STAT 757 HOMEWORK 1 SOLUTION

1. 1.5. "EY_i = \beta_0 + \beta_1 X_i + \epsilon_i."
LHS = EY_i = expected value of Y_i, which is a number.
RHS = \beta_0 + \beta_1 X_i + \epsilon_i which is a random variable, because \epsilon_i is a random variable.
So LHS \neq RHS, thus the statement the student wrote is false.

The simple linear regression model is Y_i = \beta_0 + \beta_1 X_i + \epsilon_i. The regression function is EY_i = \beta_0 + \beta_1 X_i, where the average of the response depends linearly on the value of the predictor.

2. 1.7. Regression model 1.1: Y_i = \beta_0 + \beta_1 X_i + \epsilon_i, errors are uncorrelated, mean 0 and variance \sigma^2. With given values for the parameters, we have
Y_i = 100 + 20X_i + \epsilon_i, \sigma^2 = 25, X = 5.

a. I cannot state any probabilities for Y, because I do not have the distribution for Y. I do not have the distribution for Y because I do not have the distribution for \epsilon_i.

b. We assume normal error model, that is \epsilon_i \sim N(0, \sigma^2), in this case \epsilon_i \sim N(0, 25). Then

\[ P(195 < Y < 205) = P\left(\frac{195 - 200}{5} < Z < \frac{205 - 200}{5}\right) = P(-1 < Z < 1) = 0.8413 - \]

3. 1.8 Note the equation for the average/expected value of the response for a given value of the predictor: EY = \beta_0 + \beta_1 X. The average response depends only on the predictor. Thus, if the predictor does not change, EY does not change. So, given X = 45, EY is always 104.
However, a new observation Y has an equation Y = \beta_0 + \beta_1 X + \epsilon = \beta_0 + \beta_1 45 + \epsilon, which is a random variable, because of the presence of \epsilon. Thus, the new value of Y will likely be different from 108. We simply do not know what the new observation of Y will be.

4. 1.16. The least squares method does not require any distributional assumptions to