

Solutions Manual to Accompany

Statistics for Business and Economics

Eighth Edition

David R. Anderson

University of Cincinnati

Dennis J. Sweeney

University of Cincinnati

Thomas A. Williams

Rochester Institute of Technology

© 2002 by South-Western/Thomson Learning™
Cincinnati, Ohio

Contents

Chapter

1. Data and Statistics
2. Descriptive Statistics: Tabular and Graphical Approaches
3. Descriptive Statistics: Numerical Methods
4. Introduction to Probability
5. Discrete Probability Distributions
6. Continuous Probability Distributions
7. Sampling and Sampling Distributions
8. Interval Estimation
9. Hypothesis Testing
10. Statistical Inference about Means and Proportions With Two Populations
11. Inferences about Population Variances
12. Tests of Goodness of Fit and Independence
13. Analysis of Variance and Experimental Design
14. Simple Linear Regression
15. Multiple Regression
16. Regression Analysis: Model Building
17. Index Numbers
18. Forecasting
19. Nonparametric Methods
20. Statistical Methods for Quality Control
21. Sample Survey

Preface

The purpose of *Statistics for Business and Economics* is to provide students, primarily in the fields of business administration and economics, with a sound conceptual introduction to the field of statistics and its many applications. The text is applications-oriented and has been written with the needs of the nonmathematician in mind.

The solutions manual furnishes assistance by identifying learning objectives and providing detailed solutions for all exercises in the text.

Acknowledgements

We would like to provide special recognition to Catherine J. Williams for her efforts in preparing the solutions manual.

David R. Anderson

Dennis J. Sweeney

Thomas A. Williams

Chapter 1

Data and Statistics

Learning Objectives

1. Obtain an appreciation for the breadth of statistical applications in business and economics.
2. Understand the meaning of the terms elements, variables, and observations as they are used in statistics.
3. Obtain an understanding of the difference between qualitative, quantitative, crosssectional and time series data.
4. Learn about the sources of data for statistical analysis both internal and external to the firm.
5. Be aware of how errors can arise in data.
6. Know the meaning of descriptive statistics and statistical inference.
7. Be able to distinguish between a population and a sample.
8. Understand the role a sample plays in making statistical inferences about the population.

Solutions:

1. Statistics can be referred to as numerical facts. In a broader sense, statistics is the field of study dealing with the collection, analysis, presentation and interpretation of data.
2.
 - a. 9
 - b. 4
 - c. Country and room rate are qualitative variables; number of rooms and the overall score are quantitative variables.
 - d. Country is nominal; room rate is ordinal; number of rooms and overall score are ratio.
3.
 - a. Average number of rooms = $808/9 = 89.78$ or approximately 90 rooms
 - b. 2 of 9 are located in England; approximately 22%
 - c. 4 of 9 have a room rate of \$\$; approximately 44%
4.
 - a. 10
 - b. Fortune 500 largest U.S. industrial corporations
 - c. Average revenue = $\$142,275.9/10 = \$14,227.59$ million
 - d. Using the sample average, statistical inference would let us estimate the average revenue for the population of 500 corporations as \$14,227.59 million.
5.
 - a. 3
 - b. Industry code is qualitative; revenues and profit are quantitative.
 - c. Average profit = $10,652.1/10 = \$1065.21$ million
 - d. 8 of 10 had a profit over \$100 million; 80%
 - e. 1 of 10 had an industry code of 3; 10%
6. Questions a, c, and d are quantitative.
Questions b and e are qualitative.
7.
 - a. The data are numeric and the variable is qualitative.
 - b. Nominal
8.
 - a. 2,013
 - b. Qualitative
 - c. Percentages since we have qualitative data

$$i = \left(\frac{25}{100} \right) 15 = 3.75$$

$$Q_1 \text{ (4th position)} = 15$$

For Q_3 ,

$$i = \left(\frac{75}{100} \right) 15 = 11.25$$

$$Q_3 \text{ (12th position)} = 19$$

c. For the 70th percentile,

$$i = \left(\frac{70}{100} \right) 15 = 10.5$$

Rounding up we see the 70th percentile is in position 11.

$$70\text{th percentile} = 18$$

14. a. $\bar{x} = \frac{\sum x_i}{n} = \frac{12,780}{20} = \639

b. $\bar{x} = \frac{\sum x_i}{n} = \frac{1976}{20} = 98.8$ pictures

c. $\bar{x} = \frac{\sum x_i}{n} = \frac{2204}{20} = 110.2$ minutes

d. This is not an easy choice because it is a multicriteria problem. If price was the only criterion, the lowest price camera (Fujifilm DX-10) would be preferred. If maximum picture capacity was the only criterion, the maximum picture capacity camera (Kodak DC280 Zoom) would be preferred. But, if battery life was the only criterion, the maximum battery life camera (Fujifilm DX10) would be preferred. There are many approaches used to select the best choice in a multicriteria situation. These approaches are discussed in more specialized books on decision analysis.

15. Range $20 - 10 = 10$

10, 12, 16, 17, 20

$$i = \frac{25}{100}(5) = 1.25$$

$$Q_1 \text{ (2nd position)} = 12$$

Middle Managers: $\sigma^2 = \sum (x - \bar{x})^2 f(x) = 1.1344$

d. Executives: $\sigma = 1.1169$

Middle Managers: $\sigma = 1.0651$

- e. The senior executives have a higher average score: 4.05 vs. 3.84 for the middle managers. The executives also have a slightly higher standard deviation.

22. a. $E(x) = \sum x f(x)$

$$= 300(.20) + 400(.30) + 500(.35) + 600(.15) = 445$$

The monthly order quantity should be 445 units.

b. Cost: $445 @ \$50 = \$22,250$

Revenue: $300 @ \$70 = \underline{21,000}$
\$ 1,250 Loss

23. a. Laptop: $E(x) = .47(0) + .45(1) + .06(2) + .02(3) = .63$

Desktop: $E(x) = .06(0) + .56(1) + .28(2) + .10(3) = 1.42$

b. Laptop: $\text{Var}(x) = .47(-.63)^2 + .45(.37)^2 + .06(1.37)^2 + .02(2.37)^2 = .4731$

Desktop: $\text{Var}(x) = .06(-1.42)^2 + .56(-.42)^2 + .28(.58)^2 + .10(1.58)^2 = .5636$

- c. From the expected values in part (a), it is clear that the typical subscriber has more desktop computers than laptops. There is not much difference in the variances for the two types of computers.

24. a. Medium $E(x) = \sum x f(x)$

$$= 50(.20) + 150(.50) + 200(.30) = 145$$

Large: $E(x) = \sum x f(x)$

$$= 0(.20) + 100(.50) + 300(.30) = 140$$

Medium preferred.

- b. Medium

x	$f(x)$	$x - \bar{x}$	$(x - \bar{x})^2$	$(x - \bar{x})^2 f(x)$
50	.20	-95	9025	1805.0
150	.50	5	25	12.5
200	.30	55	3025	<u>907.5</u>
				$\uparrow^2 = 2725.0$

Large

y	$f(y)$	$y - \bar{x}$	$(y - \bar{x})^2$	$(y - \bar{x})^2 f(y)$
0	.20	-140	19600	3920

$$-1.6 \pm 2.57 \quad (-4.17 \text{ to } +.97)$$

10. a. $\bar{x}_1 = 17.54 \quad \bar{x}_2 = 15.36$

$$\bar{x}_1 - \bar{x}_2 = 17.54 - 15.36 = \$2.18 \text{ per hour greater for union workers.}$$

b. $s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} = \frac{14(2.24)^2 + 19(1.99)^2}{15 + 20 - 2} = 4.41$

c. $\bar{x}_1 - \bar{x}_2 \pm t_{r/2} s_{\bar{x}_1 - \bar{x}_2}$

$$s_{\bar{x}_1 - \bar{x}_2} = \sqrt{4.41 \left(\frac{1}{15} + \frac{1}{20} \right)} = 0.72$$

$$17.54 - 15.36 \pm t_{r/2} (.72)$$

$$2.18 \pm t_{r/2} (.72)$$

Note: Values for $t_{.025}$ are not listed for 33 degrees of freedom; for 30 d.f. $t_{.025} = 2.042$ and for 40 d.f. $t_{.025} = 2.021$. We will use the more conservative value of 2.042 as an approximation.

$$2.18 \pm 2.042 (.72)$$

$$2.18 \pm 1.47 \text{ or } 0.71 \text{ to } 3.65$$

11. a. $s_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = \sqrt{\frac{(5.2)^2}{40} + \frac{(6)^2}{50}} = 1.18$

$$z = \frac{(25.2 - 22.8)}{1.18} = 2.03$$

Reject H_0 if $z > 1.645$

Reject H_0 ; conclude H_a is true and $\sim_1 - \sim_2 > 0$.

b. $p\text{-value} = .5000 - .4788 = .0212$

12. a. $s_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = \sqrt{\frac{(8.4)^2}{80} + \frac{(7.6)^2}{70}} = 1.31$

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\sim_1 - \sim_2)}{s_{\bar{x}_1 - \bar{x}_2}} = \frac{(104 - 106) - 0}{1.31} = -1.53$$

Reject H_0 if $z < -1.96$ or $z > 1.96$

Do not reject H_0

b. $p\text{-value} = 2(.5000 - .4370) = .1260$

Computation of slope:

$$b_1 = \frac{\sum tY_t - (\sum t \sum Y_t) / n}{\sum t^2 - (\sum t)^2 / n} = \frac{6491 - (28)(1575) / 7}{140 - (28)^2 / 7} = 6.8214$$

Computation of intercept:

$$b_0 = \bar{Y} - b_1 \bar{t} = 225 - 6.8214(4) = 197.714$$

$$\text{Equation for linear trend: } T_t = 197.714 + 6.821t$$

$$\text{Forecast: } T_8 = 197.714 + 6.821(8) = 252.28$$

$$T_9 = 65.025 + 4.735(9) = 259.10$$

35. The following values are needed to compute the slope and intercept:

$$\sum t = 120 \quad \sum t^2 = 1240 \quad \sum Y_t = 578,400 \quad \sum tY_t = 5,495,900$$

Computation of slope:

$$b_1 = \frac{\sum tY_t - (\sum t \sum Y_t) / n}{\sum t^2 - (\sum t)^2 / n} = \frac{5,495,900 - (120)(578,400) / 15}{1240 - (120)^2 / 15} = 3102.5$$

Computation of intercept:

$$b_0 = \bar{Y} - b_1 \bar{t} = (578,400 / 15) - 3102.5(120 / 15) = 13,740$$

$$\text{Equation for linear trend: } T_t = 13,740 + 3102.5 t$$

b. 1995 forecast: $T_t = 13,740 + 3102.5 (16) = 60,277.5$

$$1996 \text{ forecast: } T_t = 13,740 + 3102.5 (17) = 66,482.5$$

36. a. A graph of these data shows a linear trend.

- b. The following values are needed to compute the slope and intercept:

$$\sum t = 15 \quad \sum t^2 = 55 \quad \sum Y_t = 200 \quad \sum tY_t = 750$$

Computation of slope:

$$b_1 = \frac{\sum tY_t - (\sum t \sum Y_t) / n}{\sum t^2 - (\sum t)^2 / n} = \frac{750 - (15)(200) / 5}{55 - (15)^2 / 5} = 15$$

Computation of intercept:

$$b_0 = \bar{Y} - b_1 \bar{t} = 40 - 15(3) = -5$$

$$\text{Equation for linear trend: } T_t = -5 + 15t$$