國立政治大學 114 學年度第二學期
統計學(二) 期中 R 程式加分考-解答卷
(助教：江韋築)

系級:______________ 學號:______________ 姓名:_____________

考試科目: 統計學(二)
考試日期:2026/04/14
本試題共 6 大題 (共 120%)
考試時間:15:00~16:00 (共 60 分鐘)

注意事項:

1. 從教學網站下載電子考卷 (**`114-2-Stat-R-Midterm.zip`**)，並於自己的筆電解壓縮。壓縮檔包含題目卷、答案卷和資料集。
2. 本次考題以 R 程式(Rgui 或 RStudio)方式作答，其他程式不允許。
3. 考試過程中可查詢書本、教學講義或上網，禁止利用 messenger, IG, Line 等等通訊軟體。
4. 禁止使用 AI 模式搜尋。禁止使用 chatGPT 或類似的 AI 平台/工具。禁止疑似作弊行為。
1. 本答案卷上請務必於 R Console 內**複制「執行後的程式碼及結果(含圖形)」**，於本答案卷貼上(`Courier New, 10 點字，白底黑字`)，不是只有程式碼，不是只有報表。最後，將每小題之答案以黃色底高亮起來(不能只印出報表，要助教去找答案)。
5. 請依序註明題號: (1)a, (1)b, (2)a 等等。
6. 作答完請將此 word 檔存檔，檔名為「學號-姓名-Stat-R-Midterm.docx」(更改成自己「學號、姓名」)並上傳至教學網站【作業考試上傳區】或 http://hmwu.nccu.edu.tw/login.html
7. 帳號: stat114，密碼: 上課教室號碼，資料夾: 「20260414-Midterm」
8. 如果上傳網站出現「空白頁」，請將滑鼠移至「網址列」後，按「Enter」即可。若再不行，請換其它瀏覽器(IE/Edge/Firefox/Chrome)
9. 上傳檔案無法刪除，若要上傳更新檔，請於主檔名後加「-2」，例如:「學號-姓名-Stat-R-Midterm-2.docx」。

______________________祝考試順利______________________

(1) **(5 分)**	**用 R 印出下列字句(姓名改為自己的姓名):** "本人(學號)(姓名)恪遵各項考試規則，若如違反，願受校方最嚴厲處罰，謹誓。"

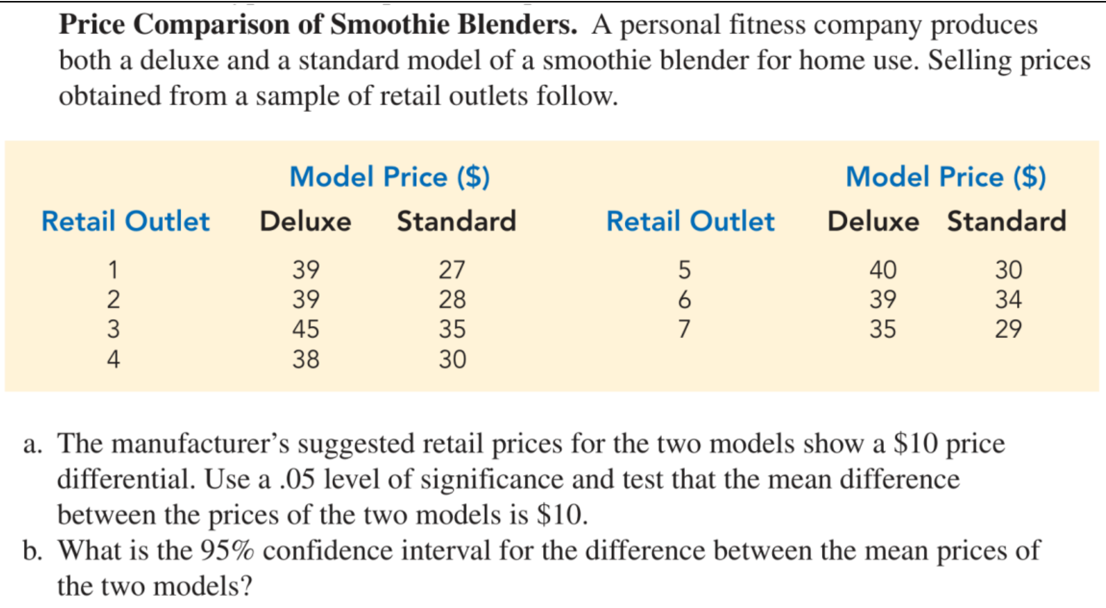
```
> cat("本人(115354501)(江韋築)恪遵各項考試規則，若如違反，願受校方最嚴厲處罰，謹誓。")
本人(115354501)(江韋築)恪遵各項考試規則，若如違反，願受校方最嚴厲處罰，謹誓。
```

(2) **(20 分)**	**Data file: Hotel.xlsx** **Hotel Price Comparison.** Suppose that you are responsible for making arrangements for a business convention and that you have been charged with choosing a city for the convention that has the least expensive hotel rooms. You have narrowed your choices to Atlanta and Houston. The file named *Hotel* contains samples of prices for rooms in Atlanta and Houston that are consistent with a *SmartMoney* survey conducted by Smith Travel Research. Because considerable historical data on the prices of rooms in both cities are available, the population standard deviations for the prices can be assumed to be $20 in Atlanta and $25 in Houston. Based on the sample data, can you conclude that the mean price of a hotel room in Atlanta is lower than one in Houston?

```
> library(readxl)
> data1 <- read_excel("Hotel.xlsx")
> sample1 <- na.omit(data1[[1]])#Atlanta
> sample2 <- na.omit(data1[[2]])#Huston
> sigma1 <- 20
> sigma2 <- 25
> alpha = 0.05
> result1 <- z.test(sample1, sample2, sigma.x = sigma1, sigma.y = sigma2
+ , conf.level = 1 - alpha
+ , alternative="less")# Left-tailed test
> result1$p.value
[1] 0.0352023

> # Conclusion: p-value < 0.05, Reject H0, conclude that Atlanta < Huston
```

(3)
(25 分)



Price Comparison of Smoothie Blenders. A personal fitness company produces both a deluxe and a standard model of a smoothie blender for home use. Selling prices obtained from a sample of retail outlets follow.

	Model Price ($)			Model Price ($)	
Retail Outlet	Deluxe	Standard	Retail Outlet	Deluxe	Standard
1	39	27	5	40	30
2	39	28	6	39	34
3	45	35	7	35	29
4	38	30			

a. The manufacturer's suggested retail prices for the two models show a $10 price differential. Use a .05 level of significance and test that the mean difference between the prices of the two models is $10.

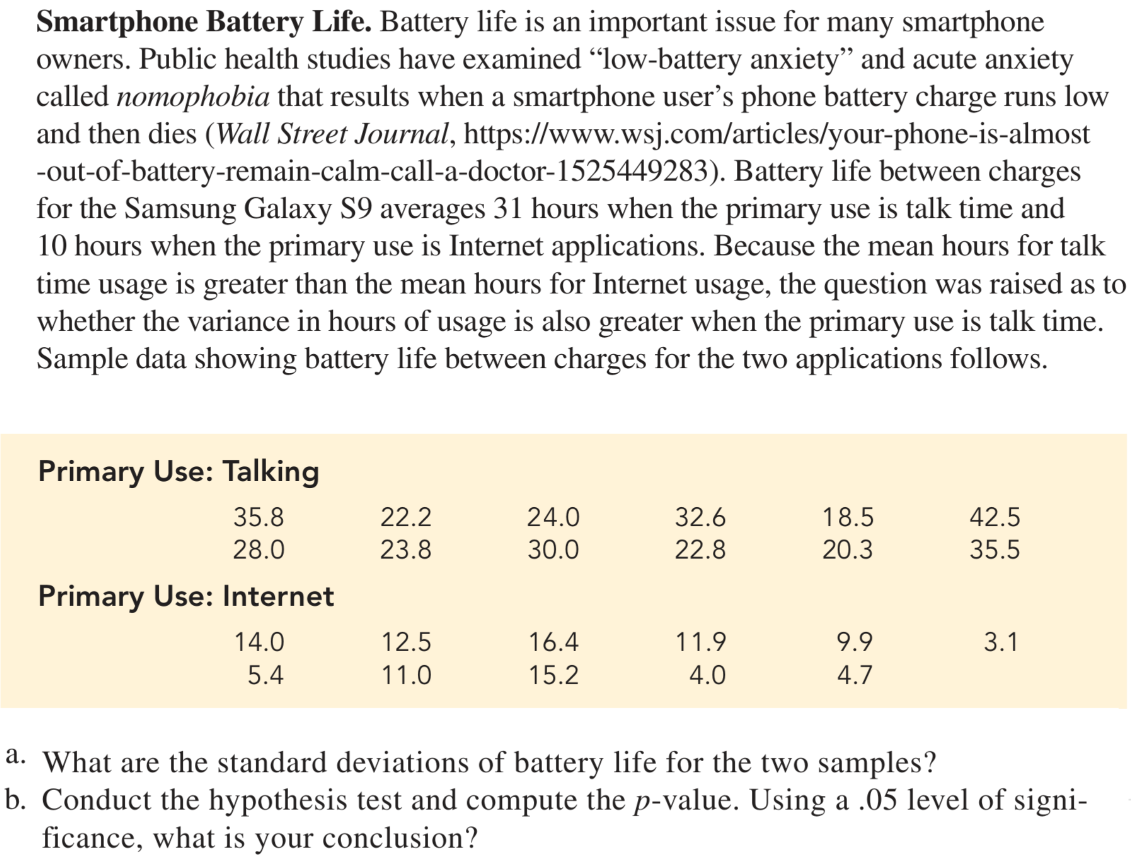
b. What is the 95% confidence interval for the difference between the mean prices of the two models?

((a)(b)禁止用 R 直接加減乘除做計算，請利用 R 套件或指令運算)

```
> sample1 <- c(39,39,45,38,40,39,35) # D
> sample2 <- c(27,28,35,30,30,34,29) # S
> alpha = 0.05
> result <- t.test(sample1, sample2, mu=10,
+ , conf.level = 1-alpha, alternative="two.sided"
+ , var.equal = FALSE)
>
> #(a)
> result$p.value
[1] 0.4879642
> # Conclusion: p-value> 0.05, do not reject H0,
> # mean difference between two models is not significantly different from 10
>
> #(b)
> result$conf.int
[1]  5.377159 12.337127
attr(,"conf.level")
[1] 0.95
```

(4)
(25 分)

Data file: BatteryLife.xlsx



Smartphone Battery Life. Battery life is an important issue for many smartphone owners. Public health studies have examined "low-battery anxiety" and acute anxiety called *nomophobia* that results when a smartphone user's phone battery charge runs low and then dies (*Wall Street Journal*, https://www.wsj.com/articles/your-phone-is-almost-out-of-battery-remain-calm-call-a-doctor-1525449283). Battery life between charges for the Samsung Galaxy S9 averages 31 hours when the primary use is talk time and 10 hours when the primary use is Internet applications. Because the mean hours for talk time usage is greater than the mean hours for Internet usage, the question was raised as to whether the variance in hours of usage is also greater when the primary use is talk time. Sample data showing battery life between charges for the two applications follows.

Primary Use: Talking

35.8	22.2	24.0	32.6	18.5	42.5
28.0	23.8	30.0	22.8	20.3	35.5

Primary Use: Internet

14.0	12.5	16.4	11.9	9.9	3.1
5.4	11.0	15.2	4.0	4.7	

a. What are the standard deviations of battery life for the two samples?

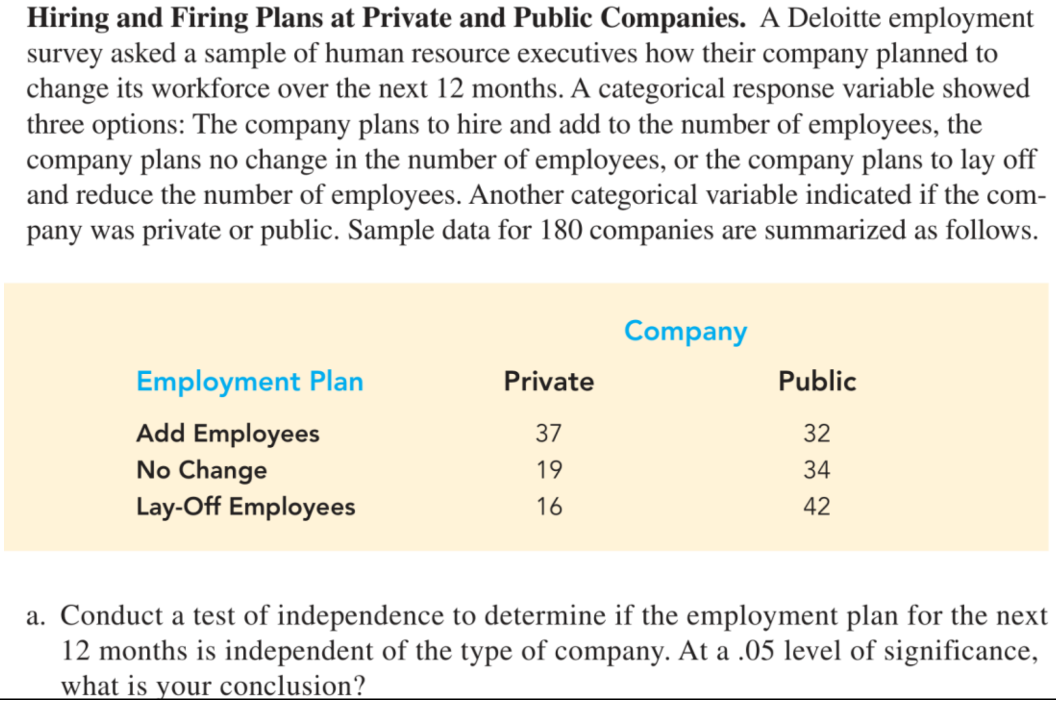
b. Conduct the hypothesis test and compute the *p*-value. Using a .05 level of significance, what is your conclusion?

((a)(b)禁止用 R 直接加減乘除做計算，請利用 R 套件或指令運算)

```
> data1 <- read_excel("BatteryLife.xlsx")
>
> sample1 <- na.omit(data1[[1]]) #Talking
> sample2 <- na.omit(data1[[2]]) #Internet
>
> #(a)
> sd(sample1)
[1] 7.363053
> sd(sample2)
[1] 4.768667
>
> #(b)
> alpha = 0.05
> result <- var.test(sample1, sample2, conf.level = 1-alpha,
+ alternative='greater')
> result$p.value
[1] 0.0910851
>
> # Conclusion: p-value > 0.05, do not reject H0,
> # the variance of talking is not significantly higher than variance of internet
```

(5)
(20 分)

Data file: WorkforcePlan.xlsx



Hiring and Firing Plans at Private and Public Companies. A Deloitte employment survey asked a sample of human resource executives how their company planned to change its workforce over the next 12 months. A categorical response variable showed three options: The company plans to hire and add to the number of employees, the company plans no change in the number of employees, or the company plans to lay off and reduce the number of employees. Another categorical variable indicated if the company was private or public. Sample data for 180 companies are summarized as follows.

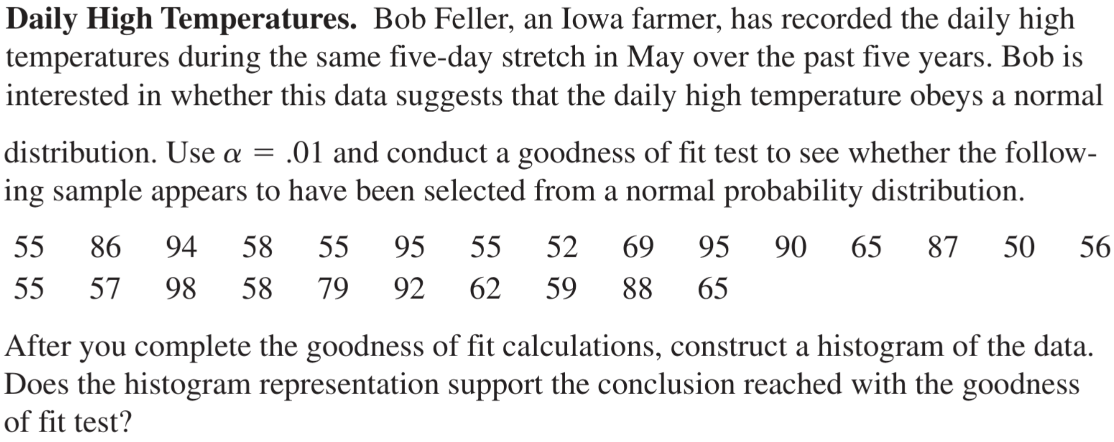
	Company	
Employment Plan	Private	Public
Add Employees	37	32
No Change	19	34
Lay-Off Employees	16	42

a. Conduct a test of independence to determine if the employment plan for the next 12 months is independent of the type of company. At a .05 level of significance, what is your conclusion?

```
> data3 <- read_excel("WorkForcePlan.xlsx")
> contingency_table <- table(data3[2:3])
> result2 <- chisq.test(contingency_table)
> result2$p.value
[1] 0.008913441
> # Conclusion: p-value < 0.05, reject H0,
> # we conclude that employment plan and type of company is not independent.
```

(6)
(20 分)

Temperatures.xlsx



Daily High Temperatures. Bob Feller, an Iowa farmer, has recorded the daily high temperatures during the same five-day stretch in May over the past five years. Bob is interested in whether this data suggests that the daily high temperature obeys a normal

distribution. Use $\alpha = .01$ and conduct a goodness of fit test to see whether the following sample appears to have been selected from a normal probability distribution.

55	86	94	58	55	95	55	52	69	95	90	65	87	50	56
55	57	98	58	79	92	62	59	88	65					

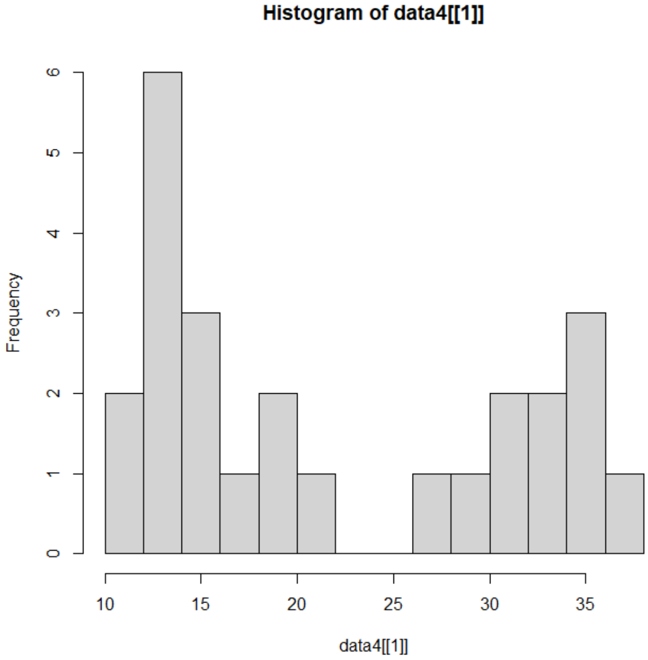
After you complete the goodness of fit calculations, construct a histogram of the data. Does the histogram representation support the conclusion reached with the goodness of fit test?

```
> data4 <- read_excel("Temperatures.xlsx")
>
> Percentage <- seq(0.1, 0.9, 0.1)
```

```
> z <- qnorm(Percentage)
> xbar <- mean(data4[[1]])
> s <- sd(data4[[1]])
> BinValues <- xbar + z * s
>
> BinInterval <- cut(data4[[1]],
+ breaks = c(min(data4[[1]]) - 1,
+                            BinValues ,
+    max(data4[[1]]) + 1))
> Observed_Frequency <- c(table(BinInterval))
> Expected_Frequency <- rep(nrow(data4)/10, 10)
>
> chi2 <- sum((Observed_Frequency - Expected_Frequency)^2 / Expected_Frequency)
> p.value <- pchisq(chi2, df = 7, lower.tail = F)
> p.value
[1] 0.001451553

> # Conclusion: p.value < 0.01, reject H0, we conclude that observed sample have
significant difference with normal distribution.

> hist(data4[[1]],breaks=10)
```



```
> # the histogram shows that the distribution of the observed sample is a bimodal
distribution, which apparently is different from normal. So, the histogram supports
the conclusion of goodness-of-fit test.
```