Lab 3 Linear Regression

Problem Statement

Suppose a fire insurance company wants to relate the amount of fire damage in major residential fires to the distance between the burning house and the nearest fire station.

The study is to be conducted in a large suburb of a major city; a sample of 15 recent fires in this suburb is selected.

The amount of damage, y, and the distance between the fire and the nearest fire station, x, are recorded for each fire.

Figure 1: Figure 1 Amount of Damage vs. Distance

Dataset

Table 10.5 Fire Damage Data	
Distance from Fire Station, x (miles)	Fire Damage, y (thousands of dollars)
3.4	26.2
1.8	17.8
4.6	31.3
2.3	23.1
3.1	27.5
5.5	36.0
.7	14.1
3.0	22.3
2.6	19.6
4.3	31.3
2.1	24.0
1.1	17.3
6.1	43.2
4.8	36.4
3.8	26.1

Data Set: FIREDAM

Figure 2: Fire Damage Data

Questions

We assume a simple linear relationship between `y` and `x`, and hypothesise that the model is of the following form:

$$y = \beta_0 + \beta_1 x + \varepsilon$$

1) Use R to estimate the unknown parameters of the hypothesised model. Find the least squares estimates of the slope $\hat{\beta}_1$ and intercept $\hat{\beta}_0$ on the printout. Write down the least squares equation for this model.

2) Plot the dataset and the least squares line on the same figure.

3) How to interpret the slope $\hat{\beta}_1$ and y-intercept $\hat{\beta}_0$ of the least squares line?

4) Measuring the extent to which the model fits the data.

a. What is the RSE of this model? How to interpret it?
b. What is the R-squared value? How to interpret it?

5) Test the null hypothesis that $\hat{\beta}_1$ is 0. Find the t-value and p-value. What conclusions can you draw from the two values?

6) Predict the value of damage (`y`) for a new set of distances (`x = 0.5, 1.5, 2.5, 3.5`). Can you obtain the above new `y`'s simultaneously?

7) Find the confidence interval for the prediction of damage (`y`) for a new set of distance (`x = 0.5, 3, 5.5`). Can you obtain the above confidence intervals simultaneously?

8) Find the prediction interval for the prediction of damage (`y`) for a new set of distance (`x = 0.5, 3, 5.5`). Can you obtain the above confidence intervals simultaneously?

9) Compare your results in 7) and 8) and comment on the confidence intervals and prediction intervals.

10) Plot the confidence intervals and prediction intervals in the same figure.