

## Problem Set 4

Due by 3/10

Consider a simple regression model

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

for  $i = 1, 2, \dots, n$ . Here  $\{y_i, x_i\}$ s are random observations (i.i.d.), and  $\beta_0$  and  $\beta_1$  are unknown coefficients. The least-squares methodology tries to infer about the unknown coefficients by finding a best fit to sample data from the population. The best fit refers to values of  $\beta_0$  and  $\beta_1$  so that the sum of the squared residuals (or prediction errors) is the least possible, i.e.

$$(\hat{\beta}_0, \hat{\beta}_1)' = \arg \min_{\beta_0, \beta_1} \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2.$$

It can be easily shown that the least squares estimator for the slope coefficient  $\beta_1$  is given by

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

and the least-squares estimator for the intercept is given by

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x},$$

where  $\bar{y}$  and  $\bar{x}$  denote the sample means of  $y$  and  $x$ , respectively.

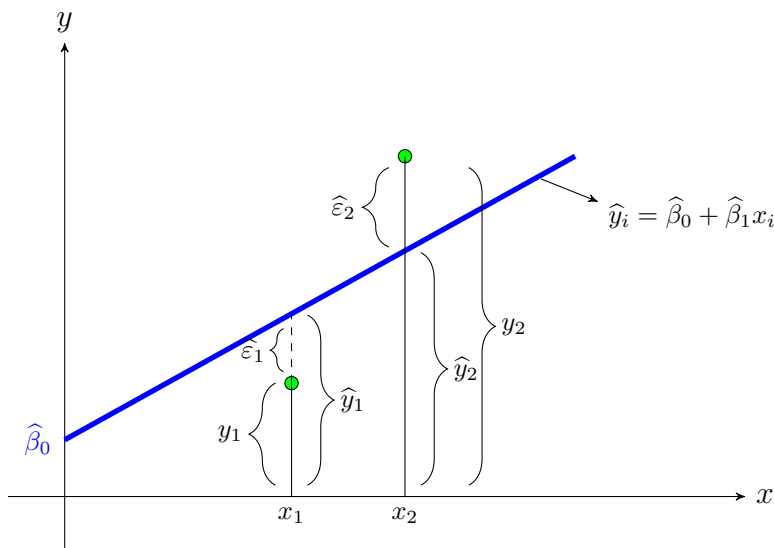


FIGURE 1. The least-squares regression line.

- (1) Write an R function which will take  $y$  and  $x$  as inputs and return estimates of  $\beta_0$  and  $\beta_1$  using the least-squares methodology.
- (2) Set the seed to 37 for the random number generator, i.e., `set.seed(37)`.
- (3) Generate 1000 observations on  $x$  by drawing randomly from the standard normal distribution.
- (4) Generate  $y$  as

$$y_i = 0.5 + 1.8x_i + \varepsilon_i$$

for  $i = 1, 2, \dots, 1000$ , where  $\varepsilon_i$  is a random draw from the standard normal distribution.

- (5) Using your function estimate the simple linear regression model.
- (6) Are the estimates you obtained for  $\beta_0$  and  $\beta_1$  the same as the true values you set, i.e. 0.5 and 1.8? Why?
- (7) Regenerate  $y$  by regenerating  $\varepsilon$ . Fit the model again using your function. Are the estimates the same the estimates you obtained from the initial estimation? Why?