Solutions Manual to

Introduction to Probability and Statistics for Engineers and Scientists

Original Text by

Sheldon M. Ross

이명규 지음

Myeongkyu Lee

mgklee@kaist.ac.kr

Contents

Chapter 2	Descriptive Statistics	1
Chapter 3	Elements of Probability	3
Chapter 4	Random Variables and Expectation	5
Chapter 5	Special Random Variables	12
Chapter 6	Distributions of Sampling Statistics	18
Chapter 7	Parameter Estimation	19
Chapter 8	Hypothesis Testing	24
Chapter 9	Regression	29
Chapter 10	Analysis of Variance	35

Chapter 2

Descriptive Statistics

2.11. The sample mean of the entire data set is

$$\frac{99 \cdot 120 + 99 \cdot 100}{198} = 110.$$

(a) 50 values are less than or equal to 100, and other 50 values are greater than or equal to 120. The sample median is between 100 and 120.

(b) Nothing can be inferred about the sample mode with current information.

$\blacksquare$

2.14. Let x and y be the unknown values. By definition, we know that

$$\frac{x + y + 102 + 100 + 105}{5} = 104,$$
$$\frac{(x-104)^2 + (y-104)^2 + (102-104)^2 + (100-104)^2 + (105-104)^2}{5-1} = 16.$$

From the first equation, $y = 213 - x$. By substitution,

$$(x-104)^2 + (109-x)^2 = 43,$$
$$x^2 - 213x + 11327 = 0.$$

Therefore, the two values are $\frac{213+\sqrt{61}}{2}$ and $\frac{213-\sqrt{61}}{2}$. $\blacksquare$

2.18. **(a)** By the formulas,

$$x_1 = 3,$$
$$\overline{x}_2 = 3 + \frac{4-3}{2} = \frac{7}{2},$$
$$\overline{x}_3 = \frac{7}{2} + \frac{7-\frac{7}{2}}{3} = \frac{14}{3},$$
$$\overline{x}_4 = \frac{14}{3} + \frac{2-\frac{14}{3}}{4} = 4,$$
$$\overline{x}_5 = 4 + \frac{9-4}{5} = 5,$$

$$\overline{x}_6 = 5 + \frac{6-5}{6} = \frac{31}{6},$$

and

$$\begin{aligned}
s_2^2 &= \frac{0}{1} \cdot 0 + 2\left(\frac{7}{2} - 3\right)^2 = \frac{1}{2}, \\
s_3^2 &= \frac{1}{2} \cdot \frac{1}{2} + 3\left(\frac{14}{3} - \frac{7}{2}\right)^2 = \frac{13}{3}, \\
s_4^2 &= \frac{2}{3} \cdot \frac{13}{3} + 4\left(4 - \frac{14}{3}\right)^2 = \frac{14}{3}, \\
s_5^2 &= \frac{3}{4} \cdot \frac{14}{3} + 5\,(5-4)^2 = \frac{17}{2}, \\
s_6^2 &= \frac{4}{5} \cdot \frac{17}{2} + 6\left(\frac{31}{6} - 5\right)^2 = \frac{209}{30}.
\end{aligned}$$

(b) By usual computation,

$$\begin{aligned}
\overline{x} &= \frac{3+4+7+2+9+6}{6} = \frac{31}{6}, \\
s^2 &= \frac{1}{6-1} \sum_{i=1}^{6} \left(x_i - \frac{31}{6}\right)^2 = \frac{209}{30}.
\end{aligned}$$

(c) By definition,

$$(j+1)\overline{x}_{j+1} = \sum_{i=1}^{j} x_i + x_{j+1} = j\overline{x}_j + x_{j+1} = (j+1)\overline{x}_j + x_{j+1} - \overline{x}_j.$$

Dividing both sides by $j+1$ yields the formula.

■

Chapter 3

Elements of Probability

3.2. Let H and T denote a head and a tail, respectively. The sample space of this experiment is

$$S = \{\text{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT}\}.$$

HHH, HHT, HTH, and THH have more heads than tails. ■

3.11. The proof is by induction.

(i) For the base case $n = 1$, $P(E_1) \leq P(E_1)$.

(ii) For the inductive step, assume the inequality for $n = k$. Then

$$\begin{aligned} P\left(\bigcup_{i=1}^{k+1} E_i\right) &= P\left(\bigcup_{i=1}^{k} E_i \cup E_{k+1}\right) \\ &= P\left(\bigcup_{i=1}^{k} E_i\right) + P(E_{k+1}) - P\left(\bigcup_{i=1}^{k} E_i \cap E_{k+1}\right) \\ &\leq P\left(\bigcup_{i=1}^{k} E_i\right) + P(E_{k+1}) \\ &\leq \sum_{i=1}^{k} P(E_i) + P(E_{k+1}) \\ &= \sum_{i=1}^{k+1} P(E_i). \end{aligned}$$

The inequality also holds for $n = k + 1$.

By induction, Boole's inequality is true for every $n \in \mathbb{N}$. ■

3.30. Let A_n be the event that the first n balls chosen are colored red.
Let B_n be the event that n balls in the urn are colored red.

(a) By Bayes' formula,

$$P(B_2 \mid A_2) = \frac{P(A_2 \mid B_2)P(B_2)}{\sum_{i=0}^{2} P(A_2 \mid B_i)P(B_i)}$$

$$= \frac{1 \cdot \frac{1}{4}}{0 \cdot \frac{1}{4} + (\frac{1}{2})^2 \cdot \frac{1}{2} + 1 \cdot \frac{1}{4}} = \frac{2}{3}.$$

(b) By the law of total probability,

$$\begin{aligned} P(A_3 \mid A_2) = \frac{P(A_3)}{P(A_2)} &= \frac{\sum_{i=0}^{2} P(A_3 \mid B_i)P(B_i)}{\sum_{i=0}^{2} P(A_2 \mid B_i)P(B_i)} \\ &= \frac{0 \cdot \frac{1}{4} + (\frac{1}{2})^3 \cdot \frac{1}{2} + 1 \cdot \frac{1}{4}}{0 \cdot \frac{1}{4} + (\frac{1}{2})^2 \cdot \frac{1}{2} + 1 \cdot \frac{1}{4}} = \frac{5}{6}. \end{aligned}$$

■

3.47. $P(A) = 0.2$, $P(B) = 0.3$, $P(C) = 0.4$

(a) If $P(A \cap B) = 0$, then

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.2 + 0.3 - 0 = 0.5.$$

(b) If $P(A \cap B) = P(A)P(B)$, then

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.2 + 0.3 - 0.2 \cdot 0.3 = 0.44.$$

(c) If A, B, and C are independent, by definition,

$$P(A \cap B \cap C) = P(A)P(B)P(C) = 0.024.$$

(d) If $P(A \cap B) = P(A \cap C) = P(B \cap C) = 0$, then $P(A \cap B \cap C) = 0$.

■

3.50. Given that $P(A) = 0.6$ and

$$P(B \mid A^c) = \frac{P(B \cap A^c)}{P(A^c)} = \frac{P(B \cap A^c)}{1 - 0.6} = 0.1,$$

$P(B \cap A^c) = 0.04$. It follows from

$$P(B) = P(B \cap A) + P(B \cap A^c)$$

that

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = P(A) + P(B \cap A^c) = 0.64.$$

■

Chapter 4

Random Variables and Expectation

4.5. (a) Since f is a probability density function, we must have that

$$\int_{-\infty}^{\infty} f(x)\,dx = \int_0^1 cx^3\,dx = \frac{c}{4} = 1,$$

so $c = 4$.

(b) Hence,

$$P\{0.4 < X < 0.8\} = \int_{0.4}^{0.8} 4x^3\,dx = (0.8)^4 - (0.4)^4 = 0.384.$$

■

4.7. The probability that such a tube in a radio set will have to be replaced within the first 150 hours of operation is

$$P\{X \le 150\} = \int_{-\infty}^{150} f(x)\,dx = \int_{100}^{150} \frac{100}{x^2}\,dx = \frac{1}{3}.$$

Therefore, the probability of our interest is

$$\binom{5}{2}\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)^3 = \frac{80}{243}.$$

■

4.11. Since $X_1, X_2, \cdots, X_n$ are independent uniform random variables,

$$F_M(x) = P\{M \le x\} = P\{X_1 \le x, \cdots, X_n \le x\} = \prod_{i=1}^{n} P\{X_i \le x\} = \prod_{i=1}^{n} x = x^n$$

for $x \in [0,1]$. Therefore, the probability density function of M is

$$f_M(x) = \frac{d}{dx} F_M(x) = nx^{n-1}$$

for $x \in [0,1]$ and $f_M(x) = 0$ elsewhere. ■

4.13. By Equations (4.3.5) and (4.3.6),

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\,dy = \int_x^1 2\,dy = 2(1-x) \text{ for } x \in (0,1),$$
$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\,dx = \int_0^y 2\,dx = 2y \text{ for } y \in (0,1).$$

X and Y are not independent because $f(x,y) \neq f_X(x) f_Y(y)$ for some x and y. ■

4.16. Since X and Y are independent continuous random variables, $f(x,y) = f_X(x)f_Y(y)$ for all x and y.

(a)

$$\begin{aligned} P\{X+Y \le a\} &= \iint_{x+y\le a} f(x,y)\,dxdy \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{a-y} f_X(x)f_Y(y)\,dxdy \\ &= \int_{-\infty}^{\infty} f_Y(y)\left[\int_{-\infty}^{a-y} f_X(x)\,dx\right] dy \\ &= \int_{-\infty}^{\infty} F_X(a-y)f_Y(y)\,dy. \end{aligned}$$

(b)

$$\begin{aligned} P\{X \le Y\} &= \iint_{x\le y} f(x,y)\,dxdy \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{y} f_X(x)f_Y(y)\,dxdy \\ &= \int_{-\infty}^{\infty} f_Y(y)\left[\int_{-\infty}^{y} f_X(x)\,dx\right] dy \\ &= \int_{-\infty}^{\infty} F_X(y)f_Y(y)\,dy. \end{aligned}$$

■

4.20. We assume that $p_Y(y) > 0$ and $f_Y(y) > 0$ for all y. For all x and y,

(a) $p_{X|Y}(x \mid y) = \dfrac{p(x,y)}{p_Y(y)} = p_X(x)$ if and only if $p(x,y) = p_X(x)p_Y(y)$.

(b) $f_{X|Y}(x \mid y) = \dfrac{f(x,y)}{f_Y(y)} = f_X(x)$ if and only if $f(x,y) = f_X(x)f_Y(y)$.

Therefore, each condition is equivalent to independence of X and Y. ■

4.25. **(a)** $E[X]$ should be larger than $E[Y]$ because a student with more students on the same bus is more likely to be selected than one with fewer students.

(b) $E[X] = 40 \cdot \dfrac{40}{148} + 33 \cdot \dfrac{33}{148} + 25 \cdot \dfrac{25}{148} + 50 \cdot \dfrac{50}{148} = \dfrac{2907}{74}$,

$E[Y] = \dfrac{40+33+25+50}{4} = 37.$

■

4.29. Let $M = \max\{X_1, \cdots, X_n\}$ and $N = \min\{X_1, \cdots, X_n\}$. Since $X_1, X_2, \cdots, X_n$ are independent uniform random variables,

$$\begin{aligned}
F_M(x) &= P\{M \le x\} = P\{X_1 \le x, \cdots, X_n \le x\} \\
&= \prod_{i=1}^{n} P\{X_i \le x\} = \prod_{i=1}^{n} x = x^n, \\
F_N(x) &= P\{N \le x\} = 1 - P\{N > x\} = 1 - P\{X_1 > x, \cdots, X_n > x\} \\
&= 1 - \prod_{i=1}^{n} P\{X_i > x\} = 1 - \prod_{i=1}^{n}(1-x) = 1-(1-x)^n.
\end{aligned}$$

for $x \in [0,1]$. Hence, the probability density functions of M and N are

$$f_M(x) = \frac{d}{dx}F_M(x) = \begin{cases} nx^{n-1} & \text{if } x \in (0,1), \\ 0 & \text{otherwise.} \end{cases}$$

$$f_N(x) = \frac{d}{dx}F_N(x) = \begin{cases} n(1-x)^{n-1} & \text{if } x \in (0,1), \\ 0 & \text{otherwise.} \end{cases}$$

(a) $E[M] = \displaystyle\int_0^1 nx^n \, dx = \frac{n}{n+1}$.

(b) By integration by parts,

$$\begin{aligned}
E[N] &= \int_0^1 nx(1-x)^{n-1}\, dx \\
&= \left[nx \cdot \frac{-(1-x)^n}{n} \right]_0^1 - n\int_0^1 \frac{-(1-x)^n}{n}\, dx \\
&= \left[\frac{-(1-x)^{n+1}}{n+1} \right]_0^1 = \frac{1}{n+1}.
\end{aligned}$$

■

4.40. We know that $p_3 = 1 - p_1 - p_2$ and that $E[X] = p_1 + 2p_2 + 3p_3 = 2$.
Hence, $p_1 + 2p_2 + 3(1 - p_1 - p_2) = 2$, so $2p_1 + p_2 = 1$.
$\mathrm{Var}(X) = (1-2)^2 p_1 + (2-2)^2 p_2 + (3-2)^2 p_3 = p_1 + p_3 = 1 - p_2$. Since $p_i \in [0,1]$,

(a) When $(p_1, p_2, p_3) = (\frac{1}{2}, 0, \frac{1}{2})$, $\mathrm{Var}(X)$ attains its maximum 1.

(b) When $(p_1, p_2, p_3) = (0, 1, 0)$, $\mathrm{Var}(X)$ attains its minimum 0.

■

4.45. **(a)** The marginal probability distributions of X_1 and X_2 are

$$P\{X_1 = x_1\} = \begin{cases} \frac{3}{16} & \text{if } x_1 = 0, \\ \frac{1}{8} & \text{if } x_1 = 1, \\ \frac{5}{16} & \text{if } x_1 = 2, \\ \frac{3}{8} & \text{if } x_1 = 3, \end{cases} \qquad P\{X_2 = x_2\} = \begin{cases} \frac{1}{2} & \text{if } x_2 = 1, \\ \frac{1}{2} & \text{if } x_2 = 2. \end{cases}$$

respectively.

(b) $E[X_1] = 0 \cdot \dfrac{3}{16} + 1 \cdot \dfrac{1}{8} + 2 \cdot \dfrac{5}{16} + 3 \cdot \dfrac{3}{8} = \dfrac{15}{8}$.

$E[X_2] = 1 \cdot \dfrac{1}{2} + 2 \cdot \dfrac{1}{2} = \dfrac{3}{2}$.

Recall that $\mathrm{Var}(X) = E\left[X^2\right] - (E[X])^2$. (Equation 4.6.1)

$\mathrm{Var}(X_1) = 0^2 \cdot \dfrac{3}{16} + 1^2 \cdot \dfrac{1}{8} + 2^2 \cdot \dfrac{5}{16} + 3^2 \cdot \dfrac{3}{8} - \left(\dfrac{15}{8}\right)^2 = \dfrac{79}{64}$.

$\mathrm{Var}(X_2) = 1^2 \cdot \dfrac{1}{2} + 2^2 \cdot \dfrac{1}{2} - \left(\dfrac{3}{2}\right)^2 = \dfrac{1}{4}$.

Recall that $\mathrm{Cov}(X_1, X_2) = E[X_1 X_2] - E[X_1]E[X_2]$. (Equation 4.7.1)

$\mathrm{Cov}(X_1, X_2) = 0 \cdot \dfrac{3}{16} + 1 \cdot \dfrac{1}{16} + 2 \cdot \dfrac{1}{4} + 3 \cdot \dfrac{1}{8} + 4 \cdot \dfrac{1}{8} + 6 \cdot \dfrac{1}{4} = \dfrac{47}{16} - \dfrac{15}{8} \cdot \dfrac{3}{2} = \dfrac{1}{8}$.

■

4.49. $\mathrm{Var}(X) = \sigma_x^2$ and $\mathrm{Var}(Y) = \sigma_y^2$.

$$\begin{aligned} 0 \le \mathrm{Var}\left(\frac{X}{\sigma_x} + \frac{Y}{\sigma_y}\right) &= \mathrm{Var}\left(\frac{X}{\sigma_x}\right) + \mathrm{Var}\left(\frac{Y}{\sigma_y}\right) + 2\,\mathrm{Cov}\left(\frac{X}{\sigma_x}, \frac{Y}{\sigma_y}\right) \\ &= \frac{\mathrm{Var}(X)}{\sigma_x^2} + \frac{\mathrm{Var}(Y)}{\sigma_y^2} + \frac{2\,\mathrm{Cov}(X,Y)}{\sigma_x \sigma_y} \\ &= 2 + 2\,\mathrm{Corr}(X,Y). \\ \therefore -1 &\le \mathrm{Corr}(X,Y). \\ 0 \le \mathrm{Var}\left(\frac{X}{\sigma_x} - \frac{Y}{\sigma_y}\right) &= \mathrm{Var}\left(\frac{X}{\sigma_x}\right) + \mathrm{Var}\left(\frac{Y}{\sigma_y}\right) - 2\,\mathrm{Cov}\left(\frac{X}{\sigma_x}, \frac{Y}{\sigma_y}\right) \\ &= \frac{\mathrm{Var}(X)}{\sigma_x^2} + \frac{\mathrm{Var}(Y)}{\sigma_y^2} - \frac{2\,\mathrm{Cov}(X,Y)}{\sigma_x \sigma_y} \\ &= 2 - 2\,\mathrm{Corr}(X,Y). \\ \therefore 1 &\ge \mathrm{Corr}(X,Y). \end{aligned}$$

Therefore, we conclude that $-1 \le \mathrm{Corr}(X,Y) \le 1$.

$$\begin{aligned} \mathrm{Corr}(X,Y) = \pm 1 &\iff \mathrm{Var}\left(\frac{X}{\sigma_x} \mp \frac{Y}{\sigma_y}\right) = 0 \\ &\implies \frac{X}{\sigma_x} \mp \frac{Y}{\sigma_y} = c \quad \text{for some constant } c \\ &\iff Y = \mp c\sigma_y \pm \frac{\sigma_y}{\sigma_x} X. \end{aligned}$$

$b = \pm \dfrac{\sigma_y}{\sigma_x}$ is positive when $\mathrm{Corr}(X,Y) = 1$ and negative when $\mathrm{Corr}(X,Y) = -1$. ■

4.50. For $i \in [n] = \{1, \cdots, n\}$, let

$$X_i = \begin{cases} 1 & \text{if trial } i \text{ results in outcome 1,} \\ 0 & \text{if trial } i \text{ does not result in outcome 1.} \end{cases}$$

$E[X_i] = 1 \cdot p_1 + 0 \cdot (1 - p_1) = p_1$. Similarly, for $j \in [n]$, let

$$Y_j = \begin{cases} 1 & \text{if trial } j \text{ results in outcome 2,} \\ 0 & \text{if trial } j \text{ does not result in outcome 2.} \end{cases}$$

$E[Y_j] = 1 \cdot p_2 + 0 \cdot (1 - p_2) = p_2$. It follows from the definitions of N_1 and N_2 that

$$N_1 = \sum_{i=1}^{n} X_i, \quad N_2 = \sum_{j=1}^{n} Y_j.$$

By Proposition 4.7.2,

$$\mathrm{Cov}(N_1, N_2) = \mathrm{Cov}\left(\sum_{i=1}^{n} X_i, \sum_{j=1}^{n} Y_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{Cov}(X_i, Y_j).$$

Because each trial is independent, X_i and Y_j must be independent if they are from two distinct trials, that is, $i \neq j$. By Theorem 4.7.4, $\mathrm{Cov}(X_i, Y_j) = 0$ if $i \neq j$. Also, observe that

$$(X_i, Y_i) = \begin{cases} (1,0) & \text{if trial } i \text{ results in outcome 1,} \\ (0,1) & \text{if trial } i \text{ results in outcome 2,} \\ (0,0) & \text{if trial } i \text{ results in outcome 3,} \end{cases}$$

so $X_i Y_i = 0$ for each trial $i \in [n]$. Hence,

$$\begin{aligned} \mathrm{Cov}(N_1, N_2) &= \sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{Cov}(X_i, Y_j) \\ &= \sum_{i=1}^{n} \mathrm{Cov}(X_i, Y_i) \\ &= \sum_{i=1}^{n} (E[X_i Y_i] - E[X_i]E[Y_i]) \\ &= -\sum_{i=1}^{n} E[X_i]E[Y_i] \\ &= -\sum_{i=1}^{n} p_1 p_2 = -np_1p_2. \end{aligned}$$

Intuitively, the covariance is negative because $N_1 + N_2 \leq n$; the sum of N_1 and N_2 is bounded above by n. The more trials result in outcome 1, the fewer trials can result in outcome 2, so N_1 and N_2 are negatively correlated. ■

4.54. We first compute the moment generating function $\phi_X(t)$ of X.

$$\begin{aligned}
\phi_X(t) = E\left[e^{tX}\right] &= \int_0^1 e^{tx}\,dx \\
&= \begin{cases} \dfrac{e^t-1}{t} & \text{if } t \neq 0, \\ 1 & \text{if } t = 0 \end{cases} \\
&= \begin{cases} \dfrac{1}{t}\left(\displaystyle\sum_{n=0}^{\infty}\frac{t^n}{n!}-1\right) & \text{if } t \neq 0, \\ 1 & \text{if } t = 0 \end{cases} \\
&= \sum_{n=1}^{\infty}\frac{t^{n-1}}{n!} = \sum_{n=0}^{\infty}\frac{t^n}{(n+1)!}.
\end{aligned}$$

For each $n \in \mathbb{N}$, $\phi_X^{(n)}(0) = \dfrac{n!}{(n+1)!} = \dfrac{1}{n+1}$. By definition, the nth moment of X is

$$E[X^n] = \int_0^1 x^n\,dx = \frac{1}{n+1}.$$

Therefore, we have verified that $\phi_X^{(n)}(0) = E[X^n]$ for each $n \in \mathbb{N}$. ■

4.56. Let X be the random variable with $E[X] = 75$. X takes only nonnegative values.

(a) By Markov's inequality (Proposition 4.9.1),

$$P\{X > 85\} \leq \frac{E[X]}{85} = \frac{15}{17}.$$

(b) By Chebyshev's inequality (Proposition 4.9.2),

$$\begin{aligned}
P\{65 \leq X \leq 85\} &= P\{|X - E[X]| \leq 10\} \\
&= 1 - P\{|X - E[X]| \geq 10\} \\
&\geq 1 - \frac{\mathrm{Var}(X)}{10^2} \\
&= 1 - \frac{1}{4} = \frac{3}{4}.
\end{aligned}$$

(c) For n test scores $X_1, \cdots, X_n$ of n students taking her final examination, let $\overline{X} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} X_i$ be the class average. Since X_i's are independent and identically distributed,

$$\begin{aligned}
E\left[\overline{X}\right] &= \frac{1}{n}\sum_{i=1}^{n} E[X_i] = E[X], \\
\mathrm{Var}\left(\overline{X}\right) &= \frac{1}{n^2}\sum_{i=1}^{n} \mathrm{Var}\left(X_i\right) = \frac{\mathrm{Var}(X)}{n}.
\end{aligned}$$

By Chebyshev's inequality,

$$\begin{aligned}
P\left\{\left|\overline{X}-E[X]\right| \leq 5\right\} &= 1-P\left\{\left|\overline{X}-E\left[\overline{X}\right]\right| \geq 5\right\} \\
&\geq 1-\frac{\operatorname{Var}\left(\overline{X}\right)}{5^2} \\
&= 1-\frac{1}{n} \\
&\geq 0.9.
\end{aligned}$$

Therefore, $n \geq 10$.

■

Chapter 5

Special Random Variables

5.1. Let X be the number of components independently in working condition. Then $X \sim B(4, 0.6)$. Since the system can function adequately if $X \geq 2$, the probability is

$$P\{X \geq 2\} = \binom{4}{2}\left(\frac{3}{5}\right)^2\left(\frac{2}{5}\right)^2 + \binom{4}{3}\left(\frac{3}{5}\right)^3\left(\frac{2}{5}\right) + \binom{4}{4}\left(\frac{3}{5}\right)^4 = \frac{513}{625}.$$

■

5.5. A 4-engine plane can operate if at least 2 engines function, while a 2-engine plane can if at least 1 engine functions. We find values of the probability $p \in [0, 1]$ such that

$$\begin{gathered}
\binom{4}{2}p^2(1-p)^2 + \binom{4}{3}p^3(1-p) + \binom{4}{4}p^4 > \binom{2}{1}p(1-p) + \binom{2}{2}p^2 \\
p^2\left[6(1-p)^2 + 4p(1-p) + p^2\right] > p[2(1-p) + p] \\
p\left(3p^3 - 8p^2 + 7p - 2\right) = p(p-1)^2(3p-2) > 0.
\end{gathered}$$

Therefore, $\frac{2}{3} < p < 1$. ■

5.10. $X \sim B(n, p)$. Poisson approximation: $P\{X = k\} \approx \frac{e^{-np}(np)^k}{k!}$.

(a) $P\{X = 2\} = \binom{10}{2}\left(\frac{1}{10}\right)^2\left(\frac{9}{10}\right)^8 = 0.1937102445.$

By approximation, $P\{X = 2\} \approx \frac{e^{-1}}{2!} \approx 0.1839397206.$

(b) $P\{X = 0\} = \binom{10}{0}\left(\frac{9}{10}\right)^{10} = 0.3486784401.$

By approximation, $P\{X = 0\} \approx \frac{e^{-1}}{2!} \approx 0.3678794412.$

(c) $P\{X = 4\} = \binom{9}{4}\left(\frac{1}{5}\right)^4\left(\frac{4}{5}\right)^5 = 0.066060288.$

By approximation, $P\{X = 4\} \approx \frac{e^{-1.8}(1.8)^4}{4!} \approx 0.072301734.$

■

5.11. Let X be the number of prizes I will win. Given that $X \sim B(50, 0.01)$, Recall that, for a large n and a small p, we can approximate

$$P\{X = k\} \approx e^{-\lambda} \frac{\lambda^k}{k!}$$

where $\lambda = np = 0.5$.

(a) $P\{X \geq 1\} = 1 - P\{X = 0\} = 1 - \left(\frac{99}{100}\right)^{50} \approx 1 - e^{-1/2}$.

(b) $P\{X = 1\} = \binom{50}{1} \left(\frac{99}{100}\right)^{49} \frac{1}{100} \approx \frac{e^{-1/2}}{2}$.

(c) $P\{X \geq 2\} = 1 - P\{X \leq 1\} = 1 - \left(\frac{99}{100}\right)^{50} - \binom{50}{1} \left(\frac{99}{100}\right)^{49} \frac{1}{100} \approx 1 - \frac{3e^{-1/2}}{2}$.

■

5.12. Let A be the event that an individual who tries the drug for a year has 0 colds in that time. Let B be the event that the drug is beneficial for him or her. By Bayes' formula,

$$\begin{aligned} P(B \mid A) &= \frac{P(A \mid B)P(B)}{P(A \mid B)P(B) + P(A \mid B^c)P(B^c)} \\ &= \frac{e^{-2} \cdot 3/4}{e^{-2} \cdot 3/4 + e^{-3} \cdot 1/4} = \frac{3e}{3e + 1}. \end{aligned}$$

■

5.17. If X is a Poisson random variable with $E[X] = \lambda$, then we know that $X \sim \text{Poisson}(\lambda)$. Hence, $P\{X = i\} = \frac{e^{-\lambda}\lambda^i}{i!}$ for $i \in \mathbb{Z}_{\geq 0}$. Observe that

$$\frac{P\{X = i + 1\}}{P\{X = i\}} = \frac{\lambda}{i + 1} \geq 1 \iff i + 1 \leq \lambda.$$

Therefore, $P\{X = i\}$ first increases and then decreases as i increases. To be more precise, $P\{X = \lambda - 1\} = P\{X = \lambda\}$ if $\lambda \in \mathbb{N}$. It reaches its maximum value when $i = \lfloor \lambda \rfloor$ anyway. ■

5.20. We shall assume that $p \in (0, 1]$.

(a) $X = k$ if the first $k - 1$ trials are all failures (with probability $1 - p$) and the kth trial is a success (with probability p). Thus, $P\{X = k\} = (1 - p)^{k-1}p$ for $k \in \mathbb{N}$.

(b) We know that

$$\frac{1}{1 - x} = \sum_{k=0}^{\infty} x^k$$

for $|x| < 1$. It can be shown that this power series allows term-by-term differentiation, so

$$\frac{1}{(1 - x)^2} = \sum_{k=1}^{\infty} kx^{k-1}$$

for $|x| < 1$. Hence,

$$E[X] = \sum_{k=1}^{\infty} kP\{X = k\} = p \sum_{k=1}^{\infty} k(1-p)^{k-1} = p \cdot \frac{1}{p^2} = \frac{1}{p}.$$

(c) In order for Y to equal k, $r-1$ successes must result in the first $k-1$ trials and a success must be the outcome of the kth trial. Thus,

$$P\{Y = k\} = \binom{k-1}{r-1} p^{r-1}(1-p)^{k-r} \cdot p = \binom{k-1}{r-1} p^r (1-p)^{k-r}$$

for $k \geq r$.

(d) Write $Y = \sum_{i=1}^{r} Y_i$ where Y_i is the number of trials needed to go from a total of $i-1$ to a total of i successes. Then each Y_i is a geometric random variable with probability p. Therefore, by part (b),

$$E[Y] = E\left[\sum_{i=1}^{r} Y_i\right] = \sum_{i=1}^{r} E[Y_i] = \sum_{i=1}^{r} \frac{1}{p} = \frac{r}{p}.$$

■

5.21. Since U is uniformly distributed on $(0,1)$, its cumulative distribution function is

$$F_U(x) = \begin{cases} 0 & \text{if } x \leq 0, \\ x & \text{if } 0 < x < 1, \\ 1 & \text{if } x \geq 1. \end{cases}$$

For $a, b \in \mathbb{R}$ with $a < b$, let $V = a + (b-a)U$. Then $U = \dfrac{V-a}{b-a}$. The cumulative distribution function of V is

$$F_U\left(\frac{y-a}{b-a}\right) = \begin{cases} 0 & \text{if } \dfrac{y-a}{b-a} \leq 0, \\ \dfrac{y-a}{b-a} & \text{if } 0 < \dfrac{y-a}{b-a} < 1, \\ 1 & \text{if } \dfrac{y-a}{b-a} \geq 1 \end{cases}$$

$$F_V(y) = \begin{cases} 0 & \text{if } y \leq a, \\ \dfrac{y-a}{b-a} & \text{if } a < y < b, \\ 1 & \text{if } y \geq b. \end{cases}$$

Hence, the probability density function of V is

$$f_V(y) = \frac{dF_V(y)}{dy} = \begin{cases} \dfrac{1}{b-a} & \text{if } a < y < b, \\ 0 & \text{otherwise.} \end{cases}$$

Therefore, $V = a + (b-a)U$ is uniform on (a, b). ■

5.27. Let X be the output of a bulb. We know that $X \sim N\left(\mu, \sigma^2\right)$ where $\mu = 2000$ and $\sigma = 85$. We find the value of L such that

$$\begin{aligned} P\{X \geq L\} &= P\left\{Z = \frac{X-\mu}{\sigma} \geq \frac{L-2000}{85}\right\} \\ &= 1 - \Phi\left(\frac{L-2000}{85}\right) = 0.95 \end{aligned}$$

where $\Phi(x)$ is the standard normal distribution function,

$$\Phi(z) = P\{Z \leq z\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-x^2/2}\, dx,$$

for $Z \sim N\left(0, 1^2\right)$. Therefore,

$$L = 85\Phi^{-1}(0.05) + 2000 \approx 85 \cdot (-1.645) + 2000 = 1860.175.$$

■

5.29. Let $I = \int_{-\infty}^{\infty} e^{-x^2/2}\, dx$.

(a) Let $t = (x-\mu)/\sigma$. Then $dx = \sigma dt$.

$$\begin{aligned} 1 &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} e^{-(x-\mu)^2/(2\sigma^2)}\, dx \\ &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} e^{-t^2/2}\, \sigma dt \\ &= \frac{I}{\sqrt{2\pi}} \end{aligned}$$

is equivalent to $I = \sqrt{2\pi}$.

(b) We evaluate the double integral by means of a change of variables to polar coordinates. Note that

$$J = \begin{vmatrix} \frac{dx}{dr} & \frac{dx}{d\theta} \\ \frac{dy}{dr} & \frac{dy}{d\theta} \end{vmatrix} = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} = r,$$

so $dxdy = |J| drd\theta = rdrd\theta$. Thus,

$$\begin{aligned} I^2 &= \int_{-\infty}^{\infty} e^{-x^2/2}\, dx \int_{-\infty}^{\infty} e^{-y^2/2}\, dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-(x^2+y^2)/2}\, dxdy \\ &= \int_{0}^{2\pi} \int_{0}^{\infty} e^{-r^2/2}\, rdrd\theta \\ &= \int_{0}^{\infty} re^{-r^2/2}\, dr \int_{0}^{2\pi} d\theta \\ &= 2\pi \left[-e^{-r^2/2}\right]_0^{\infty} = 2\pi. \end{aligned}$$

Since I is an integral of a positive function over $\mathbb{R}$, $I \geq 0$. Therefore, $I = \sqrt{2\pi}$.

■

5.30. If $\log X \sim N(\mu, \sigma^2)$,

$$P\{X \leq x\} = \begin{cases} 0 & \text{if } x \leq 0, \\ P\{\log X \leq \log x\} & \text{if } x > 0 \end{cases}$$
$$= \begin{cases} 0 & \text{if } x \leq 0, \\ \dfrac{1}{\sqrt{2\pi}\sigma} \displaystyle\int_{-\infty}^{\log x} e^{-\frac{(t-\mu)^2}{2\sigma^2}}\, dt & \text{if } x > 0. \end{cases}$$

■

5.37. Repair time (in hours): $X \sim \mathrm{Exp}(1)$.

(a) $P\{X > 2\} = \displaystyle\int_2^\infty e^{-x}\, dx = e^{-2}$.

(b) $P\{X \geq 3 \mid X > 2\} = \dfrac{P\{X \geq 3\}}{P\{X > 2\}} = \dfrac{\int_3^\infty e^{-x}\, dx}{\int_2^\infty e^{-x}\, dx} = \dfrac{1}{e}$.

■

5.41. Recall that a Poisson process has independent and stationary increments. By Proposition 5.6.2,

$$P\{N(t) = k\} = e^{-\lambda t} \frac{(\lambda t)^k}{k!}$$

for $k \in \mathbb{Z}_{\geq 0}$. If t is in years, here $\lambda = 5$.

(a) $P\{N(1/2) \geq 2\} = 1 - P\{N(1/2) = 0\} - P\{N(1/2) = 1\} = 1 - 7/2e^{-5/2}$.

(b) The number of events that occur in disjoint time intervals are independent (independent increments), so $P\{N(3/4) = 0\} = e^{-15/4}$.

(c)

$$P\{N(3/4) \geq 4 \mid N(1/2) \geq 2\} = \frac{P\{N(3/4) \geq 4 \wedge N(1/2) \geq 2\}}{P\{N(1/2) \geq 2\}}$$
$$= \frac{P\{N(1/2) = 2 \wedge N(3/4) - N(1/2) \geq 2\}}{P\{N(1/2) \geq 2\}} + \frac{P\{N(1/2) = 3 \wedge N(3/4) - N(1/2) \geq 1\}}{P\{N(1/2) \geq 2\}} + \frac{P\{N(1/2) \geq 4\}}{P\{N(1/2) \geq 2\}}$$
$$= \frac{P\{N(1/2) = 2\}P\{N(1/4) \geq 2\}}{P\{N(1/2) \geq 2\}} + \frac{P\{N(1/2) = 3\}P\{N(1/4) \geq 1\}}{P\{N(1/2) \geq 2\}} + \frac{P\{N(1/2) \geq 4\}}{P\{N(1/2) \geq 2\}}$$
$$= \frac{1}{1 - 7/2e^{-5/2}} \left[\frac{e^{-5/2}(5/2)^2}{2} \left(1 - e^{-5/4} - 5/4e^{-5/4}\right) \right.$$

$$+\frac{e^{-5/2}(5/2)^3}{6}\left(1-e^{-5/4}\right)+1-\sum_{k=0}^{3}\frac{e^{-5/2}(5/2)^k}{k!}\Bigg].$$

■

Chapter 6

Distributions of Sampling Statistics

6.4. Let X_i be the number I win for each bet. $X_1, \cdots, X_n$ are independent and identically distributed random variables with $P\{X_i = 35\} = 1/38$ and $P\{X_i = -1\} = 37/38$.

Let Y_n be the number of bets for which I win 35 among n bets; then $Y_n \sim B(n, 1/38)$. We have

$$\sum_{i=1}^{n} X_i = 35Y_n - (n - Y_n) > 0 \iff Y_n > \frac{n}{36}.$$

If n is large, then approximately $Y_n \sim N\left(\frac{n}{38}, \left(\frac{\sqrt{37n}}{38}\right)^2\right)$ by the central limit theorem.

(a) $P\left\{Y_{34} > \frac{34}{36}\right\} = 1 - P\{Y_{34} = 0\} = 1 - \left(\frac{37}{38}\right)^{34} \approx 0.596.$

(b) $P\left\{Y_{1000} > \frac{1000}{36}\right\} \approx P\left\{Z \geq \frac{\frac{1000}{36} - \frac{1000}{38}}{\frac{\sqrt{37000}}{38}}\right\} \approx 1 - \Phi(0.29) \approx 0.3859.$

(c) $P\left\{Y_{100000} > \frac{100000}{36}\right\} \approx P\left\{Z \geq \frac{\frac{100000}{36} - \frac{100000}{38}}{\frac{\sqrt{3700000}}{38}}\right\} \approx 1 - \Phi(2.89) \approx 0.0019.$

■

6.20. By Theorem 6.5.1, $\frac{9}{4}S_1^2 \sim \chi_9^2$ and $\frac{4}{2}S_2^2 \sim \chi_4^2$.

$$P\left\{S_1^2 < S_2^2\right\} = P\left\{\frac{S_1^2}{S_2^2} < 1\right\} = P\left\{\frac{4\chi_9^2/9}{2\chi_4^2/4} < 1\right\} = P\left\{F_{9,4} < \frac{1}{2}\right\}.$$

■

Chapter 7

Parameter Estimation

7.1. The likelihood function is

$$f(x_1, \cdots, x_n \mid \theta) = \begin{cases} \exp\left(n\theta - \sum_{i=1}^n x_i\right) & \text{if } \theta \le x_i \text{ for each } 1 \le i \le n, \\ 0 & \text{otherwise.} \end{cases}$$

f is an increasing function of θ subject to $\theta \le x_i$ for each $1 \le i \le n$. Hence, $\hat{\theta} = \min\{x_1, \cdots x_n\}$ maximizes f. Therefore, the maximum likelihood estimator of θ is $\min\{X_1, \cdots, X_n\}$. ■

7.5. Let $\sigma^2 > 0$ be the common variance. We have

$$\begin{aligned} f(x_1, \cdots, x_n, y_1, \cdots, y_n, w_1, \cdots, w_n \mid \mu_1, \mu_2) &= \prod_{i=1}^n f_{X_i}(x_i) \prod_{i=1}^n f_{Y_i}(y_i) \prod_{i=1}^n f_{W_i}(w_i) \\ &= \left(\frac{1}{\sqrt{2\pi}|\sigma|}\right)^{3n} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n \left[(x_i - \mu_1)^2 + (y_i - \mu_2)^2 + (w_i - \mu_1 - \mu_2)^2\right]\right). \end{aligned}$$

Taking logs,

$$\begin{aligned} g(\mu_1, \mu_2) &= \log f(x_1, \cdots, x_n, y_1, \cdots, y_n, w_1, \cdots, w_n \mid \mu_1, \mu_2) \\ &= -3n \log\left(\sqrt{2\pi}|\sigma|\right) - \frac{1}{2\sigma^2} \sum_{i=1}^n \left[(x_i - \mu_1)^2 + (y_i - \mu_2)^2 + (w_i - \mu_1 - \mu_2)^2\right]. \end{aligned}$$

To find critical points of g, we solve

$$\begin{aligned} \frac{\partial g}{\partial \mu_1} &= \frac{1}{\sigma^2} \sum_{i=1}^n [(x_i - \mu_1) + (w_i - \mu_1 - \mu_2)] = 0, \\ \frac{\partial g}{\partial \mu_2} &= \frac{1}{\sigma^2} \sum_{i=1}^n [(y_i - \mu_2) + (w_i - \mu_1 - \mu_2)] = 0, \end{aligned}$$

that is,

$$\begin{bmatrix} 2n & n \\ n & 2n \end{bmatrix} \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n x_i + \sum_{i=1}^n w_i \\ \sum_{i=1}^n y_i + \sum_{i=1}^n w_i \end{bmatrix}.$$

Observe that $\dfrac{\partial^2 g}{\partial \mu_1^2} = -\dfrac{2n}{\sigma^2} < 0$ and

$$\begin{vmatrix} \dfrac{\partial^2 g}{\partial \mu_1^2} & \dfrac{\partial^2 g}{\partial \mu_1 \partial \mu_2} \\ \dfrac{\partial^2 g}{\partial \mu_2 \partial \mu_1} & \dfrac{\partial^2 g}{\partial \mu_2^2} \end{vmatrix} = \left(-\frac{2n}{\sigma^2}\right)\left(-\frac{2n}{\sigma^2}\right) - \left(-\frac{n}{\sigma^2}\right)\left(-\frac{n}{\sigma^2}\right) = \frac{3n^2}{\sigma^4} > 0.$$

By the second derivative test, the maximum likelihood estimates of μ_1 and μ_2 are

$$\begin{aligned} \begin{bmatrix} \hat{\mu}_1 \\ \hat{\mu}_2 \end{bmatrix} &= \frac{1}{3n^2} \begin{bmatrix} 2n & -n \\ -n & 2n \end{bmatrix} \begin{bmatrix} \sum_{i=1}^n x_i + \sum_{i=1}^n w_i \\ \sum_{i=1}^n y_i + \sum_{i=1}^n w_i \end{bmatrix} \\ &= \frac{1}{3n} \begin{bmatrix} 2\sum_{i=1}^n x_i + \sum_{i=1}^n w_i - \sum_{i=1}^n y_i \\ 2\sum_{i=1}^n y_i + \sum_{i=1}^n w_i - \sum_{i=1}^n x_i \end{bmatrix}. \end{aligned}$$

Therefore, the maximum likelihood estimators of μ_1 and μ_2 are

$$\frac{1}{3n}\left(2\sum_{i=1}^n X_i + \sum_{i=1}^n W_i - \sum_{i=1}^n Y_i\right) \text{ and } \frac{1}{3n}\left(2\sum_{i=1}^n Y_i + \sum_{i=1}^n W_i - \sum_{i=1}^n X_i\right),$$

respectively. ■

7.9. $\overline{x} = 11.48$, $\sigma = 0.08$, and $n = 10$.

(a) A 95 percent confidence interval for the PCB level of this fish is

$$\left(11.48 - \frac{1.96 \cdot 0.08}{\sqrt{10}},\, 11.48 + \frac{1.96 \cdot 0.08}{\sqrt{10}}\right) \approx (11.4304, 11.5296).$$

(b) A 95 percent lower confidence interval is

$$\left(-\infty,\, 11.48 + \frac{1.645 \cdot 0.08}{\sqrt{10}}\right) \approx (-\infty, 11.5216).$$

(c) A 95 percent upper confidence interval is

$$\left(11.48 - \frac{1.645 \cdot 0.08}{\sqrt{10}},\, \infty\right) \approx (11.4384, \infty).$$

■

7.18. Assume that the sample is normal. $\overline{x} \approx 133.222$, $s \approx 10.213$, and $n = 18$. $t_{0.05,17} = 1.740$, $t_{0.025,17} = 2.110$ according to Table A3.

(a) A 95 percent confidence interval estimate of the average IQ score of all students at the university is

$$\left(133.222 - \frac{2.110 \cdot 10.213}{\sqrt{18}},\, 133.222 + \frac{2.110 \cdot 10.213}{\sqrt{18}}\right) \approx (128.143, 138.301).$$

(b) A 95 percent lower confidence interval estimate is

$$\left(-\infty, 133.222 + \frac{1.740 \cdot 10.213}{\sqrt{18}}\right) \approx (-\infty, 137.411).$$

(c) A 95 percent upper confidence interval estimate is

$$\left(133.222 - \frac{1.740 \cdot 10.213}{\sqrt{18}}, \infty\right) \approx (129.033, \infty).$$

■

7.26. $\overline{x} = 2062.75$, $s \approx 104.343$, and $n = 20$. $t_{0.05,19} = 1.729$, $t_{0.025,19} = 2.093$, and $t_{0.005,19} = 2.861$ according to Table A3.

(a) A 95 percent two-sided confidence interval for the mean number of steps is

$$\left(2062.75 - \frac{2.093 \cdot 104.343}{\sqrt{20}}, 2062.75 + \frac{2.093 \cdot 104.343}{\sqrt{20}}\right) \approx (2013.92, 2111.58).$$

(b) A 99 percent two-sided confidence interval for the mean number of steps is

$$\left(2062.75 - \frac{2.861 \cdot 104.343}{\sqrt{20}}, 2062.75 + \frac{2.861 \cdot 104.343}{\sqrt{20}}\right) \approx (1996.00, 2129.50).$$

(c) A 95 percent upper confidence interval is

$$\left(2062.75 - \frac{1.729 \cdot 104.343}{\sqrt{20}}, \infty\right) \approx (2022.41, \infty).$$

Therefore, $v \approx 2022.41$.

■

7.38. The pooled estimate of σ^2 is

$$s_p^2 = \frac{4s_1^2 + 2s_2^2}{6} = \frac{4 \cdot 2.5 + 2 \cdot 7}{6} = 4.$$

Thus, an estimate of σ is $s_p = 2$. ■

7.39. $\mu = 3.180$, $s^2 \approx 6.484 \times 10^{-5}$, and $n = 8$.

(a) By Equation (7.2.3), the maximum likelihood estimate $\hat{\sigma}$ of σ is

$$\left[\frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2\right]^{1/2} \approx 7.706 \times 10^{-3}.$$

(b) Because each $(X_i - \mu)/\sigma \sim \mathcal{N}(0,1)$ is independent, it follows that

$$\sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma} \right)^2 \sim \chi_n^2.$$

According to Table A2, $\chi^2_{0.05,8} = 15.507$ and $\chi^2_{0.95,8} = 2.733$. A 90 percent confidence interval for σ^2 is

$$\left(\frac{7 \cdot 6.484 \times 10^{-5}}{15.507}, \frac{7 \cdot 6.484 \times 10^{-5}}{2.733} \right) \approx (3.063 \times 10^{-5}, 1.738 \times 10^{-4}).$$

Taking square roots, a 90 percent confidence interval for σ is

$$(5.535 \times 10^{-3}, 1.318 \times 10^{-2}).$$

■

7.41. We assume that the samples are normal with a common variance. $\overline{x} = 3358.1$, $\overline{y} = 3130.4$, and $n = m = 10$. According to Table A3, $t_{0.025,18} = 2.101$ and $t_{0.05,18} = 1.734$.

$$s_p = \sqrt{\frac{(n-1)s_1^2 + (m-1)s_2^2}{n+m-2}} \approx \sqrt{\frac{124419 + 17729}{2}} \approx 266.597.$$

(a) By Equation (7.4.4), a 95 percent two-sided confidence interval for $\mu_1 - \mu_2$ is

$$\left(\overline{x} - \overline{y} - t_{0.025,18} s_p \sqrt{\frac{1}{n} + \frac{1}{m}}, \overline{x} - \overline{y} + t_{0.025,18} s_p \sqrt{\frac{1}{n} + \frac{1}{m}} \right) \approx (-22.79, 478.19).$$

(b) A 95 percent one-sided upper confidence interval for $\mu_1 - \mu_2$ is

$$\left(\overline{x} - \overline{y} - t_{0.05,18} s_p \sqrt{\frac{1}{n} + \frac{1}{m}}, \infty \right) \approx (20.96, \infty).$$

(c) A 95 percent one-sided upper confidence interval for $\mu_1 - \mu_2$ is

$$\left(-\infty, \overline{x} - \overline{y} + t_{0.05,18} s_p \sqrt{\frac{1}{n} + \frac{1}{m}} \right) \approx (-\infty, 434.44).$$

■

7.44. $\overline{x} = 532.1$, $\overline{y} = 548.6$, and $n = m = 10$. According to Table A3, $t_{0.005,18} = 2.878$.

$$s_p = \sqrt{\frac{(n-1)s_1^2 + (m-1)s_2^2}{n+m-2}} \approx \sqrt{\frac{2932.8 + 1191.6}{2}} \approx 45.412.$$

By Equation (7.4.4), a 99 percent two-sided confidence interval for $\mu_1 - \mu_2$ is

$$\left(\overline{x} - \overline{y} - t_{0.005,18} s_p \sqrt{\frac{1}{n} + \frac{1}{m}}, \overline{x} - \overline{y} + t_{0.005,18} s_p \sqrt{\frac{1}{n} + \frac{1}{m}} \right) \approx (-74.949, 41.949).$$

■

7.55. Let p be the probability that a person contracting lung cancer will die within 5 years. Let $\{X_1, \cdots, X_{100}\}$ be a random sample from a Bernoulli distribution with parameter p; then $E[X_i] = p$. We know that $\sum_{i=1}^{100} x_i = 67$.

(a) By Example 7.7a, an unbiased estimate of p is $\hat{p} = \frac{1}{100}\sum_{i=1}^{100} x_i = 0.67$.

(b) A 95 percent two-sided confidence interval for p is

$$\left(\hat{p} - 1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}},\, \hat{p} + 1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right).$$

We find some $n \in \mathbb{N}$ such that

$$1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} < 0.02 \iff n \geq 2124.$$

Therefore, an additional sample of at least 2024 data is required to be 95 percent confident that the error in estimating the probability in part (a) is less than .02.

■

7.63. If X is an exponential random variable with $E[X] = 1/\lambda$, then $X \sim \mathrm{Exp}(\lambda)$. Hence,

$$f(x \mid \lambda) = \lambda e^{-\lambda x}\vec{1}_{[0,\infty)}(x),$$

where

$$\vec{1}_A(x) = \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \notin A. \end{cases}$$

We know that $\frac{1}{20}\sum_{i=1}^{20} x_i = 4.6$. The posterior density function of λ is

$$\begin{aligned} f(\lambda \mid x_1, \cdots x_{20}) &= \frac{f(x_1, \cdots, x_{20} \mid \lambda)g(\lambda)}{\int_0^\infty f(x_1, \cdots, x_{20} \mid \lambda)g(\lambda)\, d\lambda}\vec{1}_{(0,\infty)}(\lambda) \\ &= \frac{\lambda^{20}\exp\left(-\lambda\sum_{i=1}^{20} x_i\right)\frac{1}{2}e^{-\lambda}\lambda^2}{\int_0^\infty \lambda^{20}\exp\left(-\lambda\sum_{i=1}^{20} x_i\right)\frac{1}{2}e^{-\lambda}\lambda^2\, d\lambda}\vec{1}_{(0,\infty)}(\lambda) \\ &= \frac{\lambda^{22}e^{-93\lambda}}{\int_0^\infty \lambda^{22}e^{-93\lambda}\, d\lambda}\vec{1}_{(0,\infty)}(\lambda). \end{aligned}$$

Therefore, the Bayes estimate of λ is

$$\begin{aligned} E\left[\lambda \mid x_1, \cdots x_{20}\right] &= \frac{\int_0^\infty \lambda^{23}e^{-93\lambda}\, d\lambda}{\int_0^\infty \lambda^{22}e^{-93\lambda}\, d\lambda} \\ &= \frac{1}{\int_0^\infty \lambda^{22}e^{-93\lambda}\, d\lambda}\left(\left[\lambda^{23}\frac{e^{-93\lambda}}{-93}\right]_0^\infty - \int_0^\infty 23\lambda^{22}\frac{e^{-93\lambda}}{-93}\, d\lambda\right) \\ &= \frac{23}{93}. \end{aligned}$$

■

Chapter 8

Hypothesis Testing

8.2. We test the null hypothesis $H_0 : \mu = 32$ against the alternative hypothesis $H_1 : \mu \neq 32$. We first compute the test statistic

$$T = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{30.4 - 32}{4/\sqrt{25}} = -2.$$

$|T| > z_{0.025} = 1.96$, so we reject H_0 at the 5% level of significance. ■

8.5. We test the null hypothesis $H_0 : \mu \geq 200$ against the alternative hypothesis $H_1 : \mu < 200$. We first compute the test statistic

$$T = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{199.1125 - 200}{5/\sqrt{8}} \approx -0.502.$$

The p-value is $P\{Z \leq t\} \approx \Phi(-0.502) \approx 0.308$ where $Z \sim N(0,1)$. It is greater than $\alpha = 0.05$ and $\alpha = 0.1$, so we accept H_0 at both levels of significance. ■

8.10. We test the null hypothesis $H_0 : \mu \geq 7.6$ against the alternative hypothesis $H_1 : \mu < 7.6$. We first compute the test statistic

$$T = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{7.2 - 7.6}{1.2/\sqrt{16}} = -\frac{4}{3}.$$

The p-value is $P\{Z \leq -4/3\} = P\{Z \geq 4/3\} \approx 1 - \Phi(1.33) = 0.0918$ where $Z \sim N(0,1)$. It is greater than $\alpha = 0.01$ and $\alpha = 0.05$, so we accept H_0 at both levels of significance. ■

8.14. We test the null hypothesis $H_0 : \mu = 24$ against the alternative hypothesis $H_1 : \mu \neq 24$. We first compute the test statistic

$$T = \frac{\overline{X} - \mu_0}{S/\sqrt{n}} = \frac{22.5 - 24}{3.1/\sqrt{36}} = -\frac{90}{31} \approx -2.903.$$

$|T| > 2.045 > t_{0.025,35}$, so we reject H_0 at the 5% level of significance. ■

8.20. We test the null hypothesis $H_0 : \mu \geq 30$ against the alternative hypothesis $H_1 : \mu < 30$. We first compute the test statistic

$$T = \frac{\overline{X} - \mu_0}{S/\sqrt{n}} = \frac{26.4 - 30}{2\sqrt{\frac{46}{15}\frac{1}{\sqrt{10}}}} = -9\sqrt{\frac{3}{23}}.$$

The p-value is $P\left\{T_9 < -9\sqrt{\dfrac{3}{23}}\right\}$ where T_9 is a t-random variable with 9 degrees of freedom. We accept H_0 if the significance level α is less than the p-value. ■

8.29. We test the null hypothesis $H_0 : \mu_1 \leq \mu_2$ against the alternative hypothesis $H_1 : \mu_1 > \mu_2$. We first compute the test statistic

$$T = \frac{\overline{X} - \overline{Y}}{\sqrt{\sigma_1^2/n + \sigma_2^2/m}}.$$

Let $Z \sim N(0,1)$.

(a) When $\sigma_2 = 5$, $T \approx 2.65$, so the p-value is approximately $P\{Z > 2.65\} = 1 - \Phi(2.65) = 0.0041$.

(b) When $\sigma_2 = 10$, $T \approx 2.10$, so the p-value is approximately $P\{Z > 2.10\} = 1 - \Phi(2.10) = 0.0179$.

(c) When $\sigma_2 = 20$, $T \approx 1.33$, so the p-value is approximately $P\{Z > 1.33\} = 1 - \Phi(1.33) = 0.0918$.

■

8.31. We test the null hypothesis $H_0 : \mu_1 = \mu_2$ against the alternative hypothesis $H_1 : \mu_1 \neq \mu_2$. We first compute the test statistic

$$T = \frac{\overline{X} - \overline{Y}}{\sqrt{S_p^2(1/n + 1/m)}} \approx 0.4371,$$

where S_p^2 is the pooled estimator of the common variance σ^2 given by

$$S_p^2 = \frac{(n-1)S_1^2 + (m-1)S_2^2}{n+m-2}.$$

The p-value is $2P\{T_{11} \geq 0.4371\} \approx 0.67$ where T_{11} is a t-random variable with 11 degrees of freedom. ■

8.44. By Theorem 6.5.1, $(n-1)S^2/\sigma^2 \sim \chi^2_{n-1}$. Suppose that H_0 is true. Then we have

$$\frac{(n-1)S^2}{\sigma_0^2} \leq \frac{(n-1)S^2}{\sigma^2}.$$

Hence,

$$P\left\{\frac{(n-1)S^2}{\sigma_0^2} > \chi^2_{\alpha,n-1}\right\} \leq P\left\{\frac{(n-1)S^2}{\sigma^2} > \chi^2_{\alpha,n-1}\right\} = \alpha.$$

Therefore, we

$$\text{accept } H_0 \text{ if } \frac{(n-1)S^2}{\sigma_0^2} \leq \chi^2_{\alpha,n-1},$$
$$\text{reject } H_0 \text{ if } \frac{(n-1)S^2}{\sigma_0^2} > \chi^2_{\alpha,n-1}$$

at the significance level α. ■

8.45. Because each $(X_i - \mu)/\sigma \sim \mathcal{N}(0,1)$ is independent, it follows that

$$\sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma} \right)^2 \sim \chi_n^2.$$

Suppose that H_0 is true. Then we have

$$\sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma_0} \right)^2 \le \sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma} \right)^2.$$

Hence,

$$P\left\{ \sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma_0} \right)^2 > \chi_{\alpha,n}^2 \right\} \le P\left\{ \sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma} \right)^2 > \chi_{\alpha,n}^2 \right\} = \alpha.$$

Therefore, we

$$\text{accept } H_0 \text{ if } \sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma_0} \right)^2 \le \chi_{\alpha,n}^2,$$
$$\text{reject } H_0 \text{ if } \sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma_0} \right)^2 > \chi_{\alpha,n}^2$$

at the significance level α. ■

8.50. We assume that $\{X_i\}_{i=1}^n$ and $\{Y_j\}_{j=1}^m$ are independent samples from two normal populations $\mathcal{N}(\mu_x, \sigma_x^2)$ and $\mathcal{N}(\mu_y, \sigma_y^2)$, respectively. Let

$$S_x^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left(X_i - \overline{X} \right)^2 \text{ and } S_y^2 = \frac{1}{m-1} \sum_{j=1}^{m} \left(Y_j - \overline{Y} \right)^2.$$

By Theorem 6.5.1, $(n-1)S_x^2/\sigma_x^2 \sim \chi_{n-1}^2$ and $(m-1)S_y^2/\sigma_y^2 \sim \chi_{m-1}^2$. When H_0 is true,

$$\frac{S_x^2/\sigma_x^2}{S_y^2/\sigma_y^2} \ge \frac{S_x^2}{S_y^2} \sim F_{n-1,m-1}.$$

Hence,

$$P\left\{ \frac{S_x^2}{S_y^2} > F_{\alpha,n-1,m-1} \right\} \le P\left\{ \frac{S_x^2/\sigma_x^2}{S_y^2/\sigma_y^2} > F_{\alpha,n-1,m-1} \right\} = \alpha.$$

Therefore, we

$$\text{accept } H_0 \text{ if } \frac{S_x^2}{S_y^2} \le F_{\alpha,n-1,m-1},$$
$$\text{reject } H_0 \text{ if } \frac{S_x^2}{S_y^2} > F_{\alpha,n-1,m-1}$$

at the significance level α. ■

8.60. Let p_1 and p_2 be the probabilities that the first and the second treatments are effective, respectively.

(a) We test the null hypothesis $H_0 : p_1 = p_2$ against the alternative hypothesis $H_1 : p_1 \neq p_2$ using the Fisher-Irwin test. We compute $P\{X \leq 39\} \approx 0.649$ and $P\{X \geq 39\} \approx 0.475$, where

$$P\{X = i\} = \frac{\binom{72}{i}\binom{84}{83-i}}{\binom{156}{83}}$$

for $0 \leq i \leq 83$. The p-value is considerably large, so we accept H_0 in general.

(b) Let $Y \sim \mathrm{B}(156, p)$. We test the null hypothesis $H_0 : p = 0.5$ against the alternative hypothesis $H_1 : p \neq 0.5$. Suppose that H_0 is true. We compute the p-value $2P\{Y \geq 84\} \approx 0.379$. Hence, we accept H_0 in general. Therefore, the fact is consistent with the claim that the determination of the treatment to be given to each patient was made in a totally random fashion.

■

8.63. Let $X_1 \sim \mathrm{B}(n_1, p_1)$ and $X_2 \sim \mathrm{B}(n_2, p_2)$. When n_1 and n_2 are large, by the central limit theorem, $X_1 \dot{\sim} \mathcal{N}(n_1p_1, n_1p_1(1-p_1))$ and $X_2 \dot{\sim} \mathcal{N}(n_2p_2, n_2p_2(1-p_2))$. Hence,

$$\begin{aligned}
\frac{X_1}{n_1} &\dot{\sim} \mathcal{N}\left(p_1, \frac{p_1(1-p_1)}{n_1}\right), \\
\frac{X_2}{n_2} &\dot{\sim} \mathcal{N}\left(p_2, \frac{p_2(1-p_2)}{n_2}\right), \\
\frac{X_1}{n_1} - \frac{X_2}{n_2} &\dot{\sim} \mathcal{N}\left(p_1 - p_2, \frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}\right), \\
\frac{\frac{X_1}{n_1} - \frac{X_2}{n_2} - (p_1 - p_2)}{\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}} &\dot{\sim} \mathcal{N}(0,1).
\end{aligned}$$

We test the null hypothesis $H_0 : p_1 = p_2$ against the alternative hypothesis $H_1 : p_1 \neq p_2$. When H_0 is true, their common value can be best estimated by $(X_1 + X_2)/(n_1 + n_2)$ because it is the proportion of the total $X_1 + X_2$ successes to the total $n_1 + n_2$ trials.

$$\frac{X_1/n_1 - X_2/n_2}{\sqrt{p(1-p)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \approx \frac{X_1/n_1 - X_2/n_2}{\sqrt{\frac{X_1+X_2}{n_1+n_2}\left(1 - \frac{X_1+X_2}{n_1+n_2}\right)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \dot{\sim} \mathcal{N}(0,1).$$

Therefore, we reject H_0 if the test statistic is greater than $z_{\alpha/2}$. ■

8.64. We first compute the test statistic

$$T = \frac{X_1/n_1 - X_2/n_2}{\sqrt{\frac{X_1+X_2}{n_1+n_2}\left(1 - \frac{X_1+X_2}{n_1+n_2}\right)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{\frac{39}{72} - \frac{44}{84}}{\sqrt{\frac{83}{156} \cdot \frac{73}{156}\left(\frac{1}{72} + \frac{1}{84}\right)}} \approx 0.208.$$

The p-value is $2P\{Z \geq 0.208\} \approx 0.835$. Therefore, we accept H_0 in general. ■

8.68. Let $X \sim \text{Poisson}(\lambda)$. We test the null hypothesis $H_0 : \lambda = 52$ against the alternative hypothesis $H_1 : \lambda \neq 52$. Suppose that H_0 is true. Since $\sum_{i=1}^{8} X_i \sim \text{Poisson}(416) \dot{\sim} \mathcal{N}(416, 416)$ and $46 + 62 + 60 + 58 + 47 + 50 + 59 + 49 = 431$, we compute the p-value

$$2P\left\{\sum_{i=1}^{8} X_i \geq 431\right\} \approx 2P\left\{Z \geq \frac{431 - 416}{\sqrt{416}}\right\} \approx 0.462.$$

Therefore, we accept H_0 in general. ■

8.71. To investigate the relationship between smoking and heart attacks, all the other factors should be controlled. Thus, the scientist should choose samples such that each smoker is associated with a nonsmoker of almost the same age. ■

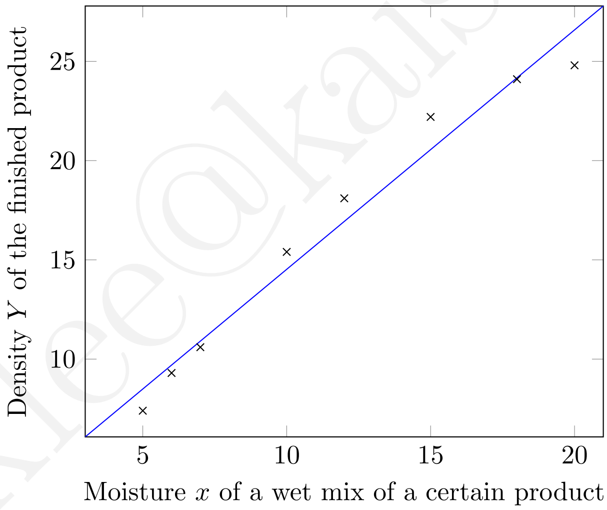
Chapter 9

Regression

9.1. $n = 8$, $\overline{x} = \dfrac{93}{8}$, $\overline{Y} = \dfrac{1319}{80}$, $S_{xx} = \dfrac{1775}{8}$, and $S_{xY} = \dfrac{21413}{80}$. Hence, the least squares estimators are

$$A = \overline{Y} - B\overline{x} = \frac{43727}{17750} \text{ and } B = \frac{S_{xY}}{S_{xx}} = \frac{21413}{17750}.$$

The estimated regression line is $y = A + Bx = \dfrac{43727 + 21413x}{17750} \approx 2.463 + 1.206x$.



■

9.8. By Proposition 9.2.1,

$$\begin{aligned} A &= \overline{Y} - B\overline{x} \\ &= \frac{1}{n}\sum_{i=1}^{n} Y_i - \frac{\overline{x}}{S_{xx}}\sum_{i=1}^{n}(x_i - \overline{x})Y_i \\ &= \sum_{i=1}^{n}\left[\frac{1}{n} - \frac{\overline{x}(x_i - \overline{x})}{S_{xx}}\right] Y_i. \end{aligned}$$

Because $Y_1, \cdots, Y_n$ are independent normal random variables with common variance σ^2,

$$\begin{aligned}
\mathrm{Var}(A) &= \sum_{i=1}^{n} \left[\frac{1}{n} - \frac{\overline{x}(x_i - \overline{x})}{S_{xx}}\right]^2 \mathrm{Var}(Y_i) \\
&= \sigma^2 \sum_{i=1}^{n} \left[\frac{1}{n^2} - \frac{2\overline{x}(x_i - \overline{x})}{nS_{xx}} + \frac{\overline{x}^2(x_i - \overline{x})^2}{S_{xx}^2}\right] \\
&= \sigma^2 \left(\frac{1}{n} + \frac{\overline{x}^2}{S_{xx}}\right) \\
&= \frac{\sigma^2}{nS_{xx}} \left(S_{xx} + n\overline{x}^2\right) \\
&= \frac{\sigma^2}{nS_{xx}} \sum_{i=1}^{n} x_i^2.
\end{aligned}$$

■

9.10. By definition,

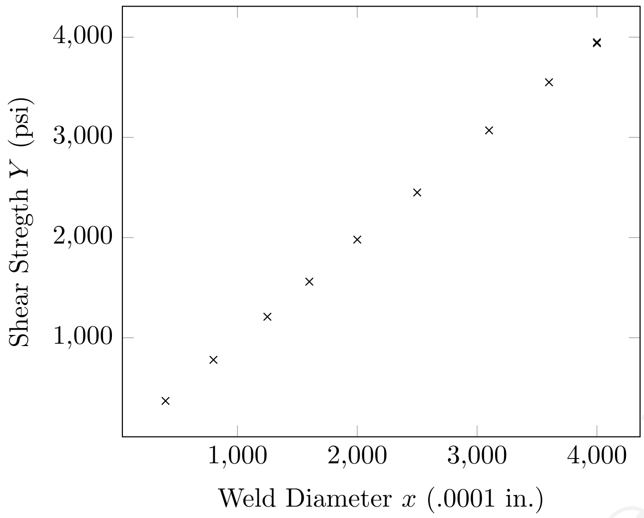
$$\begin{aligned}
SS_R &= \sum_{i=1}^{n} (Y_i - A - Bx_i)^2 \\
&= \sum_{i=1}^{n} \left[Y_i - \overline{Y} - \frac{S_{xY}}{S_{xx}}(x_i - \overline{x})\right]^2 \\
&= \sum_{i=1}^{n} \left[\left(Y_i - \overline{Y}\right)^2 - \frac{2S_{xY}}{S_{xx}}(x_i - \overline{x})\left(Y_i - \overline{Y}\right) + \frac{S_{xY}^2}{S_{xx}^2}(x_i - \overline{x})^2\right] \\
&= S_{YY} - \frac{2S_{xY}^2}{S_{xx}} + \frac{S_{xY}^2}{S_{xx}} \\
&= \frac{S_{xx}S_{YY} - S_{xY}^2}{S_{xx}}.
\end{aligned}$$

■

9.26. $n = 10$, $\overline{x} = 2325$, $\overline{Y} = 2286$, $S_{xx} = 15686250$, $S_{xY} = 15573000$, $S_{YY} = 15461440$, and

$$SS_R = \frac{S_{xx}S_{YY} - S_{xY}^2}{S_{xx}} \approx 872.369.$$

(a)



(b) $A = \overline{Y} - B\overline{x} \approx -22.214$ and $B = \dfrac{S_{xY}}{S_{xx}} \approx 0.9928$.

(c) We test the null hypothesis $H_0 : \beta = 1$ against the alternative hypothesis $H_1 : \beta \neq 1$. By Equation 9.4.2,

$$\sqrt{\frac{(n-2)S_{xx}}{SS_R}}(B - \beta) \sim t_{n-2}.$$

If H_0 is true, then the test statistic is

$$\sqrt{\frac{8S_{xx}}{SS_R}}(B - 1) = -2.738 < -t_{0.025,8} = -2.306.$$

Hence, we reject H_0 at the 5% level of significance.

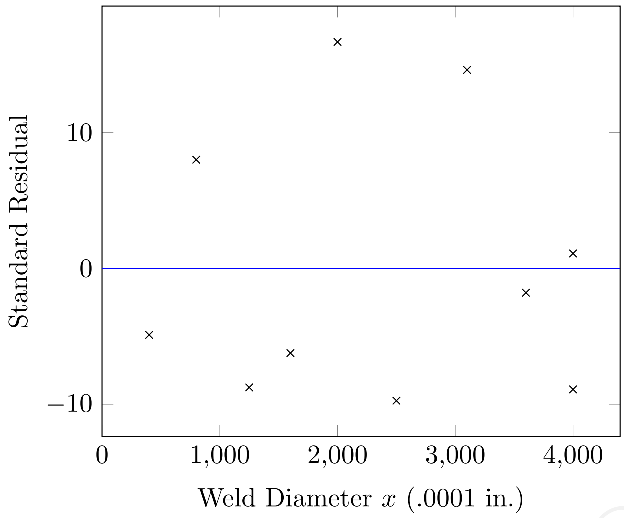
(d) $E[\alpha + 2500\beta] = A + 2500B = 2459.737$.

(e) We use the **Prediction Interval for a Response at the Input Level x_0** on page 381:

$$A + Bx_0 \pm t_{a/2,n-2}\sqrt{\left[\frac{n+1}{n} + \frac{(x_0 - \overline{x})^2}{S_{xx}}\right]\frac{SS_R}{n-2}}.$$

Such a prediction interval is $(2186.282, 2236.801)$.

(f)



(g) As indicated both by its scatter diagram and the random nature of its standardized residuals, the plot appears to fit the straight-line model quite well.

■

9.30. $n = 8$, $\overline{x} = 24.6875$, $\overline{Y} = 126.5$, $S_{xx} = 222.38875$, $S_{xY} = 425.65$, $S_{YY} = 1084$, and

$$SS_R = \frac{S_{xx}S_{YY} - S_{xY}^2}{S_{xx}} \approx 269.310.$$

We use the **Prediction Interval for a Response at the Input Level $\boldsymbol{x_0}$** on page 381:

$$A + Bx_0 \pm t_{a/2,n-2}\sqrt{\left[\frac{n+1}{n} + \frac{(x_0 - \overline{x})^2}{S_{xx}}\right]\frac{SS_R}{n-2}}.$$

Such a prediction interval is $(111.564, 146.460)$. ■

9.32. Taking logs,

$$\log S = \log A - m \log N \iff \log N = \frac{\log A}{m} - \frac{\log S}{m}.$$

Let $Y = \log N$, $x = \log S$, $\alpha = (\log A)/m$, and $\beta = -1/m$. Then we obtain the usual regression equation

$$Y = \alpha + \beta x + e.$$

$n = 10$, $\overline{x} \approx 3.698$, $\overline{Y} \approx 3.286$, $S_{xx} \approx 0.2902$, and $S_{xY} \approx -4.068$. Hence, the least squares estimators are $A' = \overline{Y} - B'\overline{x} \approx 55.124$ and $B' = \dfrac{S_{xY}}{S_{xx}} \approx -14.017$. Thus, we can estimate

$$m = -\frac{1}{B'} \approx 0.07134 \text{ and } A = \exp\left(-\frac{A'}{B'}\right) \approx 51.038.$$

■

9.47. Let

$$\vec{X} = \begin{bmatrix} 1 & 7.1 & 0.68 & 4 \\ 1 & 9.9 & 0.64 & 1 \\ 1 & 3.6 & 0.58 & 1 \\ 1 & 9.3 & 0.21 & 3 \\ 1 & 2.3 & 0.89 & 5 \\ 1 & 4.6 & 0.00 & 8 \\ 1 & 0.2 & 0.37 & 5 \\ 1 & 5.4 & 0.11 & 3 \\ 1 & 8.2 & 0.87 & 4 \\ 1 & 7.1 & 0.00 & 6 \\ 1 & 4.7 & 0.76 & 0 \\ 1 & 5.4 & 0.87 & 8 \\ 1 & 1.7 & 0.52 & 1 \\ 1 & 1.9 & 0.31 & 3 \\ 1 & 9.2 & 0.19 & 5 \end{bmatrix} \text{ and } \vec{Y} = \begin{bmatrix} 41.53 \\ 63.75 \\ 16.38 \\ 45.54 \\ 15.52 \\ 28.55 \\ 5.65 \\ 25.02 \\ 52.49 \\ 38.05 \\ 30.76 \\ 39.69 \\ 17.59 \\ 13.22 \\ 50.98 \end{bmatrix}.$$

$\vec{X}$ has full rank, so $\vec{X}^T\vec{X}$ is invertible. By Equation 9.10.3, the least squares estimators are given by

$$\begin{bmatrix} B_0 \\ B_1 \\ B_2 \\ B_3 \end{bmatrix} = \vec{B} = \left(\vec{X}^T\vec{X}\right)^{-1}\vec{X}^T\vec{Y} = \begin{bmatrix} -2.8278 \\ 5.3707 \\ 9.8157 \\ 0.4482 \end{bmatrix}.$$

Hence,

$$SS_R = \vec{Y}^T\vec{Y} - \vec{B}^T\vec{X}^T\vec{Y} = 201.9692.$$

(a) The estimated regression plane is

$$Y = -2.8278 + 5.3707x_1 + 9.8157x_2 + 0.4482x_3.$$

(b) By Equation 9.10.6,

$$\mathrm{Cov}(\vec{B}) = \sigma^2\left(\vec{X}^T\vec{X}\right)^{-1}$$

where σ^2 is the variance of a normal random error e whose mean is 0. Thus,

$$\mathrm{Var}(B_0) = \sigma^2\left[\left(\vec{X}^T\vec{X}\right)^{-1}\right]_{11} = 0.7472\sigma^2.$$

We test the null hypothesis $H_0 : \beta_0 = 0$ against the alternative hypothesis $H_1 : \beta_0 \neq 0$.

■

9.50. With $100(1-a)$ percent confidence,

(a) **Confidence Interval Estimator of $\boldsymbol{\alpha + \beta x_0}$** on page 378:

$$A + Bx_0 \pm t_{a/2,n-2}\sqrt{\left[\frac{1}{n} + \frac{(x_0-\overline{x})^2}{S_{xx}}\right]\frac{SS_R}{n-2}}.$$

(b) Prediction Interval for a Response at the Input Level x_0 on page 381:

$$A + Bx_0 \pm t_{a/2,n-2}\sqrt{\left[\frac{n+1}{n} + \frac{(x_0-\overline{x})^2}{S_{xx}}\right]\frac{SS_R}{n-2}}.$$

$A + Bx_0$ is an unbiased estimator of the mean response $\alpha + \beta x_0$ such that

$$A + Bx_0 \sim \mathcal{N}\left(\alpha + \beta x_0, \sigma^2\left[\frac{1}{n} + \frac{(x_0-\overline{x})^2}{S_{xx}}\right]\right). \tag{9.4.4}$$

Let Y denote the future response whose input level is x_0. We know that

$$Y \sim \mathcal{N}\left(\alpha + \beta x_0, \sigma^2\right).$$

To obtain a prediction interval, we consider the distribution of $Y - A - Bx_0$. Since the future response Y is independent of the past observations $Y_1, \cdots, Y_n$ that were used to determine A and B, it follows that Y is also independent of $A + Bx_0$. Hence,

$$Y - A - Bx_0 \sim \mathcal{N}\left(0, \sigma^2\left[\frac{1}{n} + \frac{(x_0-\overline{x})^2}{S_{xx}}\right] + \sigma^2\right).$$

$\mathrm{Var}(Y - A - Bx_0) = \mathrm{Var}(A + Bx_0) + \mathrm{Var}(Y)$. $\mathrm{Var}(A+Bx_0)$ is the variance of the estimator of the mean response. $\mathrm{Var}(Y) = \sigma^2$ is the variance of a single observation. Therefore, for the same data, a prediction interval for a future response always contains the corresponding confidence interval for the mean response. ■

Chapter 10

Analysis of Variance

10.3. t-tests are to examine the hypothesis that **two** normal populations have the same mean value. To test the hypothesis $H_0 : \mu_1 = \mu_2 = \cdots = \mu_m$, one may think of running t-tests on all of the $\binom{m}{2}$ pairs of samples. However, multiple testing may increase the chance of making a Type I error. In general, the probability that m *independent* tests produce at least one Type I error is $1 - (1 - \alpha)^m$. Because multiple t-tests here use the same data several times, they may not be independent and would have complicated dependencies. Therefore, the analysis of variance is the preferred method to test H_0. ■

10.6. We test the null hypothesis $H_0 : \mu_1 = \mu_2$ against the alternative hypothesis $H_1 : \mu_1 \neq \mu_2$. $m = 2$ and $n = 10$. By definition,

$$X_{..} = \frac{1}{m}\sum_{i=1}^{m} \overline{X}_i = \frac{17.5 + 17.87}{2} = 17.685,$$

$$SS_b = n\sum_{i=1}^{m}\left(\overline{X}_i - X_{..}\right)^2 = 0.6845,$$

$$SS_W = \sum_{i=1}^{m}\sum_{j=1}^{n}(X_{ij} - X_{i.})^2 = 934.681.$$

We compute the test statistic

$$F = \frac{SS_b/(m-1)}{SS_W/(nm-m)} = 0.01318.$$

$F < F_{0.05,1,18} = 4.41$, and the p-value is 0.9099. Therefore, we accept H_0 at the 5% level of significance. ■

10.8. By definition,

$$S_i^2 = \frac{1}{n-1}\sum_{j=1}^{n}(X_{ij} - X_{i.})^2.$$

Hence,

$$SS_W = \sum_{i=1}^{m}\sum_{j=1}^{n}(X_{ij} - X_{i.})^2 = \sum_{i=1}^{m}(n-1)S_i^2.$$

■

10.12. $m = 3$ and $n = 12$. By Exercise 10.8,

$$SS_W = (n-1)\sum_{i=1}^{m} S_i^2 = 11 \cdot (145 + 138 + 150) = 4763.$$

By definition,

$$X_{..} = \frac{1}{m}\sum_{i=1}^{m} \overline{X}_i = 34,$$
$$SS_b = n\sum_{i=1}^{m} \left(\overline{X}_i - X_{..}\right)^2 = 12 \cdot \left[(-2)^2 + 6^2 + (-4)^2\right] = 672.$$

(a) We test the null hypothesis $H_0 : \mu_1 = \mu_2 = \mu_3$ against the alternative hypothesis H_1 : not all the means are equal. We compute the test statistic

$$F = \frac{SS_b/(m-1)}{SS_W/(nm-m)} = \frac{1008}{433}.$$

$F < F_{0.05,2,33} = 3.29$, and the p-value is 0.1133. Therefore, we accept H_0 at the 5% level of significance.

(b) Let

$$W = \frac{1}{\sqrt{n}} C(m, nm-m, \alpha)\sqrt{\frac{SS_W}{nm-m}} = \frac{C(3, 33, 0.05)\sqrt{433}}{6}.$$

By the T-method,

$$P\{\mu_1 - \mu_2 \in (X_{1.} - X_{2.} - W, X_{1.} - X_{2.} + W) = (-8 - W, -8 + W)\} = 0.95,$$
$$P\{\mu_1 - \mu_3 \in (X_{1.} - X_{3.} - W, X_{1.} - X_{3.} + W) = (2 - W, 2 + W)\} = 0.95,$$
$$P\{\mu_2 - \mu_3 \in (X_{2.} - X_{3.} - W, X_{2.} - X_{3.} + W) = (10 - W, 10 + W)\} = 0.95.$$

■

10.13. $m = 3$ and $n = 5$. By definition,

$$X_{..} = \frac{1}{m}\sum_{i=1}^{m} \overline{X}_i = \frac{33 + 34 + 32.8}{3} = \frac{499}{15},$$
$$SS_b = n\sum_{i=1}^{m} \left(\overline{X}_i - X_{..}\right)^2 = 5 \cdot \left[\left(-\frac{4}{15}\right)^2 + \left(\frac{11}{15}\right)^2 + \left(-\frac{7}{15}\right)^2\right] = \frac{62}{15},$$
$$SS_W = \sum_{i=1}^{m}\sum_{j=1}^{n} (X_{ij} - X_{i.})^2 = \frac{744}{5}.$$

(a) We test the null hypothesis $H_0 : \mu_1 = \mu_2 = \mu_3$ against the alternative hypothesis H_1 : not all the means are equal. We compute the test statistic

$$F = \frac{SS_b/(m-1)}{SS_W/(nm-m)} = \frac{1}{6}.$$

$F < F_{0.05,2,12} = 3.89$, and the p-value is 0.8484. Therefore, we accept H_0 at the 5% level of significance.

(b) Let

$$W = \frac{1}{\sqrt{n}} C(m, nm - m, \alpha) \sqrt{\frac{SS_W}{nm - m}} \approx 5.937.$$

By the T-method,

$$\begin{aligned}
P\{\mu_1 - \mu_2 \in (X_{1.} - X_{2.} - W,\ X_{1.} - X_{2.} + W) \approx (-6.937, 4.937)\} = 0.95, \\
P\{\mu_1 - \mu_3 \in (X_{1.} - X_{3.} - W,\ X_{1.} - X_{3.} + W) \approx (-5.737, 6.137)\} = 0.95, \\
P\{\mu_2 - \mu_3 \in (X_{2.} - X_{3.} - W,\ X_{2.} - X_{3.} + W) \approx (-4.737, 7.137)\} = 0.95.
\end{aligned}$$

■

10.16. Because $x_{..}$ is the average of finitely many terms, we can change the order of summation. By definition,

$$\begin{aligned}
x_{..} &= \frac{1}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} x_{ij} = \frac{1}{nm} \sum_{j=1}^{n} \sum_{i=1}^{m} x_{ij} \\
&= \frac{1}{m} \sum_{i=1}^{m} \left(\sum_{j=1}^{n} \frac{x_{ij}}{n} \right) = \frac{1}{m} \sum_{i=1}^{m} x_{i.} \\
&= \frac{1}{n} \sum_{j=1}^{n} \left(\sum_{i=1}^{m} \frac{x_{ij}}{m} \right) = \frac{1}{n} \sum_{j=1}^{n} x_{.j}.
\end{aligned}$$

■