1.19. **Grade point average.** The director of admissions of a small college selected 120 students at random from the new freshman class in a study to determine whether a student's grade point average (GPA) at the end of the freshman year (Y) can be predicted from the ACT test score (X). The results of the study follow. Assume that first-order regression model (1.1) is appropriate.

i:	1	2	3	...	118	119	120
X_i:	21	14	28	...	28	16	28
Y_i:	3.897	3.885	3.778	...	3.914	1.860	2.948

a. Obtain the least squares estimates of β_0 and β_1, and state the estimated regression function.

b. Plot the estimated regression function and the data. Does the estimated regression function appear to fit the data well?

c. Obtain a point estimate of the mean freshman GPA for students with ACT test score $X = 30$.

d. What is the point estimate of the change in the mean response when the entrance test score increases by one point?

Least Squares Estimators. The estimators b_0 and b_1 that satisfy the least squares criterion can be found in two basic ways:

1. Numerical search procedures can be used that evaluate in a systematic fashion the least squares criterion Q for different estimates b_0 and b_1 until the ones that minimize Q are found. This approach was illustrated in Figure 1.9 for the persistence study example.
2. Analytical procedures can often be used to find the values of b_0 and b_1 that minimize Q. The analytical approach is feasible when the regression model is not mathematically complex.

Using the analytical approach, it can be shown for regression model (1.1) that the values b_0 and b_1 that minimize Q for any particular set of sample data are given by the following simultaneous equations:

$$\sum Y_i = nb_0 + b_1 \sum X_i \tag{1.9a}$$

$$\sum X_i Y_i = b_0 \sum X_i + b_1 \sum X_i^2 \tag{1.9b}$$

Equations (1.9a) and (1.9b) are called *normal equations;* b_0 and b_1 are called *point estimators* of β_0 and β_1, respectively.

The normal equations (1.9) can be solved simultaneously for b_0 and b_1:

$$b_1 = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sum(X_i - \bar{X})^2} \tag{1.10a}$$

$$b_0 = \frac{1}{n}\left(\sum Y_i - b_1 \sum X_i\right) = \bar{Y} - b_1\bar{X} \tag{1.10b}$$

where $\bar{X}$ and $\bar{Y}$ are the means of the X_i and the Y_i observations, respectively. Computer calculations generally are based on many digits to obtain accurate values for b_0 and b_1.

1.32. Derive the expression for b_1 in (1.10a) from the normal equations in (1.9).