Assignment 1 Report

LIANG Jing   IEDA

September 17, 2025

1   Idea Introduction

In the basic linear regression model, gender is assumed to affect height linearly. However, according to research by Cole or Tanner, the influence of gender on height is not linear but rather exhibits a more complex relationship with parental heights.

To avoide using a much more complex model, I attempt to split the data by gender and construct separate linear regression models to predict the heights of boys and girls, respectively.

2   Results

After conducting 100 repeated experiments, the mean and standard deviation of the Mean Squared Error (MSE) were computed, yielding the following results:

- **Single Linear Regression Model:** Mean MSE = 4.8273 ± 0.5152

- **Gender-Separated Linear Regression Model:** Mean MSE = 4.6750 ± 0.4175

Both the mean MSE and its standard deviation decreased, indicating model has slightly improved.

1the regression results of one of the experiments, and the boxplot of MSE repetitions is shown in Figure 2.
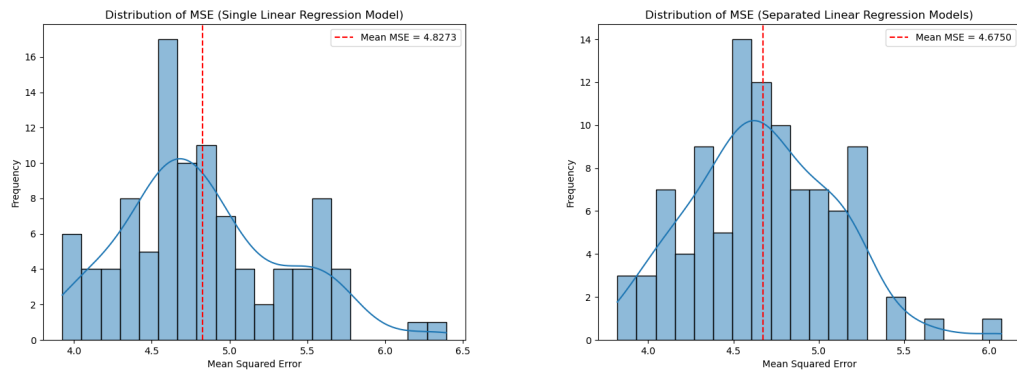
Figure 1: Left: Single Linear Regression Model; Right: Gender-Separated Linear Regression Model
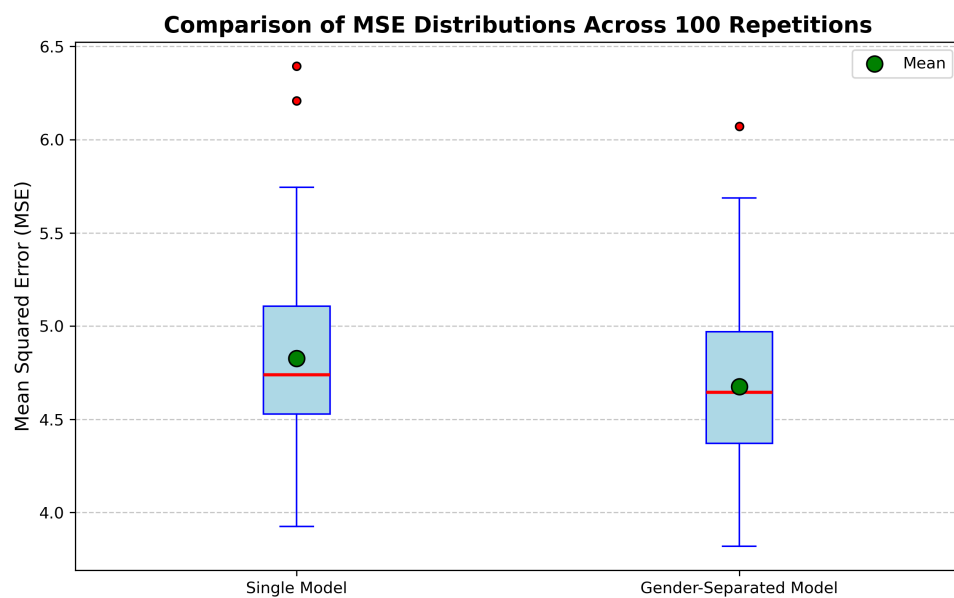

Figure 2: Boxplot of MSE