Section 1.3. Simple Linear Regression Model with Distribution of Error Terms Unspecified

Note. In this section we define the simple linear regression model, explain it, give graphical representations of examples of it, and give an alternative form of it.

Note/Definition. The model considered in Part I, "Simple Linear Regression," is the model with one predictor variable and a linear regression function. The model is then of the form

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \tag{1.1}$$

where:

Y_i is the value of the response variable in the ith trial,

β_0 and β_1 are parameters,

X_i is a known constant, namely, the value of the predictor variable in the ith trial, and

ε_i is a random error term with mean $E\{\varepsilon_i\} = 0$ and variance $\sigma^2\{\varepsilon_i\} = \sigma^2$; ε_i and ε_j are uncorrelated so that their covariance is zero (i.e., $\sigma\{\varepsilon_i, \varepsilon_j\} = 0$ for all $i \neq j$) for $i = 1, 2, \ldots, n$.

This model is said to be *simple* (since it involves one predictor variable), *linear in the parameters*, and *linear in the predictor variable.* Since it only involves the first power of the parameters β_0 and β_1 and in the predictor variable X then it is a *first-order model.*

Note 1.3.A. The response variable Y_i in the ith trial is the sum of a constant term $\beta_0 + \beta_1 X_i$ and the random term ε_i So Y_i is a random variable. Since $\beta_0 + \beta_1 X_i$ is constant, $E\{\varepsilon_i\} = 0$, and expectation $E\{\cdot\}$ is linear (see my online notes for Mathematical Statistics 1 [STAT 4047/5047] on Section 1.8. Expectation of a Random Variables; notice Theorem 1.8.2) then

$$E\{Y_i\} = E\{\beta_0 + \beta_1 X_i + \varepsilon_i\} = E\{\beta_0 + \beta_1 X_i\} + E\{\varepsilon_i\} = \beta_0 + \beta_1 X_i.$$

So when the "level" (i.e., value) of X in the ith trial is X_i, the response (random) variable Y_i has has mean $E\{Y_i\} = \beta_0 + \beta_1 X_i$. So in terms of the predictor variable X we have from the regression model that $E\{Y\} = \beta_0 + \beta_1 X$. So the regression function relates the means of the probability distributions of Y for a given level of X.

Note 1.3.B. For random variable V with finite mean and finite variance, we have $\mathrm{Var}(aV + b) = a^2\mathrm{Var}(V)$ (or $\sigma^2\{aV + b) = a^2\sigma^2\{V\}$ in the notation of Kunter et al.) for constants a and b (see Theorem 1.9.1 in my online notes for Mathematical Statistics 1 on Section 1.9. Some Special Expectations). So for random variable ε_i with variance $\sigma^2\{\varepsilon_i\} = \sigma^2$ (a constant for all ε_i), we have

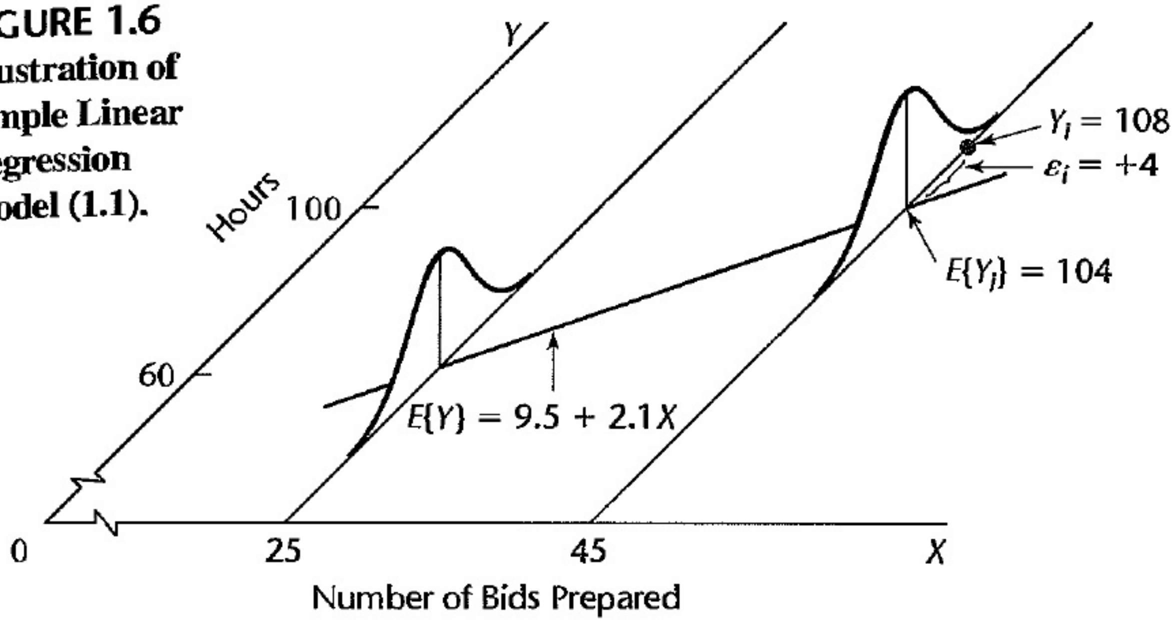
$$\sigma^2\{Y_i\} = \sigma^2\{\beta_0 + \beta_1 X_i + \varepsilon_1\} = (1)^2\sigma^2\{\varepsilon\} = \sigma^2.$$

So we see that the assumption that $\sigma^2\{\varepsilon_i\} = \sigma^2$ for all ε_i implies that the distributions of Y have variance σ^2 for all levels of predictor variable X. Since the error terms ε_i and ε_j are uncorrelated, the the responses Y_i and Y_j are uncorrelated. In summary, the simple linear regression model (1.1) gives the response variable Y_i

from a probability distribution with mean $E\{T_i\} = \beta_0 + \beta_1 X_i$ and variance the constant σ^2 (so the variance is independent of X_i).

Example. The seller of an item takes a number of bids for the item during a week and then measure the time required for the bids to be prepared by the bidders. Suppose the regression model is: $Y_i = 9.5 + 2.1X_i + \varepsilon_i$ where X is the number of bids prepared in a week and Y is the number of hours required to prepare the bids. Figure 1.6 shows the regression function $E\{Y\} = 9.5 + 2.1X$, along with the distributions of random variable Y for the predictor value at levels 25 and 45 (notice that both distributions have the same shapes since the variance is the same for each). Suppose in the ith week that $X_i = 45$ bids are prepared and the actual number of hours required is $Y_i = 108$. Since $E\{Y_i\} = 9.5 + 2.1(45) = 104$ then the error term is $\varepsilon_i = 108 - 104 = 4$. This is shown in detail in Figure 1.6.

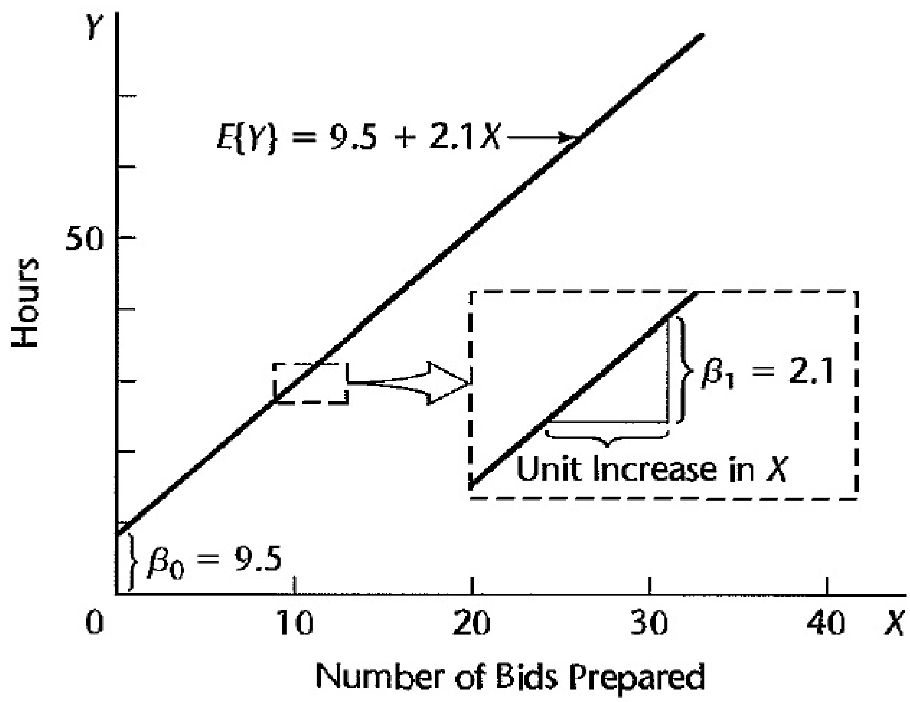
FIGURE 1.6
Illustration of Simple Linear Regression Model (1.1).



Definition. The parameters β_0 and β_1 in regression model (1.1) are *regression coefficients.* Parameter β_1 is the slope of the regression line. The parameter β_0 is the Y intercept of the regression line.

Note. Figure 1.7 shows how to geometrically interpret β_0 and β_1 for the regression function $E\{Y\} = 9.5 + 2.1X$ considered above. Regression coefficient β_1 gives the change in the mean of the probability distribution of Y per unit increase in X. Regression coefficient β_0 mean of the probability distribution of Y when $X = 0$ (though this may not have a meaning if predictor variable relates to a quantity that must be positive).

FIGURE 1.7 Meaning of Parameters of Simple Linear Regression Model (1.1).



Note 1.3.C. We can write simple linear regression model (1.1) in an alternative form. First, we can introduce $X_0 = 1$ so that we have $Y_i = \beta_0 X_0 + \beta_1 X_1 + \varepsilon_i$ (we will use this form in Chapter 5 when we introduce matrices). Let $\overline{X}$ represent the average of observed values of X (a constant). Then we can write the model by treating $X_i - \overline{X}$ as the predictor variable instead of X_i itself. We then have

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i = \beta_0 + \beta_1 X_i - \beta_1 \overline{X} + \beta_1 \overline{X} + \varepsilon_i$$

$$= (\beta_0 + \beta_1 \overline{X}) + \beta_1 (X_i - \overline{X}) + \varepsilon_i = \beta_0^* + \beta_1 (X_i - \overline{X}) + \varepsilon_i$$

where $\beta_0 = \beta_0 + \beta_1 \overline{X}$.

Revised: 9/17/2022