MTHS3003 Portfolio 3

Use the **R** markdown template provided: MTHS3003_Assessment3_Template.Rmd to answer the questions below for your data set (based on your student number).

R code must be provided for all calculations. Calculations performed without **R** code will score 0 marks.

Deadline: 12 noon Friday 28 November.

Coursework

A medical doctor has asked you to perform some statistical analysis of two new drugs, A and B against the existing drug X. They have provided you with two data sets and a set of queries for each data set. In the data sets below **Z** should be replaced by the **fifth** number in your student number.

Data set 1: **Drug_New_Z.csv** - For each patient in a study, which drug they have received A, B or X and whether or not they have experienced (Yes/No) an improvement in their condition since starting to take the drug.
Data set 2: **Drug_A_Z.csv** - A study of Drug A where a well-being score was taken prior to starting treatment (before) and 6 months after starting to take Drug A (after).

Data set 1

a. Tabulate the data. [2 marks]

b. Calculate the probability that an individual whose response was Yes was given Drug X. [2 marks]

c. Assuming independence between drug and response, compute the expected number of No responses for Drug A. [3 marks]

d. Perform a chi-squared test to check the assumption of independence between drug and response. Interpret the conclusions. [3 marks]

Data set 2

a. Plot the well-being score after treatment against the well-being score before treatment. [1 mark]

b. Find the least squares estimates of the parameters α and β for the linear regression model:

$$y = \alpha + \beta x + \epsilon$$

where y denotes well-being score after treatment, x denotes well-being score before treatment and ϵ is a mean zero residual. [2 marks]

c. Plot the least squares regression line on the plot of the data. [1 mark]

d. Is the linear regression model a reasonable model for the data provided? [3 marks]

e. Construct a 95% confidence interval for the after treatment well-being score of a patient who has a prior to treatment well-being score of 50. [3 marks]