統計學與實習上
第四次作業

1.資料集 iris 為 R 語言內建之資料集，其中包含了 150 株鳶尾花的外表性狀調查資料，請使用此資料完成以下題目。**(by R) (1 points)**

a.請分別計算出花瓣 (petal)長度的樣本平均數與標準差。在假定花瓣長度為常態分佈時，以此樣本平均和標準差計算隨機採樣一株鳶尾花，其花瓣長度介於 2 至 5 公分的機率（機率值請四捨五入至小數點第二位）。**(0.2 points)**

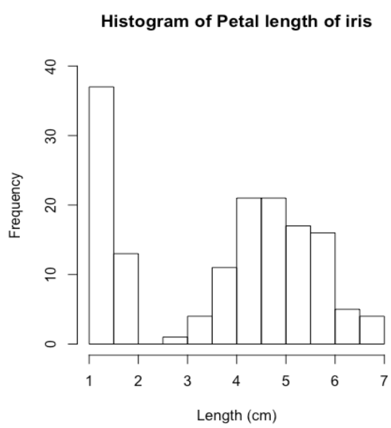
b.使用此組資料計算出實際長度介於 2 至 5 公分的樣本數佔總樣本數的比例（機率值請四捨五入至小數點第二位）。**(0.3 points)**

c.請比較 b, c 小題的結果，基於你的觀察說明兩者結果間之一致/不一致性來源為何，可適度使用統計圖表、統計值做輔助說明。**(0.5 points)**

```r
# a
data(iris)
ip_m <- mean(iris$Petal.Length); ip_s <- sd(iris$Petal.Length)
lower <- (2-ip_m)/ip_s; upper <- (5-ip_m)/ip_s
round(pnorm(upper) - pnorm(lower),2) # 0.6

# b.
n <- length(which(iris$Petal.Length >= 2 & iris$Petal.Length <= 5))
round(n/nrow(iris), 2) # 0.39

# c.
hist(iris$Petal.Length, ylim = c(0,40), main = 'Histogram of Petal length of iris', xlab = ‘Length (cm)’)
# As the histogram shows, the petal length in the iris dataset didn't follow the normal distribution.
# Therefore, using the probability of normal distribution to estimate may be incorrect.
```



2.(Text, p.264)A population consists of the following five values: 2, 2, 4, 4, and 8. **(by hand)**

a. List all samples of size 2, and compute the mean of each sample. **(0.2points)**

sample	values	sum	mean
1	2,2	4	2
2	2,4	6	3
3	2,4	6	3
4	2,8	10	5
5	2,4	6	3
6	2,4	6	3
7	2,8	10	5
8	4,4	8	4
9	4,8	12	6
10	4,8	12	6

b. Compute the mean of the distribution of sample means and the population mean. Compare the two values. **(0.2points)**

μ=(2 + 2 + 4 + 4 + 8)/5 = 4

$$\mu_{\bar{x}} = \frac{2+3+3+5+3+3+5+4+6+6}{10} = 10$$

They are equal.

c. Compare the dispersion in the population with that of the sample means. **(0.3points)**

The dispersion for the population is greater than that for the sample means. The population varies from 2 to 8, whereas the sample means only vary from 2 to 6.

3. Beer bottles are filled so that they contain an average of 330ml of beer in each bottle. Suppose that the amount of beer in a bottle is normally distributed with a standard deviation of 4ml. **(by hand)**

a. What is the probability that a randomly selected bottle will have less than 325ml of beer? **(0.4points)**

$$P(\mathrm{x} < 325) = P\left(\mathrm{x} < \frac{325-330}{4}\right) = P(z < -1.25) = 0.1056$$

b. What is the probability that a randomly selected 6-pack of bottle will have less than 325ml of beer? **(0.4points)**

$$P(\mathrm{x} < 325) = P\left(\mathrm{x} < \frac{325-330}{4/\sqrt{6}}\right) = P(z < -3.06) = 0.0011$$

c. What is the probability that a randomly selected 12-pack of bottle will have less than 325ml of beer? **(0.4points)**

$$P(\mathrm{x} < 325) = P\left(\mathrm{x} < \frac{325-330}{4/\sqrt{12}}\right) = P(z < -4.33) \approx 0$$

d. Comment on the sample size and the corresponding probabilities. **(0.5points)**

The probability that the mean weight of 12-pack of beer is less than 325 ml is much less than that of a single bottle because the variation in x is less when the sample size is bigger.

4. A small hair salon averages about 30 customers on weekdays with a standard deviation of 6. It is safe to assume that the underlying distribution if normal. In an attempt to increase the number of weekday customers, the manager offers a $2 discount on 5 consecutive weekdays. She reports that her strategy has worked since the sample mean of customers during the 5 weekday period jumps to 35. **(by hand)**

a. **<u>How</u>** unusual would it be to get a sample average of 35 or more customers if the manager has not offered the discount **(0.5points)**?

$P(x \geq 35) = P(z \geq \frac{35-30}{\frac{6}{\sqrt{5}}}) = P(z \geq 1.86) = 1 - 0.9686 = 0.0314$

Therefore, there is a 3.14% chance of getting a sample average of 35 or more without a discount.

b. Do you feel confident that the manager's discount strategy has worked? **<u>Explain</u>**. **(0.5points)**

We feel reasonably confident that the manager's discount strategy has worked since there is only a small chance of 3.14% of getting 35 or more customers if the manager had not offered the discount.

5.假設某次 TOEIC 考試台灣考生成績呈常態分配，平均μ =580，標準差σ =140。**(by hand)**

a.若某校研究所甄試入學將 TOEIC 成績高於 600 者列為審核加分的標準，假設有 49 位學生報名甄試該研究所，且所有參加甄試學生皆有 TOEIC 成績，且其成績也呈現常態分配，請問有多少甄試學生能夠獲得加分？**(0.3points)**

$\mathrm{P}\left(\mathrm{z} > \frac{600-580}{140/\sqrt{49}}\right)=$ P(z > 1)=0.5-0.3413=0.1587

49*0.1587=7.776，故取 7 位(或 8 位)

b.承上題，若假設要以 TOEIC 成績先節選前 33%的學生作為二階段口試的門檻，則應該如何指定該標準？**(0.3points)**

$P(\mathrm{x} \geq X) = 0.33$

$P\left(Z \geq \frac{x-580}{140/\sqrt{49}}\right) = 0.33$

$0.5 - P\left(Z \leq \frac{x-580}{140/\sqrt{49}}\right) = 0.33$

$P\left(Z \leq \frac{\mathrm{x}-580}{140/\sqrt{49}}\right) = 0.17 = P(Z \leq 0.44)$

$\frac{\mathrm{x}-580}{140/\sqrt{49}}$=0.44x=588.8，門檻最低分數為 589 分