Worksheet 15

1. Consider a simple linear regression where the first $n/2$ values of x_i are zero and the second $n/2$ values of x_i are 1. We can use this to model the mean of a variable that depends on whether the corresponding x_i is in group 0 or group 1. Specifically, how do the means of Y_i in these two groups correspond to the parameters b_0 and b_1?

2. Let $\bar{y}_A$ be the mean of the first $n/2$ values of Y_i and let $\bar{y}_B$ be the mean of the second $n/2$ values of Y_i. Consider the following form of the MLE for $\widehat{b}_1$:[1]

$$\widehat{b}_1 = \frac{\sum_i (y_i - \bar{y})(x_i)}{\sum_i (x_i - \bar{x})^2}.$$

[1] It can be shown that it is equivalent to the form on Worksheet 14.

Find a simple formula for $\widehat{b}_1$ in terms of $\bar{y}_A$ and $\bar{y}_B$.

3. Continuing from the previous question, find a simple formula for $\widehat{b}_0$ in terms of $\bar{y}_A$ and $\bar{y}_B$.

4. What is the connection between the linear regression here and a two-sample T-test for the means across the two groups?