MATH 4753 Laboratory 15: Simple Linear Regression

In this lab we will investigate more detail concerning the mathematical underpinning of ***Simple*** Linear Regression.

Sampling distribution of $\widehat{\boldsymbol{\beta}}_{\mathbf{1}} \sim \boldsymbol{N}\left(\boldsymbol{\beta}_{\mathbf{1}}, \frac{\boldsymbol{\sigma}}{\sqrt{\boldsymbol{SS}_{xx}}}\right) \approx \boldsymbol{N}(\boldsymbol{\beta}_{\mathbf{1}}, \frac{\boldsymbol{s}}{\sqrt{\boldsymbol{SS}_{xx}}})$

When performing hypothesis testing it pays to remember the model, namely

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i, \epsilon_i \sim N(0, \sigma^2)$$

The hypotheses are:

$$H_0: \beta_1 = 0$$
$$H_1: \beta_1 \neq 0$$

The standard error is defined as $\sigma_{\widehat{\beta}_1} = \frac{\sigma}{\sqrt{SS_{xx}}} \approx \frac{s}{\sqrt{SS_{xx}}} = s_{\widehat{\beta}_1}$

Since σ will usually be unknown, the test statistic will generally be

$$T = \frac{\hat{\beta}_1 - Hypothesised\ value}{s_{\widehat{\beta}_1}} = \frac{\hat{\beta}_1 - 0}{\frac{s}{\sqrt{SS_{xx}}}}$$

Tasks

All output made please copy and paste into **this word file**. Save and place in the dropbox when completed. Anything you are asked to make should be recorded under the question in this document. There will be two files you need to upload:

- a pdf of this document (pdf) or the word file (docx)
- a text file of all the code you used to create answers (txt)

Note: All output in RMD

You are expected to adjust the functions as needed to answer the questions within the tasks below.

- Task 1
  - Make a folder LAB15
  - Download the file "lab15.r"
  - Place this file with the others in LAB15.
  - Start Rstudio
  - Open "lab15.r" from within Rstudio.
  - Go to the "session" menu within Rstudio and "set working directory" to where the source files are located.
  - Issue the function `getwd()` and copy the output here.
  - Create your own R file and record the R code you used to complete the lab.

- Task 2
  - Please answer the question below. You can use any functions available within R

10.32 *Rate of seed germination.* Researchers at the University of North Carolina–Greensboro investigated a model for the rate of seed germination (*Journal of Experimental Botany*, Jan. 1993). In one experiment, alfalfa seeds were placed in a specially constructed germination chamber. Eleven hours later, the seeds were examined and the change in free energy (a measure of germination rate) recorded. The results for seeds germinated at seven different temperatures are given in the table. The data were used to fit a simple linear regression model, with y = change in free energy and x = temperature.

SEEDGERM

Change in Free Energy, kJ/mol	Temperature, K
7	295
6.2	297.5
9	291
9.5	289
8.5	301
7.8	293
11.2	286.5

a. Plot the points in a scattergram.
b. Find the least-squares prediction equation.
c. Plot the least-squares line, part **b**, on the scattergram of part **a.**
d. Conduct a test of model adequacy. Use $\alpha = .01$.
e. Use the plot, part **c**, to locate any unusual data points (outliers).
f. Eliminate the outlier, part **e**, from the data set, and repeat parts **a–d**.

- a)
- b)
- c)
- d) (Use a test for $\beta_1 = 0$)
- e)
- f)

################## LAB FINISHES HERE ##############################