

## 10-5 Basic Skills and Concepts

### Statistical Literacy and Critical Thinking

**1. Terminology** Using the lengths (in.), chest sizes (in.), and weights (lb) of bears from Data Set 7 in Appendix B, we get this regression equation:  $\text{Weight} = -274 + 0.426 \text{ Length} + 12.1 \text{ Chest Size}$ . Identify the response and predictor variables.

**2. Best Multiple Regression Equation** For the regression equation given in Exercise 1, the  $P$ -value is 0.000 and the adjusted  $R^2$  value is 0.925. If we were to include an additional predictor variable of neck size (in.), the  $P$ -value becomes 0.000 and the adjusted  $R^2$  becomes 0.933. Given that the adjusted  $R^2$  value of 0.933 is larger, is it better to use the regression equation with the three predictor variables of length, chest size, and neck size? Explain.

**3. Adjusted Coefficient of Determination** For Exercise 2, why is it better to use values of adjusted  $R^2$  instead of simply using values of  $R^2$ ?

**4. Interpreting  $R^2$**  For the multiple regression equation given in Exercise 1, we get  $R^2 = 0.928$ . What does that value tell us?

**Interpreting a Computer Display.** In Exercises 5–8, we want to consider the correlation between LDL cholesterol and the predictor variables of weight (WT) and systolic blood pressure (SYS) for females. Refer to the Minitab display and answer the given questions or identify the indicated items. The Minitab display is based on the measured values of LDL cholesterol, weight, and systolic blood pressure amounts for the sample of 40 females from Data Set 1 of Appendix B.

#### MINITAB

The regression equation is  
 $\text{LDL} = 47.4 + 0.085 \text{ WT} + 0.497 \text{ SYS}$

Predictor	Coef	SE Coef	T	P
Constant	47.44	34.85	1.36	0.182
WT	0.0853	0.2833	0.30	0.765
SYS	0.4968	0.2618	1.90	0.066

$S = 36.0398$      $R\text{-Sq} = 9.8\%$      $R\text{-Sq}(\text{adj}) = 4.9\%$

#### Analysis of Variance

Source	DF	SS	MS	F	P
Regression	2	5214	2607	2.01	0.149
Residual Error	37	48058	1299		
Total	39	53272			

**5. LDL Cholesterol** Identify the multiple regression equation that expresses the level of LDL cholesterol in terms of weight (WT) and diastolic blood pressure (SYS).

**6. LDL Cholesterol** Identify the following:

- The  $P$ -value corresponding to the overall significance of the multiple regression equation
- The value of the multiple coefficient of determination  $R^2$
- The adjusted value of  $R^2$

**7. LDL Cholesterol** Should the multiple regression equation be used for predicting the LDL cholesterol level based on weight and diastolic blood pressure? Why or why not?