統計學與實習上

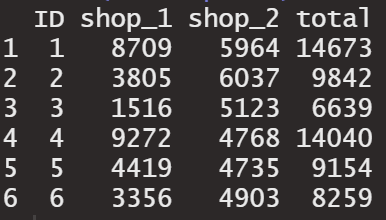
第五次作業

1.資料集「consumption.csv」為某百貨公司一部分客戶的消費紀錄，包含三個欄位：客戶編號(ID)、商店一消費金額(shop\_1、單位：元)、商店二消費金額(shop\_2、單位：元) (alpha = 0.05)。**(1.5 points)**

a. 將資料以指定路徑或預設工作路徑讀檔的方式讀入R中，命名為consumption。將資料新增一個欄位total，為兩商店消費金額的總和，即shop\_1 + shop\_2，並顯示前六筆顧客消費資料。**(0.3points)**

consumption = read.csv("C:/Users/stupi/Desktop/consumption.csv", header = T)

consumption$total = consumption$shop\_1 + consumption$shop\_2

head(consumption)

b. 商店二宣稱客戶平均來該店消費金額不等於4500元，已知該樣本的母體服從常態分佈，且母體標準差為1000元，利用單一樣本z檢定來檢驗是否符合商店二所述，請詳細列出檢定過程 (, ) 以及結果描述。 **(0.4 points)**

H0為客戶至商店二的平均消費金額 = 4500

H1為客戶至商店二的平均消費金額 ≠ 4500

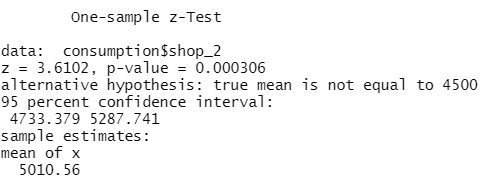
library(BSDA)

z.test(x = consumption$shop\_2,

mu = 4500,

sigma.x = 1000,

alternative = "two.sided")



95%區間估計並沒有涵蓋虛無假設H0 = 4500，故拒絕虛無假設。

c 假設消費金額總和(total)是來自𝛔=15𝟎𝟎的常態分配，請計算其95%信賴區間。 **(0.4 points)**

sigma = 1500

mean = mean(consumption$total)

alpha = 0.05

n = nrow(consumption)

upper = mean + qnorm(1 - alpha/2) \* sigma/sqrt(n)

lower = mean - qnorm(1 - alpha/2) \* sigma/sqrt(n)

c(lower,upper)

d. 延續c小題假設，若有一人宣稱到此百貨公司消費之全部客戶的消費金額總和為9000元(mu)，計算出Z統計量 (Zstat)。並計算在此統計量下，標準化後的樣本消費金額總和大於此統計量 (𝐙 > 𝐙𝐬𝐭𝐚𝐭) 的機率，並將機率結果呈現至四捨五入至小數點後第四位)。**(0.4 points)**

mu0 = 9000

sigma = 1500

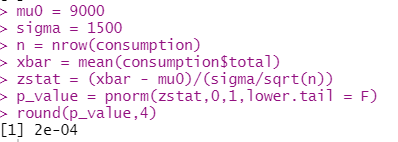
n = nrow(consumption)

xbar = mean(consumption$total)

zstat = (xbar - mu0)/(sigma/sqrt(n))

p\_value = pnorm(zstat,0,1,lower.tail = F)

round(p\_value,4)



2. (Text, p.308) In a poll to estimate presidential popularity, each person in a random  
sample of 1,000 voters was asked to agree with one of the following statements:  
1. The president is doing a good job.  
2. The president is doing a poor job.

3. I have no opinion

A total of 560 respondents selected the first statement, indicating they thought the president was doing a good job. **(by hand)**

a. Construct a 95% confidence interval for the proportion of respondents who feel the president is doing a good job. **(0.5 points)**

b. Based on your interval in part (a), is it reasonable to conclude that a majority of the population believes the president is doing a good job? **(0.5 points)**

*a.p*= 560/1000 = 0.560, from 0.529 up to 0.591, found by

b.The lower point of the interval is greater than 0.50. So we can conclude the majority feel the President is doing a good job

3.(Text, p.308) The First National Bank of Wilson has 650 checking account customers. A recent sample of 50 of these customers showed 26 have a Visa card with the bank. What is the margin of error for a 99% confidence interval for the population proportion? **(by hand) (0.5 points)**

*margin of error=*

4. 無疆社區有 20,000 個居民，其家庭每月收入分配近於常態分配，平均收入為75,000元，標準差為 15,000 元，今隨機抽取 100 戶為樣本，其平均收入為78,000 元，試以α＝0.05檢定此社區居民的家庭每月平均收入是否仍為 75,000。**(by hand) (1points)**

Step 1: State the Hypothesis

≠75000

Step 2: Select a Level of Significance

α = .05

Step3: Select the test statistic.

We’ll use z.

Step4: Formulate the decision rule

Reject Ho if z <-1.96 or z >1.96

Step5: Take sample, compute the test statistic, make decision.

the sample statistic : z=2, found by   
>>統計值落入拒絕域

C.I. for populations means:=(75060,80940)

>>此區間內為包含母體平均數 75,000

Step6: Interpret the result.

Reject , 表示樣本與母體的差異達統計顯著，無疆社區居民家庭每月平均收入可能已經不是 75,000 元

5. A recent national survey found that high school students watched an average (mean) of 6.8 movies per month with a population standard deviation of 1.8. The distribution of number of movies watched per month follows the normal distribution. A random sample of 36 college students revealed that the mean number of movies watched last month was 6.2. At the 0.05 significance level, can we conclude that college students watch fewer movies a month than high school students? **(by hand) (1points)**

Step 1: State the Hypothesis

6.8

Step 2: Select a Level of Significance

α = .05

Step3: Select the test statistic.

We’ll use z.

Step4: Formulate the decision rule

Reject if z <-1.65

Step5: Take sample, compute the test statistic, make decision.

the sample statistic : z=-2, found by   
>>統計值落入拒絕域

C.I. for populations means:=(-,6.69)

>>此區間內為包含母體平均數 6.8

Step6: Interpret the result.

Reject . The mean number of movie watching is less than 6.8 per month.