

Fitting a Line via Least Squares

Math 430, Winter 2017

SLR Model Overview

The simple linear regression (SLR) model consists for two pieces:

1. The mean function: $E(Y|X = x) = \beta_0 + \beta_1 x$
2. The variance function: $Var(Y|X = x) = \sigma^2$

An alternative (and popular) way to express this model is

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$$

where $\varepsilon_i \stackrel{iid}{\sim} \mathcal{N}(0, \sigma^2)$.

While this theoretical form of the SLR model is interesting and leads to inferential results, for now we wish to be able to find the so-called “line of best fit” for our data. Mathematically, we write the line of best fit as

$$\hat{y}_i = b_0 + b_1 x_i$$

Notice that we use b_0 and b_1 to denote actual values of the y -intercept and slope.

With the line of best fit in hand, we can

- describe the linear relationship between two variables and
- predict the value of our response variable given the value of our predictor.

Least Squares

In order to **fit** the SLR model, we must choose values for b_0 and b_1 , which we will do using the **method of least squares**. This method sets the values of b_0 and b_1 so that the predicted y_i values, which we denote by \hat{y}_i , are “as close as possible” to the actual y_i value. Complete the below steps to find b_0 and b_1 via least squares.

Step 1: Write down the objective function

The **residual sums of squares** are defined as

$$RSS = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n (y_i - b_0 - b_1 x_i)^2$$

Step 2: Remember your calculus

We wish to minimize the RSS with respect to both b_0 and b_1 . Find the necessary partial derivatives and set them equal to 0.

$$\frac{\partial}{\partial b_0} RSS =$$

$$\frac{\partial}{\partial b_1} RSS =$$

Step 3: Remember your (high school) algebra

We now have a two equation, two unknown problem. Rearrange the above equations, simplifying sums when possible:

$$\sum_{i=1}^n y_i =$$

$$\sum_{i=1}^n x_i y_i =$$

The above equations are called the **normal equations**.

Step 4: Solve for b_0 and b_1

Solving the normal equations for b_0 and b_1 yield the **least squares (LS) estimates**. The LS estimates are so commonly used that we require special notation to separate them from other possible estimates. Statisticians, being the clever folks they are, place a $\hat{}$ (a “hat”) over the β_i ’s to denote they are LS estimates.

$$\hat{\beta}_0 =$$

$$\hat{\beta}_1 =$$

Luckily, R makes fitting an SLR model via least squares very simply, so we won’t need to manually crunch any numbers to find our LS estimates, but it is important to understand the model fitting process so that we don’t treat anything as a “black box.”