
8.5.1 Simple Linear Regression Model

Here, we provide a model that is called the **simple linear regression** model. Our model is

$$Y_i = \beta_0 + \beta_1 x_i + \epsilon_i,$$

where we model ϵ_i 's as independent and zero-mean normal random variables,

$$\epsilon_i \sim N(0, \sigma^2).$$

The parameters β_0 , β_1 , and σ^2 are considered fixed but unknown. The assumption is that we have data points (x_1, y_1) , (x_2, y_2) , \dots , (x_n, y_n) and our goal is to find the "best" values for β_0 and β_1 resulting in the line that provides the "best" fit for the data points. Here, y_i 's are the observed values of the random variables Y_i 's. To have a well-defined problem we add the following assumptions. We assume $n \geq 3$. We also assume that not all x_i 's are identical.

There are several common methods for finding good values for β_0 and β_1 . These methods will result in the same answers; however, they are philosophically based on different ideas. Here, we will provide two methods for estimating β_0 and β_1 . A third method will be discussed in the Solved Problems section.