Project-CMTH380

Instructions

Data

Before you work on the project it might be useful to clear your environment. To do this you should run the command rm(list = ls()). The data for the project consists of three samples that can be found in the columns of the file data.txt . The samples are also given in an xlsx file. To analyse the files in R you could save the file in your working directory (the command getwd() will give you the working directory you have in your computer) and then run the following commands

```
read_data <- read.table("data.txt")
sample_1<-read_data[,1]
sample_2<-read_data[,2]
sample_3<-read_data[,3]
```

Submission File

To complete the project you may use any programming language. R is recommended but is not mandatory. Your submission file should consist of two sections. Section 1 will include your answers (numerical, plots) to each part of the project. Note that plots in R can be saved by using the export command in the plot window. Section 2 will include your code.

Problems

Part A

For each sample, plot the histogram and calculate the sample mean and sample variance. One of the samples is drawn from a normal distribution, determine which.

Part B

Sample 1 is drawn from a population with unknown mean and variance. Based on sample 1,

- Calculate the 99%, two sided confidence intervals on the mean of the population.
- Find the P-value of the test $H_0 : \sigma^2 = 0.5$ vs $H_1 : \sigma^2 \neq 0.5$.

Part C

Sample 3 is drawn from the production of a manufacture company. The value 1 in the data represents a non defective part and the value 0 represents a defective part. The company wants to demonstrate that the defective parts are less that 10% of their production. Perform a suitable hypothesis test to demonstrate this claim.

Usefull R commands: *hist, mean, length, qqnorm, var, sum, pnorm, qt, pchisq*