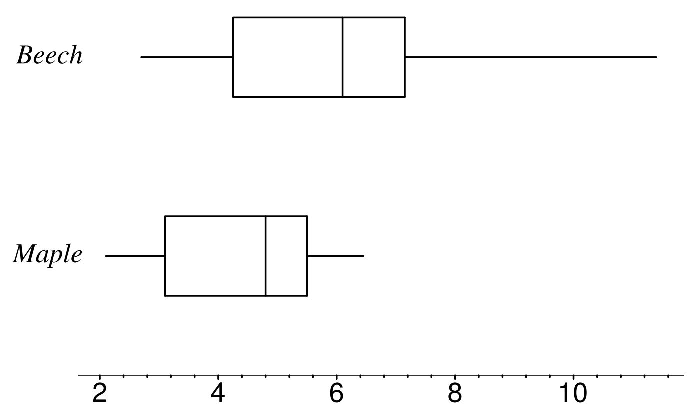


Figure 8.2–14: Tree dispersion distances in meters

8.2–16 **(a)** $F = \dfrac{s_x^2}{s_y^2} = \dfrac{12.9^2}{7.1^2} = 3.30 > 3.07 = F_{0.05}(9-1, 11-1)$, so we reject H_0.

(b) $z = \dfrac{93 - 132}{\sqrt{12.9^2/9 + 7.1^2/11}} = -8.12 < -1.645$ so H_0 is clearly rejected. A modification of this test would be to use Welch's t with 11 degrees of freedom. (See Equation 7.2-1.) This also leads to rejection of H_0.

8.3 Tests about Proportions

8.3–2 **(a)** $C = \{x : x = 0, 1, 2\}$;

(b)
$$\begin{aligned}\alpha &= P(X = 0, 1, 2;\ p = 0.6)\\ &= (0.4)^4 + 4(0.6)(0.4)^3 + 6(0.6)^2(0.4)^2 = 0.5248;\\ \beta &= P(X = 3, 4;\ p = 0.4)\\ &= 4(0.4)^3(0.6) + (0.4)^4 = 0.1792.\end{aligned}$$

OR

(a′) $C = \{x : x = 0, 1\}$;

(b′)
$$\begin{aligned}\alpha &= P(X = 0, 1;\ p = 0.6)\\ &= (0.4)^4 + 4(0.6)(0.4)^3 = 0.1792;\\ \beta &= P(X = 2, 3, 4;\ p = 0.4)\\ &= 6(0.4)^2(0.6)^2 + 4(0.4)^3(0.6) + (0.4)^4 = 0.5248.\end{aligned}$$

8.3–4 Using Table II in the Appendix,

(a) $\alpha = P(Y \geq 13;\ p = 0.40) = 1 - 0.8462 = 0.1538$;

(b)
$$\begin{aligned}\beta &= P(Y \leq 12;\ p = 0.60)\\ &= P(25 - Y \geq 25 - 12) \quad \text{where } 25 - Y \text{ is } b(25,\ 0.40)\\ &= 1 - 0.8462 = 0.1538.\end{aligned}$$

8.3–6 The value of the test statistic is

$$z = \frac{0.70 - 0.75}{\sqrt{(0.75)(0.25)/390}} = -2.280.$$

(a) Since $z = -2.280 < -1.645$, reject H_0.

(b) Since $z = -2.280 > -2.326$, do not reject H_0.

(c) p-value $\approx P(Z \leq -2.280) = 0.0113$. Note that $0.01 < p\text{-value} < 0.05$.

8.3–8 **(a)** H_0: $p = 0.14$; H_1: $p > 0.14$;

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

(c) $z = \dfrac{104/590 - 0.14}{\sqrt{(0.14)(0.86)/590}} = 2.539 > 2.326$

so H_0 is rejected and conclude that the campaign was successful.

8.3–10 **(a)** $z = \dfrac{y/n - 0.65}{\sqrt{(0.65)(0.35)/n}} \geq 1.96$;

(b) $z = \dfrac{414/600 - 0.65}{\sqrt{(0.65)(0.35)/600}} = 2.054 > 1.96$, reject H_0 at $\alpha = 0.025$.

(c) Since the p-value $\approx P(Z \geq 2.054) = 0.0200 < 0.0250$, reject H_0 at an $\alpha = 0.025$ significance level;

(d) A 95% one-sided confidence interval for p is

$$[0.69 - 1.645\sqrt{(0.69)(0.31)/600}\,, 1] = [0.659, 1].$$

8.3–12 **(a)** $|\,z\,| = \dfrac{|\,\widehat{p} - 0.20\,|}{\sqrt{(0.20)(0.80)/n}} \geq 1.96$;

(b) Only 5/54 for which $z = -1.973$ leads to rejection of H_0, so 5% reject H_0.

(c) 5%.

(d) 95%.

(e) $z = \dfrac{219/1124 - 0.20}{\sqrt{(0.20)(0.80)/1124}} = -0.43$, so fail to reject H_0.

8.3–14 **(a)** $z = \dfrac{\widehat{p}_1 - \widehat{p}_2}{\sqrt{\widehat{p}(1-\widehat{p})(1/n_1 + 1/n_2)}} \ge 1.645$;

(b) $z = \dfrac{0.15 - 0.11}{\sqrt{(0.1325)(0.8675)(1/900 + 1/700)}} = 2.341 > 1.645$, reject H_0.

(c) $z = 2.341 > 2.326$, reject H_0.

(d) The p-value $\approx P(Z \ge 2.341) = 0.0096$.

8.3–16 $z = \dfrac{204/300 - 0.73}{\sqrt{(0.73)(0.27)/300}} = \dfrac{-0.05}{0.02563} = -1.95$;

p-value $\approx P(Z < -1.95) = 0.0256 < \alpha = 0.05$ so we reject H_0. That is, the test indicates that there is progress.

8.4 The Wilcoxon Tests

8.4–2 In the following display, those observations that were negative are underlined.

$\|x\|$:	$\underline{1}$	$\underline{2}$	$\underline{2}$	$\underline{2}$	2	$\underline{3}$	$\underline{4}$	$\underline{4}$	$\underline{4}$	$\underline{5}$	$\underline{6}$	6
Ranks :	1	3.5	3.5	3.5	3.5	6	8	8	8	10	12	12

$\|x\|$:	6	7	7	8	11	12	13	$\underline{14}$	14	17	18	21
Ranks :	12	14.5	14.5	16	17	18	19	20.5	20.5	22	23	24

The value of the Wilcoxon statistic is

$$\begin{aligned} w &= -1 - 3.5 - 3.5 - 3.5 + 3.5 - 6 - 8 - 8 - 8 - 10 - 12 + 12 + 12 + \\ &\quad\ 14.5 + 14.5 + 16 + 17 + 18 + 19 - 20.5 + 20.5 + 22 + 23 + 24 \\ &= 132. \end{aligned}$$

For a one-sided alternative, the approximate p-value is, using the one-unit correction,

$$\begin{aligned} P(W \ge 132) &= P\left(\frac{W - 0}{\sqrt{24(25)(49)/6}} \ge \frac{131 - 0}{70} \right) \\ &\approx P(Z \ge 1.871) = 0.03064. \end{aligned}$$

For a two-sided alternative, p-value $= 2(0.03064) = 0.0613$.

8.4–4 In the following display, those observations that were negative are underlined.

$\|x\|$:	$\underline{0.0790}$	0.5901	$\underline{0.7757}$	$\underline{1.0962}$	$\underline{1.9415}$
Ranks :	1	2	3	4	5

$\|x\|$:	$\underline{3.0678}$	3.8545	$\underline{5.9848}$	9.3820	$\underline{74.0216}$
Ranks :	6	7	8	9	10

The value of the Wilcoxon statistic is

$$w = -1 + 2 - 3 - 4 - 5 - 6 + 7 - 8 + 9 - 10 = -19.$$

Since

$$|z| = \left| \frac{-19}{\sqrt{10(11)(21)/6}} \right| = 0.968 < 1.96,$$

we do not reject H_0.

8.4–6 **(a)** The critical region is given by

$$w \ge 1.645\sqrt{15(16)(31)/6} = 57.9.$$

(b) In the following display, those differences that were negative are underlined.

$$\begin{array}{rccccccccc} |x_i - 50|: & 2 & \underline{2} & 2.5 & 3 & 4 & \underline{4} & \underline{4.5} & \underline{6} & 7 \\ \text{Ranks}: & 1.5 & 1.5 & 3 & 4 & 5.5 & 5.5 & 7 & 8 & 9 \end{array}$$

$$\begin{array}{rcccccc} |x_i - 50|: & 7.5 & 8 & 8 & \underline{14.5} & 15.5 & 21 \\ \text{Ranks}: & 10 & 11.5 & 11.5 & 13 & 14 & 15 \end{array}$$

The value of the Wilcoxon statistic is

$$\begin{aligned} w &= 1.5 - 1.5 + 3 + 4 + 5.5 - 5.5 - 7 - 8 + 9 + 10 + 11.5 + 11.5 - 13 + 14 + 15 \\ &= 50. \end{aligned}$$

Since

$$z = \frac{50}{\sqrt{15(16)(31)/6}} = 1.420 < 1.645,$$

or since $w = 50 < 57.9$, we do not reject H_0.

(c) The approximate p-value is, using the one-unit correction,

$$\begin{aligned} p\text{-value} &= P(W \ge 50) \\ &\approx P\left(Z \ge \frac{49}{\sqrt{15(16)(31)/6}}\right) = P(Z \ge 1.3915) = 0.0820. \end{aligned}$$

8.4–8 The 24 ordered observations, with the x-values underlined and the ranks given under each observation are:

$$\begin{array}{rcccccc} & \underline{0.7794} & \underline{0.7546} & \underline{0.7565} & 0.7613 & \underline{0.7615} & \underline{0.7701} \\ \text{Ranks}: & 1 & 2 & 3 & 4 & 5 & 6 \end{array}$$

$$\begin{array}{rcccccc} & \underline{0.7712} & \underline{0.7719} & \underline{0.7719} & \underline{0.7720} & 0.7720 & 0.7731 \\ \text{Ranks}: & 7 & 8.5 & 8.5 & 10.5 & 10.5 & 12 \end{array}$$

$$\begin{array}{rcccccc} & \underline{0.7741} & \underline{0.7750} & 0.7750 & \underline{0.7776} & 0.7795 & 0.7811 \\ \text{Ranks}: & 13 & 14.5 & 14.5 & 16 & 17 & 18 \end{array}$$

$$\begin{array}{rcccccc} & 0.7815 & 0.7816 & 0.7851 & 0.7870 & 0.7876 & 0.7972 \\ \text{Ranks}: & 19 & 20 & 21 & 22 & 23 & 24 \end{array}$$

(a) The value of the Wilcoxon statistic is

$$\begin{aligned} w &= 4 + 10.5 + 12 + 14.5 + 17 + 18 + 19 + 20 + 21 + 22 + 23 + 24 \\ &= 205. \end{aligned}$$

Thus

$$p\text{-value} = P(W \ge 205) \approx P\left(Z \ge \frac{204.5 - 150}{\sqrt{12(12)(25)/12}}\right) = P(Z \ge 3.15) < 0.001$$

so that we clearly reject H_0.

(b)

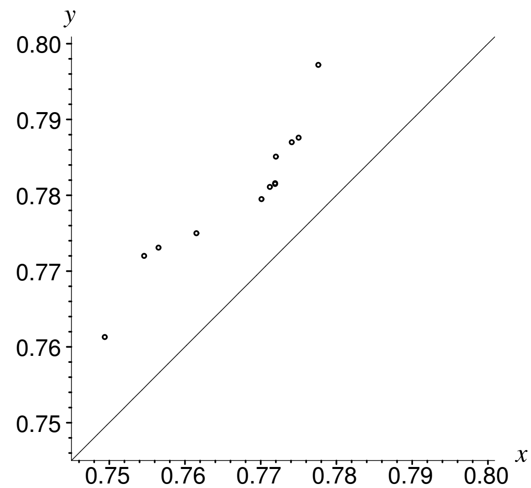


Figure 8.4–8: q-q plot of pill weights, (good, defective) = (x, y)

8.4–10 The ordered combined sample with the x observations underlined are:

	$\underline{67.4}$	$\underline{69.3}$	$\underline{72.7}$	73.1	$\underline{75.9}$	$\underline{77.2}$	$\underline{77.6}$	78.9	
Ranks:	1	2	3	4	5	6	7	8	
	82.5	$\underline{83.2}$	$\underline{83.3}$	$\underline{84.0}$	84.7	86.5	87.5		
Ranks:	9	10	11	12	13	14	15		
	$\underline{87.6}$	88.3	88.6	$\underline{90.2}$	$\underline{90.4}$	90.4	92.7	94.4	95.0
Ranks:	16	17	18	19	20.5	20.5	22	23	24

The value of the Wilcoxon statistic is

$$w = 4 + 8 + 9 + \cdots + 23 + 24 = 187.5.$$

Since

$$z = \frac{187.5 - 12(25)/2}{\sqrt{12(12)(25)/12}} = 2.165 > 1.645,$$

we reject H_0.

8.4–12 The ordered combined sample with the 48-passenger bus values underlined are:

	$\underline{104}$	184	196	197	248	$\underline{253}$	260	279
Ranks:	1	2	3	4	5	6	7	8
	$\underline{300}$	$\underline{308}$	$\underline{323}$	$\underline{331}$	355	386	393	$\underline{396}$
Ranks:	9	10	11	12	13	14	15	16
	$\underline{414}$	432	450	$\underline{452}$				
Ranks:	17	18	19	20				

The value of the Wilcoxon statistic is

$$w = 2 + 3 + 4 + 5 + 7 + 8 + 13 + 14 + 15 + 18 + 19 = 108.$$

Since

$$z = \frac{108 - 11(21)/2}{\sqrt{9(11)(21)/12}} = -0.570 > -1.645,$$

we do not reject H_0.

8.4–14 **(a)** Here is the two-sided stem-and-leaf display.

Group A leaves	Stems	Group B leaves
	0	9
7	1	2
	2	1 5 7
3	3	1 2 3 4
6 2	4	4
7 5 1 0	5	3
3 1	6	
1	7	

(b) Here is the ordered combined sample with the Group B values underlined:

	$\underline{9}$	$\underline{12}$	17	$\underline{21}$	$\underline{25}$	$\underline{27}$	$\underline{31}$	$\underline{32}$
Ranks :	1	2	3	4	5	6	7	8
	$\underline{33}$	33	$\underline{34}$	42	$\underline{44}$	46	50	51
Ranks :	9.5	9.5	11	12	13	14	15	16
	$\underline{53}$	55	57	61	63	71		
Ranks :	17	18	19	20	21	22		

The value of the Wilcoxon statistic is

$$w = 1 + 2 + 4 + 5 + 6 + 7 + 8 + 9.5 + 11 + 13 + 17 = 83.5.$$

Since

$$z = \frac{83.5 - 126.5}{\sqrt{11(11)(23)/12}} = \frac{-43}{15.2288} = -2.83 < -2.576 = z_{0.005},$$

we reject H_0.

(c) The results of the t test and the Wilcoxon test are similar.

8.4–16 **(a)** Here is the two-sided stem-and-leaf display.

Young Subjects	Stems	Older Subjects
9	$3\bullet$	
3	$4*$	
9 8 8 6 5	$4\bullet$	6
0	$5*$	3 4
8 8 7 7 7 6 6 6	$5\bullet$	7 8 9 9
	$6*$	2 2
9	$6\bullet$	5 7
	$7*$	2
	$7\bullet$	
	$8*$	1 3
	$8\bullet$	6 8
	$9*$	3

(b) The value of the Wilcoxon statistic, the sum of the ranks for the younger subjects, is $w = 198$. Since
$$z = \frac{198 - 297.5}{29.033} = -3.427,$$
we clearly reject H_0.

(c) The t test leads to the same conclusion.

8.4–18 **(a)** Using the Wilcoxon statistic, the sum of the ranks for the normal air is 102. Since
$$z = \frac{102 - 126}{\sqrt{168}} = -1.85,$$
we reject the null hypothesis. The p-value is approximately 0.03.

(b) Using a t statistic, we failed to reject the null hypothesis at an $\alpha = 0.05$ significance level.

(c) For these data, the results are a little different with the Wilcoxon statistic leading to rejection of the null hypothesis while the t test did not reject H_0.

8.5 Power of a Statistical Test

8.5–2 **(a)**
$$\begin{aligned} K(\mu) &= P(\,\overline{X} \le 354.05;\ \mu) \\ &= P\left(Z \le \frac{354.05 - \mu}{2/\sqrt{12}};\ \mu\right) \\ &= \Phi\left(\frac{354.05 - \mu}{2/\sqrt{12}}\right); \end{aligned}$$

(b) $\alpha = K(355) = \Phi\left(\dfrac{354.05 - 355}{2/\sqrt{12}}\right) = \Phi(-1.645) = 0.05;$

(c)
$$\begin{aligned} K(354.05) &= \Phi(0) = 0.5; \\ K(353.1) &= \Phi(1.645) = 0.95. \end{aligned}$$

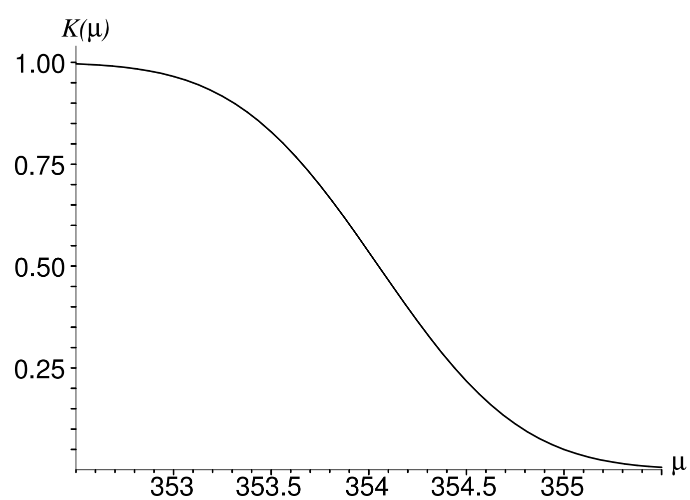


Figure 8.5–2: $K(\mu) = \Phi([354.05 - \mu]/[2/\sqrt{12}\,])$

(d) $\overline{x} = 353.83 < 354.05$, reject H_0;

(e) $$\begin{aligned} p\text{-value} &= P(\,\overline{X} \le 353.83;\ \mu = 355) \\ &= P(Z \le -2.03) = 0.0212. \end{aligned}$$

8.5–4 **(a)** $$\begin{aligned} K(\mu) &= P(\,\overline{X} \ge 83;\ \mu) \\ &= P\left(Z \ge \frac{83-\mu}{10/5}\right) = 1 - \Phi\left(\frac{83-\mu}{2}\right); \end{aligned}$$

(b) $\alpha = K(80) = 1 - \Phi(1.5) = 0.0668;$

(c) $$\begin{aligned} K(80) &= \alpha = 0.0668, \\ K(83) &= 1 - \Phi(0) = 0.5000, \\ K(86) &= 1 - \Phi(-1.5) = 0.9332; \end{aligned}$$

(d)

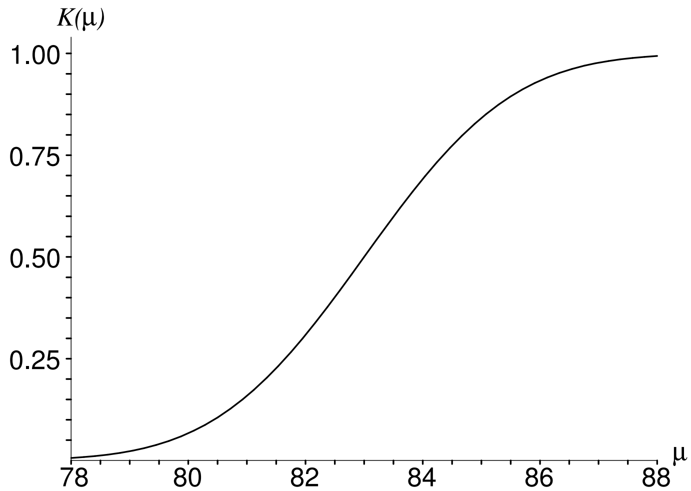


Figure 8.5–4: $K(\mu) = 1 - \Phi([83-\mu]/2)$

(e) $$\begin{aligned} p\text{-value} &= P(\,\overline{X} \ge 83.41;\ \mu = 80) \\ &= P(Z \ge 1.705) = 0.0441. \end{aligned}$$

8.5–6 **(a)** $$\begin{aligned} K(\mu) &= P(\,\overline{X} \le 668.94;\ \mu) = P\left(Z \le \frac{668.94-\mu}{140/5};\ \mu\right) \\ &= \Phi\left(\frac{668.94-\mu}{140/5}\right); \end{aligned}$$

(b) $$\begin{aligned} \alpha = K(715) &= \Phi\left(\frac{668.94-715}{140/5}\right) \\ &= \Phi(-1.645) = 0.05; \end{aligned}$$

(c) $$\begin{aligned} K(668.94) &= \Phi(0) = 0.5; \\ K(622.88) &= \Phi(1.645) = 0.95; \end{aligned}$$

(d)

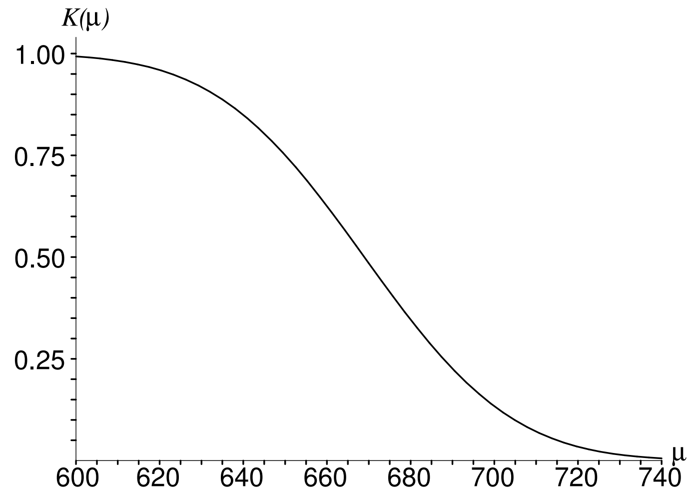


Figure 8.5–6: $K(\mu) = \Phi([668.94 - \mu]/[140/5])$

(e) $\overline{x} = 667.992 < 668.94$, reject H_0;

(f)
$$\begin{aligned} p\text{-value} &= P(\,\overline{X} \le 667.92;\ \mu = 715) \\ &= P(Z \le -1.68) = 0.0465. \end{aligned}$$

8.5–8 **(a)** and **(b)**

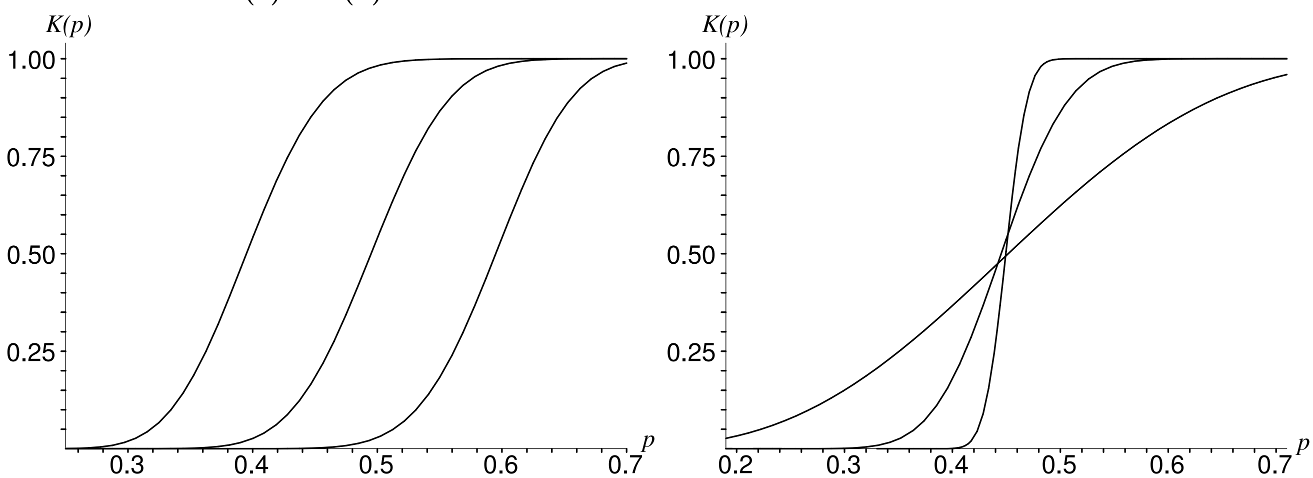


Figure 8.5–8: Power functions corresponding to different critical regions and different sample sizes

8.5–10 Let $Y = \sum_{i=1}^{8} X_i$. Then Y has a Poisson distribution with mean $\mu = 8\lambda$.

(a)
$$\begin{aligned} \alpha &= P(Y \ge 8;\ \lambda = 0.5) = 1 - P(Y \le 7;\ \lambda) = 0.5 \\ &= 1 - 0.949 = 0.051. \end{aligned}$$

(b) $K(\lambda) = 1 - \sum_{y=0}^{7} \frac{(8\lambda)^y e^{-8\lambda}}{y!}$.

(c)
$$\begin{aligned} K(0.75) &= 1 - 0.744 = 0.256, \\ K(1.00) &= 1 - 0.453 = 0.547, \\ K(1.25) &= 1 - 0.220 = 0.780. \end{aligned}$$

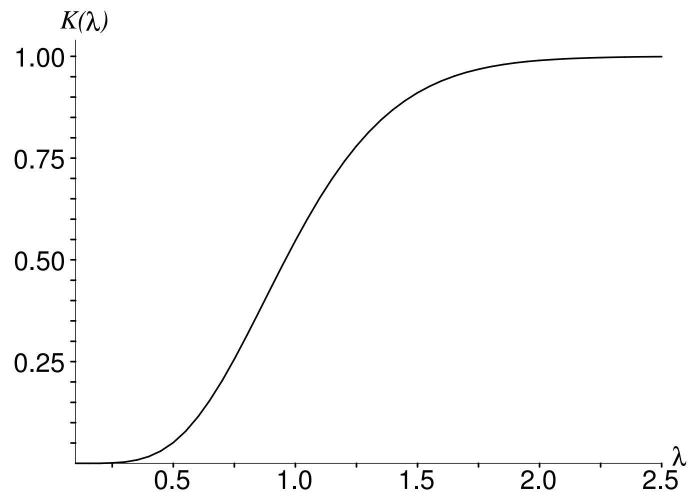


Figure 8.5–10: $K(\lambda) = 1 - P(Y \le 7;\, \lambda)$

8.5–12 **(a)** $\sum_{i=1}^{3} X_i$ has gamma distribution with parameters $\alpha = 3$ and θ. Thus

$$K(\theta) = \int_0^2 \frac{1}{\Gamma(3)\theta^3}\, x^{3-1} e^{-x/\theta} dx;$$

(b)

$$\begin{aligned} K(\theta) = \int_0^2 \frac{x^2 e^{-x/\theta}}{2\theta^3}\, dx &= \frac{1}{2\theta^3}\Big[-\theta x^2 e^{-x/\theta} - 2\theta^2 x e^{-x/\theta} - 2\theta^3 e^{-x/\theta}\Big]_0^2 \\ &= 1 - \sum_{y=0}^{2} \frac{(2/\theta)^y}{y!} e^{-2/\theta}; \end{aligned}$$

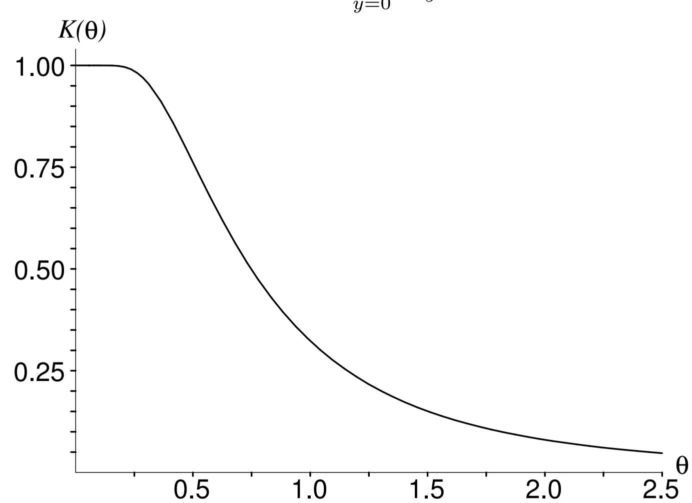


Figure 8.5–12: $K(\theta) = P(\sum_{i=1}^{3} X_i \le 2)$

(c)

$$\begin{aligned} K(2) &= 1 - \sum_{y=0}^{2} \frac{1^y e^{-1}}{y!} = 1 - 0.920 = 0.080; \\ K(1) &= 1 - 0.677 = 0.323; \\ K(1/2) &= 1 - 0.238 = 0.762; \\ K(1/4) &= 1 - 0.014 = 0.986. \end{aligned}$$

8.6 Best Critical Regions

8.6–2 **(a)**

$$\begin{aligned}
\frac{L(4)}{L(16)} &= \frac{(1/2\sqrt{2\pi}\,)^n \exp[-\Sigma\, x_i^2/8\,]}{(1/4\sqrt{2\pi}\,)^n \exp[-\Sigma\, x_i^2/32\,]} \\
&= 2^n \exp[-3\Sigma\, x_i^2/32\,] \le k \\
-\frac{3}{32}\sum_{i=1}^{n} x_i^2 &\le \ln k - \ln 2^n \\
\sum_{i=1}^{n} x_i^2 &\ge -\left(\frac{32}{3}\right)(\ln k - \ln 2^n) = c;
\end{aligned}$$

(b)

$$\begin{aligned}
0.05 &= P\left(\sum_{i=1}^{15} X_i^2 \ge c;\ \sigma^2 = 4\right) \\
&= P\left(\frac{\sum_{i=1}^{15} X_i^2}{4} \ge \frac{c}{4};\ \sigma^2 = 4\right)
\end{aligned}$$

Thus $\dfrac{c}{4} = \chi^2_{0.05}(15) = 25$ and $c = 100$.

(c)

$$\begin{aligned}
\beta &= P\left(\sum_{i=1}^{15} X_i^2 < 100; \sigma^2 = 16\right) \\
&= P\left(\frac{\sum_{i=1}^{15} X_i^2}{16} < \frac{100}{16} = 6.25\right) \approx 0.025.
\end{aligned}$$

8.6–4 **(a)**

$$\begin{aligned}
\frac{L(0.9)}{L(0.8)} = \frac{(0.9)^{\Sigma x_i}(0.1)^{n-\Sigma x_i}}{(0.8)^{\Sigma x_i}(0.2)^{n-\Sigma x_i}} &\le k \\
\left[\left(\frac{9}{8}\right)\left(\frac{2}{1}\right)\right]^{\sum_1^n x_i}\left[\frac{1}{2}\right]^n &\le k \\
\left(\sum_{i=1}^{n} x_i\right)\ln(9/4) &\le \ln k + n\ln 2 \\
y = \sum_{i=1}^{n} x_i &\le \frac{\ln k + n\ln 2}{\ln(9/4)} = c.
\end{aligned}$$

Recall that the distribution of the sum of Bernoulli trials, Y, is $b(n,\, p)$.

(b)

$$\begin{aligned}
0.10 &= P[\,Y \le n(0.85);\ p = 0.9\,] \\
&= P\left[\frac{Y - n(0.9)}{\sqrt{n(0.9)(0.1)}} \le \frac{n(0.85) - n(0.9)}{\sqrt{n(0.9)(0.1)}};\ p = 0.9\right].
\end{aligned}$$

It is true, approximately, that

$$\begin{aligned}
\frac{n(-0.05)}{\sqrt{n}(0.3)} &= -1.282 \\
n &= 59.17 \quad\text{or}\quad n = 60.
\end{aligned}$$

(c)

$$\begin{aligned}
\beta = P[\,Y > n(0.85) = 51;\ p = 0.8\,] &= P\left[\frac{Y - 60(0.8)}{\sqrt{60(0.8)(0.2)}} > \frac{51-48}{\sqrt{9.6}};\ p = 0.8\right] \\
&\approx P(Z \ge 0.97) = 0.166.
\end{aligned}$$

(d) Yes.

8.6–6 **(a)**
$$\begin{aligned} 0.05 &= P\left(\frac{\overline{X}-80}{3/4} \ge \frac{c_1-80}{3/4}\right) \\ &= 1-\Phi\left(\frac{c_1-80}{3/4}\right). \end{aligned}$$

Thus
$$\begin{aligned} \frac{c_1-80}{3/4} &= 1.645 \\ c_1 &= 81.234. \end{aligned}$$

Similarly,
$$\begin{aligned} \frac{c_2-80}{3/4} &= -1.645 \\ c_2 &= 78.766; \\ \frac{c_3}{3/4} &= 1.96 \\ c_3 &= 1.47. \end{aligned}$$

(b)
$$\begin{aligned} K_1(\mu) &= 1-\Phi([81.234-\mu]/[3/4]); \\ K_2(\mu) &= \Phi([78.766-\mu]/[3/4]; \\ K_3(\mu) &= 1-\Phi([81.47-\mu]/[3/4]) + \Phi([78.53-\mu]/[3/4]). \end{aligned}$$

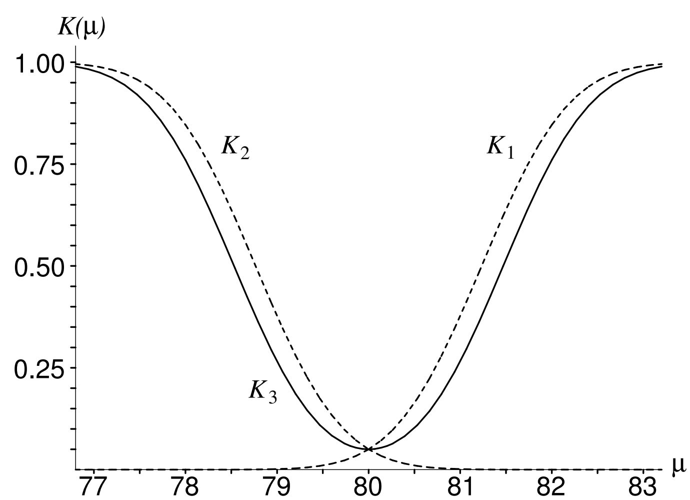


Figure 8.6–6: Three power functions

8.6–8 Let
$$L(\theta) = \theta^n \prod_{i=1}^{n} (1-x_i)^{\theta-1};$$
thus
$$L(\theta_0) = \theta_0^n \prod_{i=1}^{n} (1-x_i)^{\theta_0-1} = 1,$$
because $\theta_0 = 1$. So when $\theta > 1$, a best critical region is
$$\frac{L(\theta_0)}{L(\theta)} = \frac{1}{\theta^n \prod_{i=1}^{n}(1-x_i)^{\theta-1}} \le k.$$

That is,

$$\theta^n \left[\prod_{i=1}^{n}(1-x_i)\right]^{\theta-1} \geq \frac{1}{k}$$

or

$$\prod_{i=1}^{n}(1-x_i) \geq \left(\frac{1}{k\theta^n}\right)^{1/(\theta-1)} = c.$$

This is true for all $\theta > 1$; so it is a uniformly most powerful test.

8.7 Likelihood Ratio Tests

8.7–2 **(a)** If $\mu \in \omega$ (that is, $\mu \geq 10.35$), then $\widehat{\mu} = \overline{x}$ if $\overline{x} \geq 10.35$, but $\widehat{\mu} = 10.35$ if $\overline{x} < 10.35$.

Thus $\lambda = 1$ if $\overline{x} \geq 10.35$; but, if $\overline{x} < 10.35$, then

$$\begin{aligned}
\lambda = \frac{[1/(0.3)(2\pi)]^{n/2}\exp[-\sum_1^n (x_i-10.35)^2/(0.06)]}{[1/(0.3)(2\pi)]^{n/2}\exp[-\sum_1^n (x_i-\overline{x})^2/(0.06)]} &\leq k \\
\exp\left[-\frac{n}{0.06}(\overline{x}-10.35)^2\right] &\leq k \\
-\frac{n}{0.06}(\overline{x}-10.35)^2 &\leq \ln k \\
\frac{\overline{x}-10.35}{\sqrt{0.03/n}} &\leq \sqrt{-2\ln k} = -z_{0.05} \\
&= -1.645.
\end{aligned}$$

(b) $\dfrac{10.31-10.35}{\sqrt{0.03/50}} = -1.633 > -1.645$; do not reject H_0.

(c) p-value $= P(Z \leq -1.633) = 0.0513$.

8.7–4 **(a)**

$$\begin{aligned}
\lambda &= \frac{[1/(2\pi\sigma_0^2)]^{n/2}\exp[-\sum_{i=1}^n (x_i-\mu_0)^2/(2\sigma_0^2)]}{[1/(2\pi\sigma_0^2)]^{n/2}\exp[-\sum_{i=1}^n (x_i-\overline{x})^2/(2\sigma_0^2)]} \\
&= \exp\left[\frac{-\sum_{i=1}^n (x_i-\overline{x}+\overline{x}-\mu_0)^2}{2\sigma_0^2} + \frac{\sum_{i=1}^n (x_i-\overline{x})^2}{2\sigma_0^2}\right] \\
&= \exp\left[\frac{-n(\overline{x}-\mu_0)^2}{2\sigma_0^2}\right] \leq k
\end{aligned}$$

$$\begin{aligned}
\frac{-n(\overline{x}-\mu_0)^2}{2\sigma_0^2} &\leq \ln k \\
\frac{|\overline{x}-\mu_0|}{\sigma_0/\sqrt{n}} &\geq c \\
|z| = \frac{|\,\overline{x}-59\,|}{15/\sqrt{n}} &\geq 1.96.
\end{aligned}$$

(b) $|z| = \dfrac{|\,56.13-59\,|}{15/10} = |-1.913| < 1.96$, do not reject H_0;

(c) p-value $= P(|Z| \geq 1.913) = 0.0558$.

8.7–6 $t = \dfrac{324.8 - 335}{40/\sqrt{17}} = -1.051 > -1.337$, do not reject H_0.

8.7–8 In Ω, $\widehat{\mu} = \overline{x}$. Thus,

$$\begin{aligned}\lambda = \frac{(1/\theta_0)^n \exp[-\sum_1^n x_i/\theta_0]}{(1/\overline{x})^n \exp[-\sum_1^n x_i/\overline{x}]} &\leq k \\ \left(\frac{\overline{x}}{\theta_0}\right)^n \exp[-n(\overline{x}/\theta_0 - 1)] &\leq k.\end{aligned}$$

Plotting λ as a function of $w = \overline{x}/\theta_0$, we see that $\lambda = 0$ when $\overline{x}/\theta_0 = 0$, it has a maximum when $\overline{x}/\theta_0 = 1$, and it approaches 0 as $\overline{x}/\theta_0$ becomes large. Thus $\lambda \leq k$ when $\overline{x} \leq c_1$ or $\overline{x} \geq c_2$.

Since the distribution of $\dfrac{2}{\theta_0}\displaystyle\sum_{i=1}^n X_i$ is $\chi^2(2n)$ when H_0 is true, we could let the critical region be such that we reject H_0 if

$$\frac{2}{\theta_0}\sum_{i=1}^n X_i \leq \chi^2_{1-\alpha/2}(2n) \quad \text{or} \quad \frac{2}{\theta_0}\sum_{i=1}^n X_i \geq \chi^2_{\alpha/2}(2n).$$

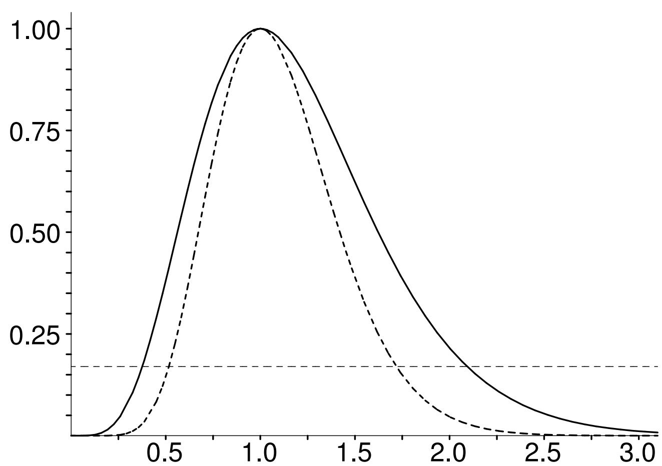


Figure 8.7–8: Likehood functions: solid, $n = 5$; dotted, $n = 10$

8.7–10 Referring to Exercise 6.4-19, we test H_0: $\gamma^2 = 1$ against H_1: $\gamma^2 \neq 1$.

$$
\begin{aligned}
\lambda &= \frac{\left(\displaystyle\prod_{i=1}^{n} \frac{1}{\sqrt{2\pi x_i^2}}\right) \exp\left[-\frac{\sum_{i=1}^{n}(y_i - \widehat{\mu})^2}{2x_i^2}\right]}{\left(\displaystyle\prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\,\widehat{\gamma^2}\,x_i^2}}\right) \exp\left[-\frac{\sum_{i=1}^{n}(y_i - \widehat{\mu})^2}{2\,\widehat{\gamma^2}\,x_i^2}\right]} \;\le\; k \\
&= (\,\widehat{\gamma^2}\,)^{n/2} \exp\left[-\frac{\sum_{i=1}^{n}(y_i - \widehat{\mu})^2}{2x_i^2} + \frac{\sum_{i=1}^{n}(y_i - \widehat{\mu})^2}{2\,\widehat{\gamma^2}\,x_i^2}\right] \;\le\; k \\
&= (\,\widehat{\gamma^2}\,)^{n/2} \exp[-n\,\widehat{\gamma^2}/2 + n/2] \;\le\; k \\
&= (\,\widehat{\gamma^2}\,)^{n/2} \exp[n(1 - \widehat{\gamma^2}\,)/2] \;\le\; k.
\end{aligned}
$$

This is a function of

$$\widehat{\gamma^2} = \frac{1}{n}\sum_{i=1}^{n} \frac{(y_i - \widehat{\mu})^2}{x_i^2}$$

where

$$\widehat{\mu} = \frac{\sum_{i=1}^{n} y_i/x_i^2}{\sum_{i=1}^{n} 1/x_i^2}.$$

Chapter 9

More Tests

9.1 Chi-Square Goodness-of-Fit Tests

9.1–2
$$\begin{aligned} q_4 &= \frac{(224-232)^2}{232} + \frac{(119-116)^2}{116} + \frac{(130-116)^2}{116} + \frac{(48-58)^2}{58} + \frac{(59-58)^2}{58} \\ &= 3.784. \end{aligned}$$

The null hypothesis will not be rejected at any reasonable significance level. Note that $E(Q_4) = 4$ when H_0 is true.

9.1–4
$$\begin{aligned} q_3 &= \frac{(124-117)^2}{117} + \frac{(30-39)^2}{39} + \frac{(43-39)^2}{39} + \frac{(11-13)^2}{13} \\ &= 0.419 + 2.077 + 0.410 + 0.308 = 3.214 < 7.815 = \chi^2_{0.05}(3). \end{aligned}$$

Thus we do not reject the Mendelian theory with these data.

9.1–6 Using Table II in Appendix B with $p = 0.30$, the hypothesized probabilities are $p_0 = P(X=0) = 0.340$, $p_1 = P(X=1) = 0.441$, $p_2 = P(X=2) = 0.189$, $p_3 = P(X=3) = 0.027$. Thus the respective expected values are 68.6, 88.2, 37.8, and 5.4. The value of the chi-square goodness of fit statistic is:

$$\begin{aligned} q &= \frac{(57-68.6)^2}{68.6} + \frac{(95-88.2)^2}{88.2} + \frac{(38-37.8)^2}{37.8} + \frac{(10-5.4)^2}{5.4} \\ &= 6.405 < 7.815 = \chi^2_{0.05}(3). \end{aligned}$$

Do not reject the hypothesis that X is $b(3,\, 0.30)$ at a 5% significance level. Limits for the p-value are $0.05 < p\text{-value} < 0.10$ because $\chi^2_{0.10}(3) = 6.251 < 6.405 < \chi^2_{0.05}(3) = 7.815$.

We find that $\widehat{p} = 201/600 = 0.335$. Thus a 95% confidence interval for p is

$$0.335 \pm 1.96\sqrt{(0.335)(0.665)/600} \quad \text{or} \quad [0.297, 0.373].$$

The pennies that were used were minted 1998 or earlier. See Figure 9.1-6. Repeat this experiment with similar pennies or with newer pennies and compare your results with those obtained by these students.

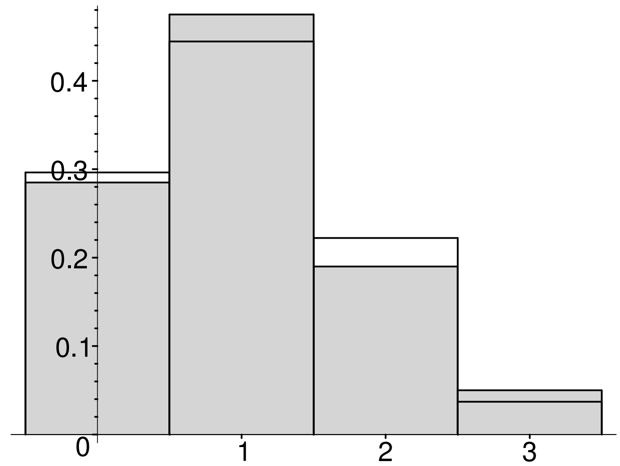


Figure 9.1–6: The $b(3,\ 3/10)$ probability histogram and the relative frequency histogram (shaded)

9.1–8 The respective probabilities and expected frequencies are 0.050, 0.149, 0.224, 0.224, 0.168, 0.101, 0.050, 0.022, 0.012 and 15.0, 44.7, 67.2, 67.2, 50.4, 30.3, 15.0, 6.6, 3.6. The last two cells could be combined to give an expected frequency of 10.2. From Exercise 3.5–12, the respective frequencies are 17, 47, 63, 63, 49, 28, 21, and 12 giving

$$q_7 = \frac{(17-15.0)^2}{15.0} + \frac{(47-44.7)^2}{44.7} + \cdots + \frac{(12-10.2)^2}{10.2} = 3.841.$$

Since $3.841 < 14.07 = \chi^2_{0.05}(7)$, do not reject. The sample mean is $\overline{x} = 3.03$ and the sample variance is $s^2 = 3.19$ which also supports the hypothesis. The following figure compares the probability histogram with the relative frequency histogram of the data.

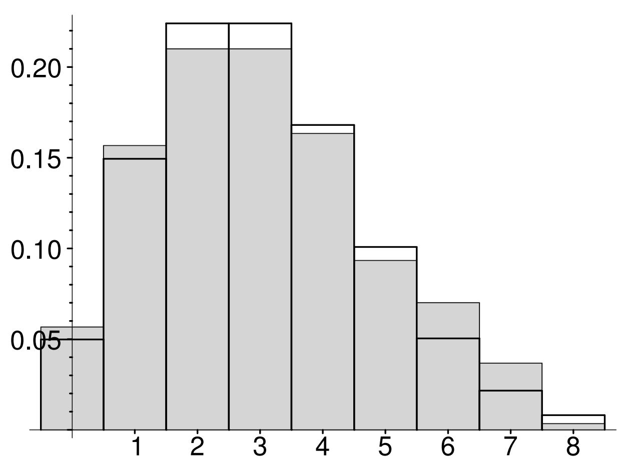


Figure 9.1–8: The Poisson probability histogram, $\lambda = 3$, and relative frequency histogram (shaded)

9.1–10 **(a)** For the infected snails, $\overline{x} = 84.74,\ s_x = 64.79$;
For the control snails, $\overline{y} = 113.16,\ s_y = 87.02$;

(b)

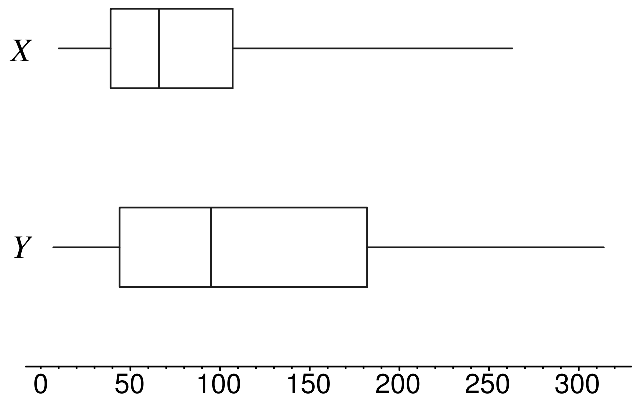


Figure 9.1–10: Box-and-whisker diagrams for infected (X) and control (Y)

(c) We shall use 5 classes with equal probability for the control snails.

A_i	Observed	Expected	q
[0, 25.25)	4	6.2	0.781
[25.25, 57.81)	9	6.2	1.264
[57.81, 103.69)	5	6.2	0.232
[193.69, 182.13)	6	6.2	0.006
[182.13, ∞)	7	6.2	0.103
	31	31.0	2.386

The p-value for $5 - 1 - 1 = 3$ degrees of freedom is 0.496 so we fail to reject the null hypothesis.

(d) We shall use 10 classes with equal probability for the infected snalis.

A_i	Observed	Expected	q
[0, 22.34)	3	3.9	0.208
[22.34, 34.62)	3	3.9	0.208
[34.62, 46.09)	8	3.9	4.310
[46.09, 57.81)	4	3.9	0.003
[57.81, 70.49)	2	3.9	0.926
[70.49, 84.94)	4	3.9	0.003
[84.94, 102.45)	4	3.9	0.003
[102.45, 125.76)	4	3.9	0.003
[125.76, 163.37)	2	3.9	0.926
[163.37, ∞)	5	3.9	0.310
	39	39.0	6.900

The p-value for $10 - 1 = 9$ degrees of freedom is 0.648 so we fail to reject the null hypothesis.

9.2 Contingency Tables

9.2–2 $10.18 < 20.48 = \chi^2_{0.025}(10)$, accept H_0.

9.2–4 In the combined sample of 45 observations, the lower third includes those with scores of 61 or lower, the middle third have scores from 62 through 78, and the higher third are those with scores of 79 and above.

	low	middle	high	Totals
Class U	9 (5)	4 (5)	2 (5)	15
Class V	5 (5)	5 (5)	5 (5)	15
Class W	1 (5)	6 (5)	8 (5)	15
Totals	15	15	15	45

Thus

$$q = 3.2 + 0.2 + 1.8 + 0 + 0 + 0 + 3.2 + 0.2 + 1.8 = 10.4.$$

Since

$$q = 10.4 > 9.488 = \chi^2_{0.05}(4),$$

we reject the equality of these three distributions. (p-value $= 0.034$.)

9.2–6 $q = 8.410 < 9.488 = \chi^2_{0.05}$, fail to reject H_0. (p-value $= 0.078$.)

9.2–8 $q = 4.268 > 3.841 = \chi^2_{0.05}(1)$, reject H_0. (p-value $= 0.039$.)

9.2–10 $q = 7.683 < 9.210 = \chi^2_{0.01}$, fail to reject H_0. (p-value $= 0.021$.)

9.2–12 $q = 8.792 > 7.378 = \chi^2_{0.025}(2)$, reject H_0. (p-value $= 0.012$.)

9.3 One-Factor Analysis of Variance

9.3–2

Source	SS	DF	MS	F	p-value
Treatment	388.2805	3	129.4268	4.9078	0.0188
Error	316.4597	12	26.3716		
Total	704.7402	15			

$F = 4.9078 > 3.49 = F_{0.05}(3, 12)$, reject H_0.

9.3–4 **(a)**

Source	SS	DF	MS	F	p-value
Treatment	184.8	2	92.4	15.4	0.00015
Error	102.0	17	6.0		
Total	286.8	19			

$F = 15.4 > 3.59 = F_{0.05}(2, 17)$, reject H_0.

(b)

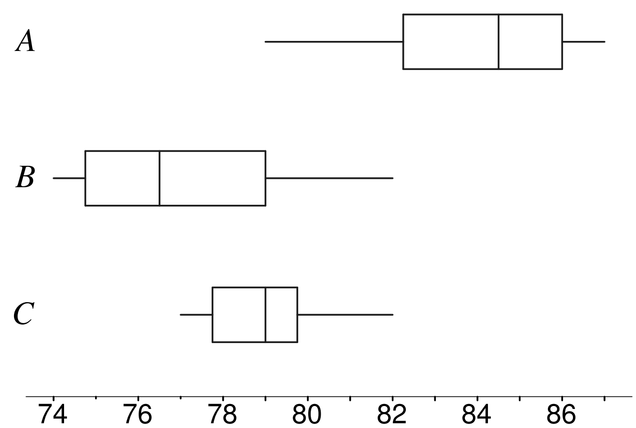


Figure 9.3–4: Box-and-whisker diagrams of strengths of different beams

9.3–6 **(a)** $F \geq F_{0.05}(3, 24) = 3.01$;

(b)

Source	SS	DF	MS	F	p-value
Treatment	12,280.86	3	4,093.62	3.455	0.0323
Error	28,434.57	24	1,184.77		
Total	40,715.43	27			

$F = 3.455 > 3.01$, reject H_0;

(c) $0.025 < p\text{-value} < 0.05$.

(d)

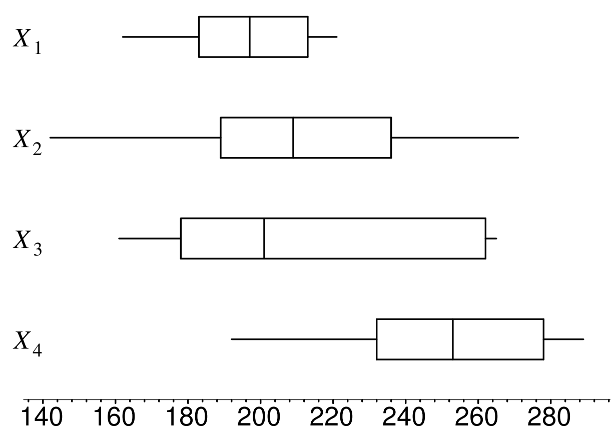


Figure 9.3–6: Box-and-whisker diagrams for cholesterol levels

9.3–8 **(a)** $F \geq F_{0.05}(4, 30) = 2.69$;

(b)

Source	SS	DF	MS	F	p-value
Treatment	0.00442	4	0.00111	2.85	0.0403
Error	0.01157	30	0.00039		
Total	0.01599	34			

$F = 2.85 > 2.69$, reject H_0;

(c)

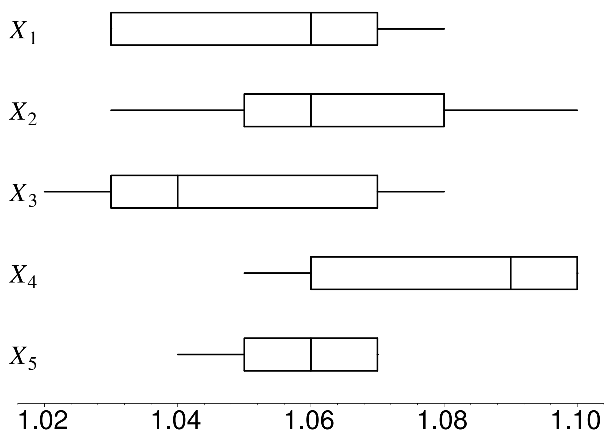


Figure 9.3–8: Box-and-whisker diagrams for nail weights

9.3–10 **(a)** $t = \dfrac{92.143 - 103.000}{\sqrt{\dfrac{6(69.139) + 6(57.669)}{12}\left(\dfrac{1}{7} + \dfrac{1}{7}\right)}} = -2.55 < -2.179$, reject H_0.

$F = \dfrac{412.517}{63.4048} = 6.507 > 4.75$, reject H_0.

The F and the t tests give the same results since $t^2 = F$.

(b) $F = \dfrac{86.3336}{114.8889} = 0.7515 < 3.55$, do not reject H_0.

9.3–12 **(a)**

Source	SS	DF	MS	F	p-value
Treatment	122.1956	2	61.0978	2.130	0.136
Error	860.4799	30	28.6827		
Total	982.6755	32			

$F = 2.130 < 3.32 = F_{0.05}(2, 30)$, fail to reject H_0;

(b)

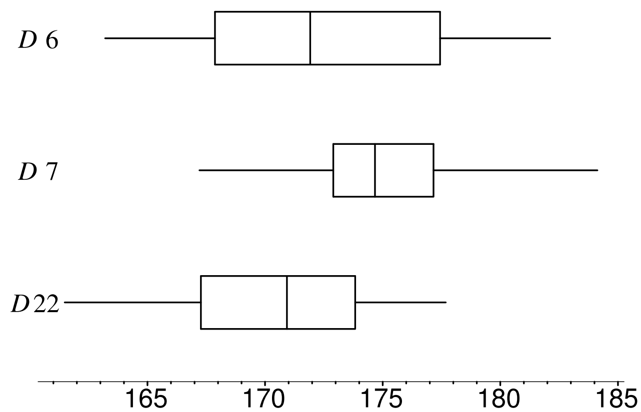


Figure 9.3–12: Box-and-whisker diagrams for resistances on three days

9.3–14 **(a)**

Source	SS	DF	MS	F	p-value
Worker	1.5474	2	0.7737	1.0794	0.3557
Error	17.2022	24	0.7168		
Total	18.7496	26			

$F = 1.0794 < 3.40 = F_{0.05}(2, 24)$, fail to reject H_0;

(b)

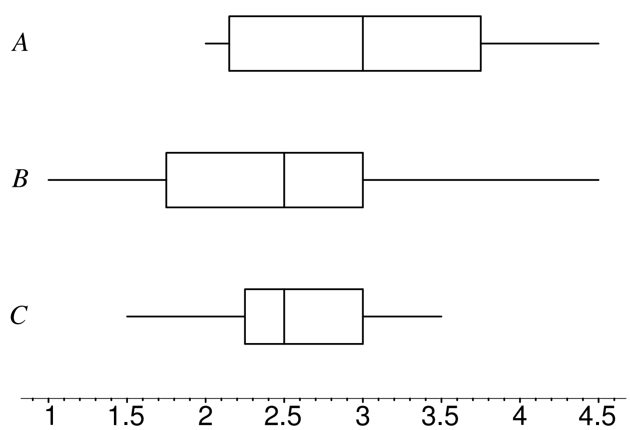


Figure 9.3–14: Box-and-whisker diagrams for workers A, B, and C

The box plot confirms the answer from part (a).

9.4 Two-Way Analysis of Variance

9.4–2

$$\begin{array}{c|cccc|c} & & & & & \mu+\alpha_i \\ & 6 & 3 & 7 & 8 & 6 \\ & 10 & 7 & 11 & 12 & 10 \\ & 8 & 5 & 9 & 10 & 8 \\ \hline \mu+\beta_j & 8 & 5 & 9 & 10 & \mu = 8 \end{array}$$

So $\alpha_1 = -2,\ \alpha_2 = 2,\ \alpha_3 = 0$ and $\beta_1 = 0,\ \beta_2 = -3,\ \beta_3 = 1,\ \beta_4 = 2$.

9.4–4 $\displaystyle\sum_{i=1}^{a}\sum_{j=1}^{b}(\overline{X}_{i\cdot} - \overline{X}_{\cdot\cdot})(X_{ij} - \overline{X}_{i\cdot} - \overline{X}_{\cdot j} + \overline{X}_{\cdot\cdot})$

$$\begin{aligned} &= \sum_{i=1}^{a}(\overline{X}_{i\cdot} - \overline{X}_{\cdot\cdot})\sum_{j=1}^{b}[(X_{ij} - \overline{X}_{i\cdot}) - (\overline{X}_{\cdot j} - \overline{X}_{\cdot\cdot})] \\ &= \sum_{i=1}^{a}(\overline{X}_{i\cdot} - \overline{X}_{\cdot\cdot})\left\{\sum_{j=1}^{b}(X_{ij} - \overline{X}_{i\cdot}) - \sum_{j=1}^{b}(\overline{X}_{\cdot j} - \overline{X}_{\cdot\cdot})\right\} \\ &= \sum_{i=1}^{a}(\overline{X}_{i\cdot} - \overline{X}_{\cdot\cdot})(0-0) = 0; \end{aligned}$$

$\displaystyle\sum_{i=1}^{a}\sum_{j=1}^{b}(\overline{X}_{\cdot j} - \overline{X}_{\cdot\cdot})(X_{ij} - \overline{X}_{i\cdot} - \overline{X}_{\cdot j} + \overline{X}_{\cdot\cdot}) = 0,$ similarly;

$$\sum_{i=1}^{a}\sum_{j=1}^{b}(\overline{X}_{i\cdot} - \overline{X}_{\cdot\cdot})(\overline{X}_{\cdot j} - \overline{X}_{\cdot\cdot}) = \left\{\sum_{i=1}^{a}(\overline{X}_{i\cdot} - \overline{X}_{\cdot\cdot})\right\}\left\{\sum_{j=1}^{b}(\overline{X}_{\cdot j} - \overline{X}_{\cdot\cdot})\right\} = (0)(0) = 0\ .$$

9.4–6

$$\begin{array}{c|cccc|c} & & & & & \mu+\alpha_i \\ & 6 & 7 & 7 & 12 & 8 \\ & 10 & 3 & 11 & 8 & 8 \\ & 8 & 5 & 9 & 10 & 8 \\ \hline \mu+\beta_j & 8 & 5 & 9 & 10 & \mu = 8 \end{array}$$

So $\alpha_1 = \alpha_2 = \alpha_3 = 0$ and $\beta_1 = 0,\ \beta_2 = -3,\ \beta_3 = 1,\ \beta_4 = 2$ as in Exercise 9.4–2. However, $\gamma_{11} = -2$ because $8 + 0 + 0 + (-2) = 6$. Similarly we obtain the other γ_{ij}'s :

$$\begin{array}{cccc} -2 & 2 & -2 & 2 \\ 2 & -2 & 2 & -2 \\ 0 & 0 & 0 & 0 \end{array}$$

9.4–8

Source	SS	DF	MS	F	p-value
Row (A)	5,103.0000	1	5,103.0000	4.307	0.049
Col (B)	6,121.2857	1	6,121.2857	5.167	0.032
Int(AB)	1,056.5714	1	1,056.5714	0.892	0.354
Error	28,434.5714	24	1,184.7738		
Total	40,715.4286	27			

(a) Since $F = 0.892 < F_{0.05}(1, 24) = 4.26$, do not reject H_{AB};

(b) Since $F = 4.307 > F_{0.05}(1, 24) = 4.26$, reject H_A;

(c) Since $F = 5.167 > F_{0.05}(1, 24) = 4.26$, reject H_B.

9.5 General Factorial and 2^k Factorial Designs

9.5–4 **(a)** $[A] = -28.4/8 = -3.55$, $[B] = -1.45$, $[C] = 3.2$, $[AB] = -1.525$, $[AC] = -0.525$, $[BC] = 0.375$, $[ABC] = -1.2$.

(b)

Identity of Effect	Ordered Effect	Percentile	Percentile from N(0,1)
[A]	−3.550	12.5	−1.15
[AB]	−1.525	25.0	−0.67
[B]	−1.450	37.5	−0.32
[ABC]	−1.200	50.0	0.00
[AC]	−0.525	62.5	0.32
[BC]	0.375	75.0	0.67
[C]	3.20	87.5	1.15

The main effects of temperature (A) and concentration (C) are significant.

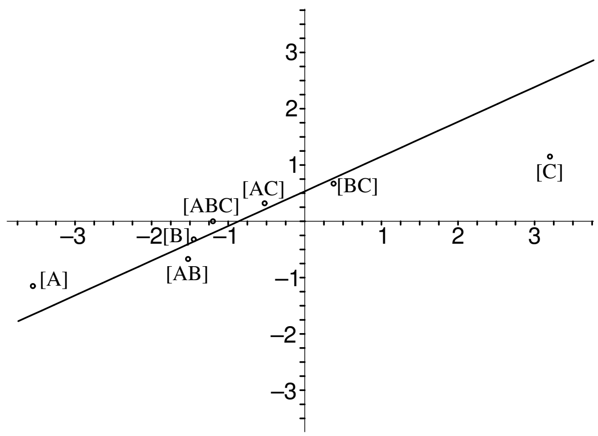


Figure 9.5–4: q-q plot

9.6 Tests Concerning Regression and Correlation

9.6–2 The critical region is $t_1 \geq t_{0.25}(8) = 2.306$. From Exercise 6.5–4,

$$\begin{aligned}\widehat{\beta} &= 4.64/5.04 \text{ and } n\widehat{\sigma^2} = 1.84924; \text{ also } \sum_{i=1}^{10}(x_i - \overline{x})^2 = 5.04, \text{ so} \\ t_1 &= \frac{4.64/5.04}{\sqrt{\dfrac{1.84924}{8(5.04)}}} = \frac{0.9206}{0.2142} = 4.299.\end{aligned}$$

Since $t_1 = 4.299 > 2.306$, we reject H_0.

9.6–4 For these data, $r = -0.413$. Since $|r| = 0.413 < 0.7292$, do not reject H_0.

9.6–6 Following the suggestion given in the hint, the expression equals

$$\begin{aligned}(n-1)S_Y^2 - \frac{2Rs_xS_Y}{s_x^2}(n-1)Rs_xS_Y + \frac{R^2s_x^2S_Y^2(n-1)s_x^2}{s_x^2} &= (n-1)S_Y^2(1 - 2R^2 + R^2) \\ &= (n-1)S_Y^2(1 - R^2).\end{aligned}$$

9.6–8

$$\begin{aligned}u(R) &\approx u(\rho) + (R-\rho)u'(\rho), \\ \text{Var}[u(\rho) + (R-\rho)u'(\rho)] &= [u'(\rho)]^2\text{Var}(R) \\ &= [u'(\rho)]^2\,\frac{(1-\rho^2)^2}{n} = c, \quad \text{which is free of } \rho, \\ u'(\rho) &= \frac{k/2}{1-\rho} + \frac{k/2}{1+\rho}, \\ u(\rho) &= -\frac{k}{2}\ln(1-\rho) + \frac{k}{2}\ln(1+\rho) = \frac{k}{2}\ln\left(\frac{1+\rho}{1-\rho}\right).\end{aligned}$$

Thus, taking $k = 1$,

$$u(R) = \left(\frac{1}{2}\right)\ln\left[\frac{1+R}{1-R}\right]$$

has a variance almost free of ρ.

9.6–10 **(a)** $r = -0.4906, |\,r\,| = 0.4906 > 0.4258$, reject H_0 at $\alpha = 0.10$;

(b) $|\,r\,| = 0.4906 < 0.4973$, fail to reject H_0 at $\alpha = 0.05$.

9.6–12 **(a)** $r = 0.339, |\,r\,| = 0.339 < 0.5325 = r_{0.025}(12)$, fail to reject H_0 at $\alpha = 0.05$;

(b) $r = -0.821 < -0.6613 = r_{0.005}(12)$, reject H_0 at $\alpha = 0.005$;

(c) $r = 0.149, |\,r\,| = 0.149 < 0.5325 = r_{0.025}(12)$, fail to reject H_0 at $\alpha = 0.05$.

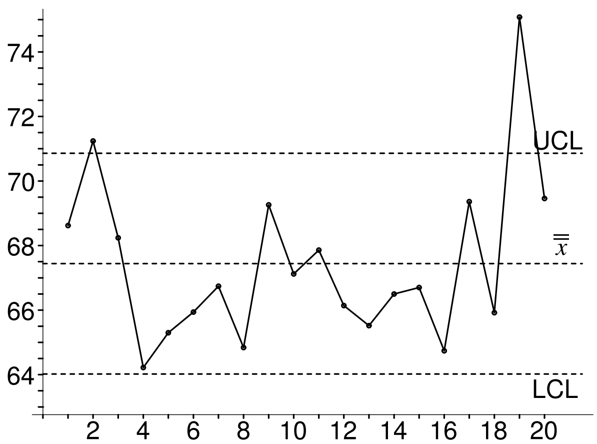
9.7 Statistical Quality Control

9.7–2 **(a)** $\overline{\overline{x}} = 67.44$, $\overline{s} = 2.392$, $\overline{R} = 5.88$;

(b) $\text{UCL} = \overline{\overline{x}} + 1.43(\overline{s}) = 67.44 + 1.43(2.392) = 70.86$;

$\text{LCL} = \overline{\overline{x}} - 1.43(\overline{s}) = 67.44 - 1.43(2.392) = 64.02$;

(c) $\text{UCL} = 2.09(\overline{s}) = 2.09(2.392) = 5.00$; $\text{LCL} = 0$;



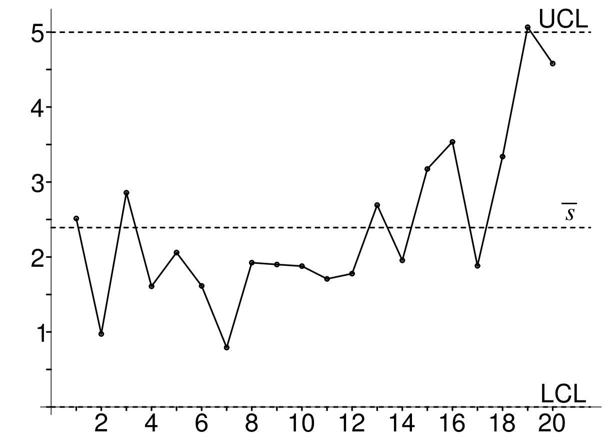
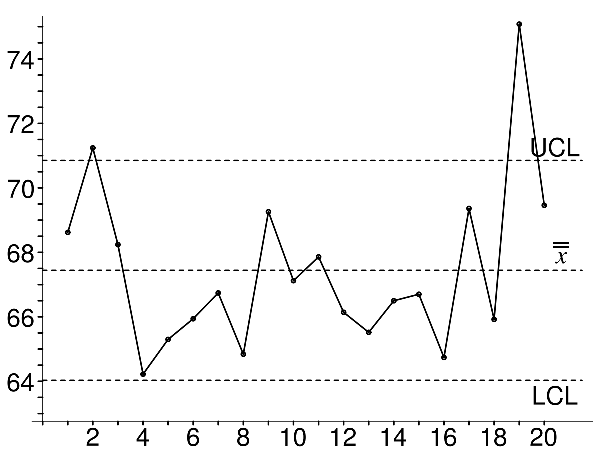


Figure 9.7–2: **(b)** $\overline{x}$-chart using $\overline{s}$ and **(c)** s-chart

(d) $\text{UCL} = \overline{\overline{x}} + 0.58(\overline{R}) = 67.44 + 0.58(5.88) = 70.85$;

$\text{LCL} = \overline{\overline{x}} - 0.58(\overline{R}) = 67.44 - 0.58(5.88) = 64.03$;

(e) $\text{UCL} = 2.11(\overline{R}) = 2.11(5.88) = 12.41$; $\text{LCL} = 0$;



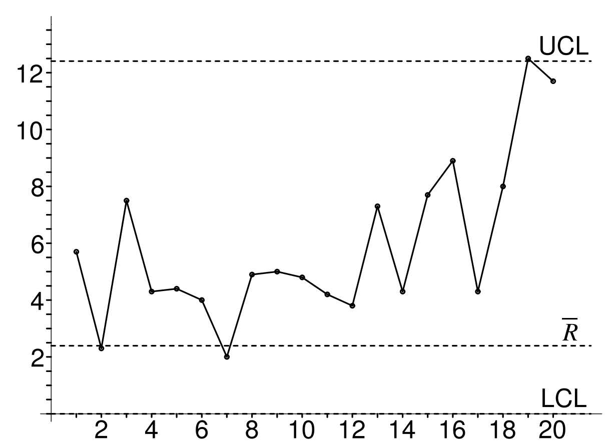


Figure 9.7–2: **(d)** $\overline{x}$-chart using $\overline{R}$ and **(e)** R-chart

(f) Quite well until near the end.

9.7–4 With $\overline{p} = 0.0254$, $\text{UCL} = \overline{p} + 3\sqrt{\overline{p}(1-\overline{p})/100} = 0.073$;

with $\overline{p} = 0.02$, $\text{UCL} = \overline{p} + 3\sqrt{\overline{p}(1-\overline{p})/100} = 0.062$;

In both cases we see that problems are arising near the end.

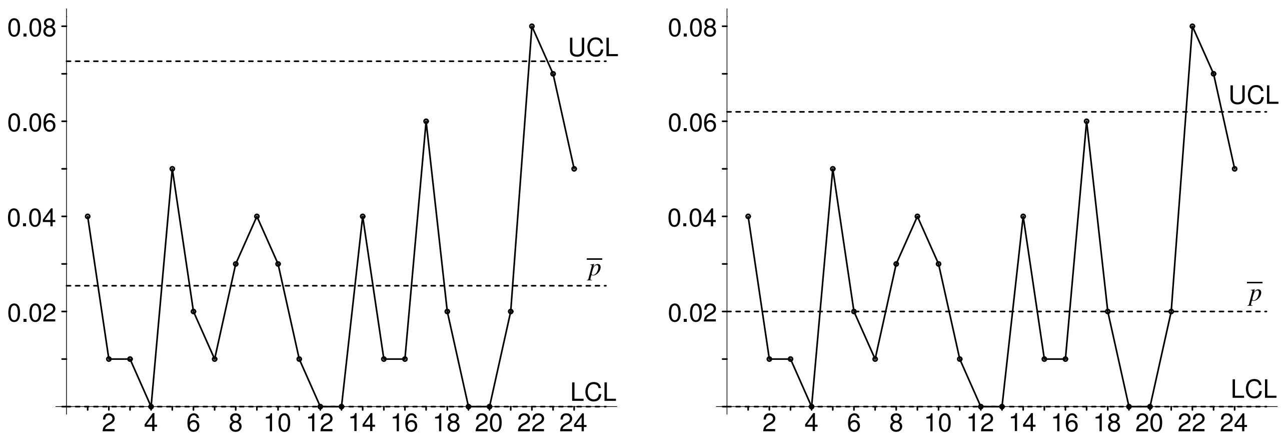


Figure 9.7–4: p-charts using $\overline{p} = 0.0254$ and $\overline{p} = 0.02$

9.7–6 **(a)** $\text{UCL} = \overline{c} + 3\sqrt{\overline{c}} = 1.80 + 3\sqrt{1.80} = 5.825$; $\text{LCL} = 0$;

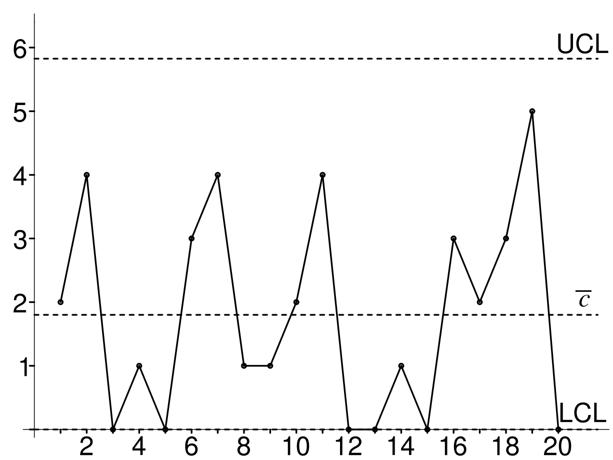


Figure 9.7–6: c-chart

(b) The process is in statistical control.

9.7–8 **(a)** $\overline{\overline{x}} = \dfrac{5101}{150} = 34.0067$, $\overline{s} = 2.8244$, $\overline{R} = 7$.

(b) $\text{UCL} = \overline{\overline{x}} + 1.43(\overline{s}) = 38.046$ for the $\overline{x}$ chart;

$\text{LCL} = \overline{\overline{x}} - 1.43(\overline{s}) = 29.968$ for the $\overline{x}$ chart;

(c) $\text{UCL} = 2.09(\overline{s}) = 2.8244$ for the s chart;

$\text{LCL} = 0(\overline{s})$ for the s chart.

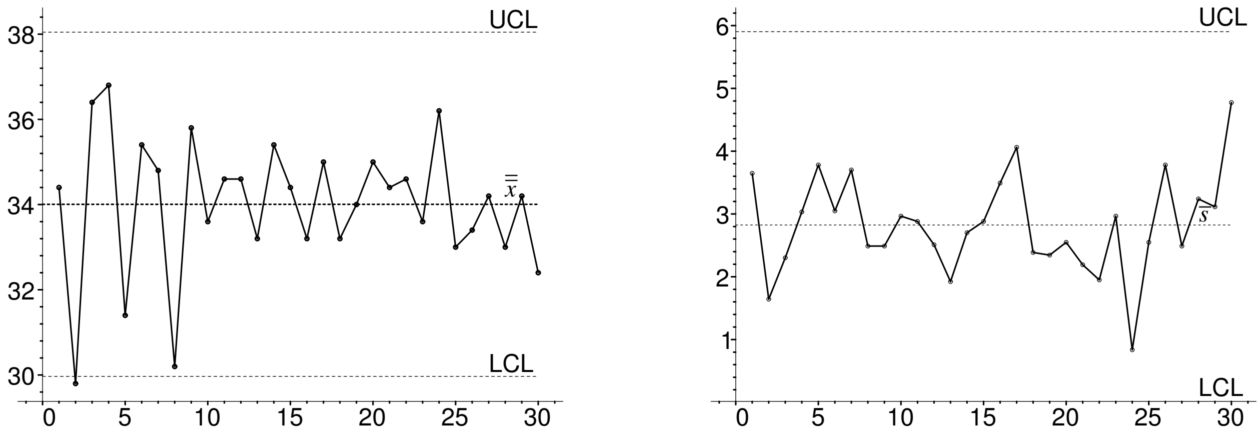


Figure 9.7–8: **(b)** $\overline{x}$-chart using $\overline{s}$ and **(c)** s-chart

(d) $\text{UCL} = \overline{\overline{x}} + 0.58(\overline{R}) = 38.0667$;

$\text{LCL} = \overline{\overline{x}} - 0.58(\overline{R}) = 29.9467$;

(e) $\text{UCL} = 2.11(\overline{R}) = 2.11(7) = 14.77$; $\text{LCL} = 0(\overline{R}) = 0$;

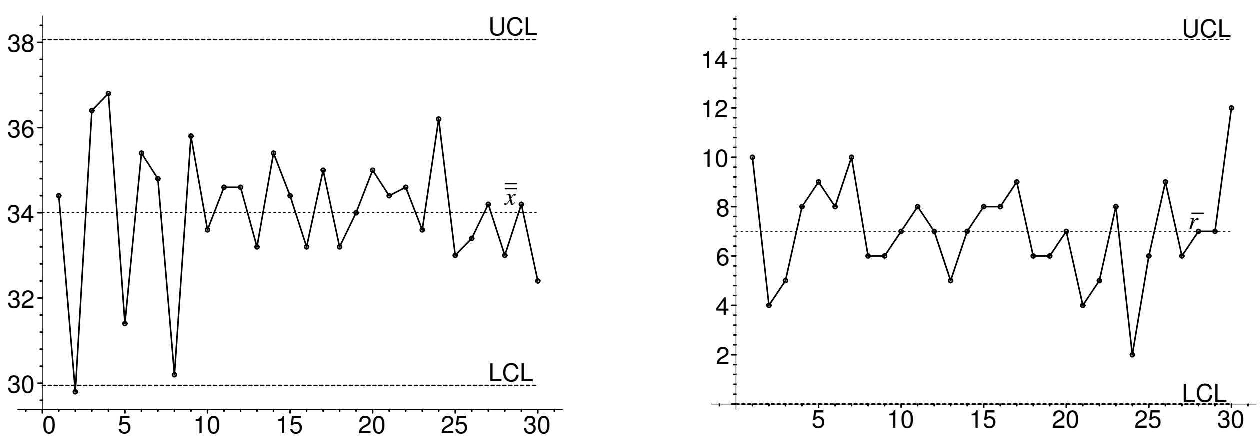


Figure 9.7–8: **(d)** $\overline{x}$-chart using $\overline{R}$ and **(e)** R-chart

(f) Yes.

9.7–10 $LCL = 1.4 - 3\sqrt{1.4} < 0$ so $LCL = 0$;

$UCL = 1.4 + 3\sqrt{1.4} = 4.9496$.

(a) If $\lambda = 3, P(X \geq 5) = 1 - P(X \leq 4) = 1 - 0.815 = 0.185$.

(b) Y is $b(10, 0.185)$. Thus

$$\begin{aligned} P(\text{at least } 1) &= 1 - P(Y = 0) \\ &= 1 - p^0 q^{10} \\ &= 1 - 0.815^{10} \\ &= 0.871. \end{aligned}$$