

R 語言與統計(三)

丘祐瑋
David Chiu

環境資訊頁面

- 所有課程補充資料、投影片皆位於
 - ▣ https://github.com/ywchiu/cdc_course

無母數檢定

無母數假設檢定

- 無母數檢定(Nonparametric Test) 又稱不需要分佈假設的統計方法 (distribution-free tests) 因為他們較少的假設
- Z檢定、t檢定只能應用在常態群體假設成，若群體分佈未知或不為常態分佈時，這些檢定方法均不適用

無母數方法的優點

■ 優點

- 可適用於檢定非常態分佈樣本之參數
- 適用於名目(Nominal)或次序(Ordinal)的資料型態
- 不需要母體參數(Population Parameter)即可檢定, 故稱無母數

■ 缺點

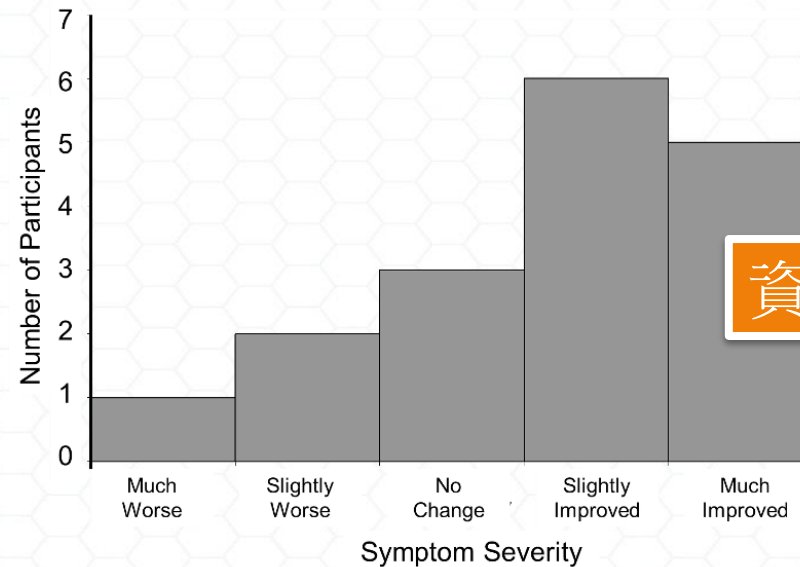
- 樣本不大時比較難否決虛無假設(null hypothesis)

適用無母數統計方法的案例

- 當資料呈現次序(Ordinal)或秩(Rank)
- 當資料有離群值時(Outliers)
- 當資料有明顯的限制時

當資料呈現次序(Ordinal)或秩(Rank)

- 一個臨床試驗，在6周的指定療程後，參與者被要求評估他們的症狀嚴重程度。症狀的嚴重程度可以用5點的順序量表來衡量，有反應選項:很糟，稍差，沒有變化，輕微改善，有很大改善
- 假設在試驗中總共有 $n=20$ 個參與者，隨機分為實驗治療或安慰劑，結果資料分佈如下圖所示。



資料明顯非常態分佈

當資料呈現次序(Ordinal)或秩(Rank)

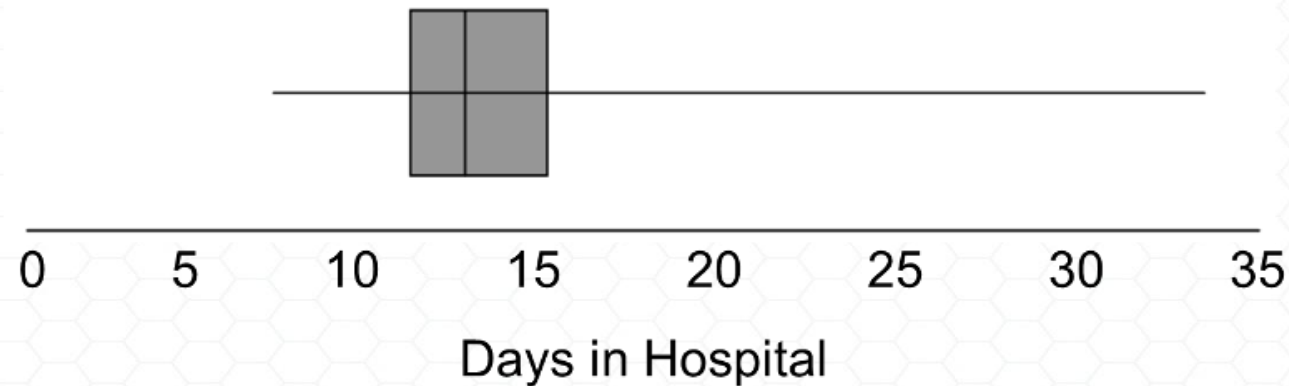
- 在一些研究中，結果以秩(Rank) 呈現
 - 例如，在婦產科研究中，經常使用APGAR 評估新生兒的健康狀況。分數範圍為1-10分，是根據嬰兒出生時的狀況而得出的五分量表的總和
 - APGAR評分一般不服從正態分佈，因為大多數新生兒的評分多在7分以上

當資料有離群值時(Outliers)

■ 結果是連續(Continuous)的卻有離群值(Outlier)

□ 例如，在醫院進行特定的外科手術後的幾天通常會出現異常值。假設在一項觀察性研究中，研究人員希望評估患者在民營醫院接受肝臟移植後的住院天數與公立醫院相比是否存在差異。假設我們在 $n=100$ 名參與者中進行移植手術，50名來民營醫院，50名來自公立醫院

□ 住院天數如箱型圖所示



當資料有明顯的限制時

■ 用不精確的方式測量連續數值

- 例如，一些儀器或化驗不能測量特定數量的存在或低於某些限度
- HIV病毒載量衡量每份血液中病毒的數量。它可以從“未檢測到”或“低於檢測限度”到數億數量。因此，在一些樣本中，參與者可能有1,254,000或874,050份病毒載量，而另一些則被測量為“未檢測到”
- 如果大量的受測者都被判定“未檢測到”，病毒載量的分佈就不會是正常分佈的

無母數統計檢定

- 檢定單一群體中心位置(中位數)或成對群體的分佈是否相同
- 進行無母數統計時，檢定的變數(Ordinal、Interval或Continuous)皆為從最低排到最高的順序(Rank)，而非測量值或原始值

檢定步驟

- 兩種假設：虛無假設 (null hypothesis) 與對立假設 (alternative hypothesis)
- 檢定步驟
 - 寫出所有假設
 - H_0 原始假設：觀測值是隨機結果
 - H_1 對立假設：存在一些因素影響結果
 - 檢查統計量
 - 檢驗數量與分佈
 - 檢驗P值
 - 觀察到極端機率有多少，P值越小越不利於原始假設
 - 比較P值與顯著性水準
 - e.g. <0.05 代表顯著

Mann Whitney U Test (Wilcoxon Rank Sum Test)

- 在一第二階段的臨床試驗，為了研究一種新藥降低兒童哮喘症狀的有效性。總共有 $n=10$ 人隨機接受新藥或安慰劑的測試。參與者被要求在接受指定治療後的1周內記錄呼吸急促的次數
- 資料如下：

Placebo	7	5	6	4	12
New Drug	3	6	4	2	1

- 與服用安慰劑的受試者相比，接受新藥物的受試者在1周內呼吸急促發作的次數是否有差異

排出資料順序

Placebo	7	5	6	4	12
New Drug	3	6	4	2	1



		Total Sample (Ordered Smallest to Largest)		Ranks	
Placebo	New Drug	Placebo	New Drug	Placebo	New Drug
7	3		1		1
5	6		2		2
6	4		3		3
4	2	4	4	4.5	4.5
12	1	5		6	
		6	6	7.5	7.5
		7		9	
		12		10	

Mann Whitney U Test (Wilcoxon Rank Sum Test)

		Total Sample (Ordered Smallest to Largest)		Ranks	
Placebo	New Drug	Placebo	New Drug	Placebo	New Drug
7	3		1		1
5	6		2		2
6	4		3		3
4	2	4	4	4.5	4.5
12	1	5		6	
		6	6	7.5	7.5
		7		9	
		12		10	

$$U_1 = n_1 n_2 + \frac{n_1(n_1+1)}{2} - R_1$$

$$U_2 = n_1 n_2 + \frac{n_2(n_2+1)}{2} - R_2$$

$$U_1 = 5(5) + \frac{5(6)}{2} - 37 = 3$$

$$U_2 = 5(5) + \frac{5(6)}{2} - 18 = 22$$

考慮以下狀況

- 若接受新藥受試者的發作次數明顯較少：

$$R_1 = 6 + 7 + 8 + 9 + 10 = 40 \text{ and } R_2 = 1 + 2 + 3 + 4 + 5 = 15$$

$$U_1 = 5 \binom{5}{2} + \frac{5(6)}{2} - 40 = 0 \text{ and } U_2 = 5 \binom{5}{2} + \frac{5(6)}{2} - 15 = 25$$

- 與服用安慰劑的受試者與新藥受試者發作次數差不多

$$R_1 = 2 + 4 + 6 + 8 + 10 = 30 \text{ and } R_2 = 1 + 3 + 5 + 7 + 9 = 25$$

$R_1 = 2+4+6+8+10 = 30 \text{ and } R_2 = 1+3+5+7+9 = 25,$

and

$$U_1 = 5 \binom{5}{2} + \frac{5(6)}{2} - 30 = 10 \text{ and } U_2 = 5 \binom{5}{2} + \frac{5(6)}{2} - 25 = 15$$

較小的U 支持對立假設
較大的U 支持虛無假設

Wilcoxon Rank Sum Test

```
placebo    <- c(7,5,6,4,12)
new_drug   <- c(3,6,4,2,1 )
wilcox.test(placebo, new_drug)
```

Wilcoxon rank sum test with continuity correction

data: placebo and new_drug

$W = 22$, p-value = 0.05855

alternative hypothesis: true location shift is not equal to 0

配對樣本檢定 (Matched Sample Test)

- 一項臨床調查評估一種減少自閉症兒童重複行為的新藥的有效性。如果藥物有效，兒童在治療時的重複行為會比未治療時減少
- 共有8名自閉症兒童參與這項研究。每個孩子在治療前和服用新藥1周後都要接受研究心理學家觀察3小時。測量每個孩子在3小時中進行重複行為的時間
- 重複行為的得分在0到100分之間，分數代表孩子在觀察時間中重複行為的百分比
 - 例如，分數為0表示在整個觀察期間，孩子沒有重複行為；分數為100則表示孩子一直在重複同樣行為

檢定配對樣本的差異

H_0 : 兩樣本的中位數相同
 H_1 : 兩樣本的中位數有差異

Child	Before Treatment	After 1 Week of Treatment	Difference (Before-After)
1	85	75	10
2	70	50	20
3	40	50	-10
4	65	40	25
5	80	20	60
6	75	65	10
7	55	40	15
8	20	25	-5

Sign Test

Child	Before Treatment	After 1 Week of Treatment	Difference (Before-After)	Sign
1	85	75	10	+
2	70	50	20	+
3	40	50	-10	-
4	65	40	25	+
5	80	20	60	+
6	75	65	10	+
7	55	40	15	+
8	20	25	-5	-

如果有差異
應該有更多的+

Sign Test

$$P(x \text{ successes}) = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$$

x=Number of Successes	P(x successes)
0	0.0039
1	0.0313
2	0.1094
3	0.2188
4	0.2734
5	0.2188
6	0.1094
7	0.0313
8	0.0039

$$P(x \leq 2) = P(0) + P(1) + P(2) = 0.0039 + 0.0313 + 0.1094 = 0.1446$$

觀察到兩個負號的機率為 $0.1446 > 0.5$
代表接受治療後的效果並不顯著

Sign Test

```
binom.test(2, 8, alternative = 'less')
```

Exact binomial test

data: 2 and 8

number of successes = 2, number of trials = 8, p-value = 0.1445

alternative hypothesis: true probability of success is less than 0.5

95 percent confidence interval:

0.0000000 0.5996894

sample estimates:

probability of success

0.25

Wilcoxon Signed Rank Test

Child	Before Treatment	After 1 Week of Treatment	Difference (Before-After)
1	85	75	10
2	70	50	20
3	40	50	-10
4	65	40	25
5	80	20	60
6	75	65	10
7	55	40	15
8	20	25	-5



Observed Differences	Ordered Absolute Values of Difference Scores	Ranks	Signed Ranks
10	-5	1	-1
20	10	3	3
-10	-10	3	-3
25	10	3	3
60	15	5	5
10	20	6	6
15	25	7	7
-5	60	8	8

Wilcoxon Signed Rank Test

- 虛無假設為真，將預期看到正負向高秩(排名前面)與低秩(排名後面)的數量相近
 - 例如： $W+$ 數量接近 $W-$
- 對立假設為真，將有更多正向高秩 (排名較前)的結果
 - 例如： $W+$ 比 $W-$ 的數量多很多

Wilcoxon Signed Rank Test

```
before <- c(85,70,40,65,80,75,55,20)
after  <- c(75,50,50,40,20,65,40,25)
wilcox.test(before, after,paired=TRUE)
```

Wilcoxon signed rank test with continuity correction

data: before and after

$V = 32$, p-value = 0.05747

alternative hypothesis: true location shift is not equal to 0

The Kruskal-Wallis Test

- Kruskal Wallis test 是一無母數統計方法用來比較超過兩個的獨立群體的中位數
- Kruskal Wallis test類似變異數分析 ANOVA. 惟將原始資料替代成秩 (Rank)
- 假設檢定如下:
 - H_0 : K 個群組的中位數相同
 - H_1 : K 個群組的中位數不同

The Kruskal-Wallis Test

- 一項臨床研究評估不同蛋白質含量的飲食對成年人白蛋白水準的影響。低蛋白飲食通常用於腎功能衰竭患者。白蛋白是血液中最豐富的蛋白質，它在血清中的濃度是以克每分升(g/dL)來測定的。臨床上，血清白蛋白濃度也被用來評估患者是否在他們的飲食中獲得足夠的蛋白質

5% Protein	10% Protein	15% Protein
3.1	3.8	4.0
2.6	4.1	5.5
2.9	2.9	5.0
	3.4	4.8
	4.2	

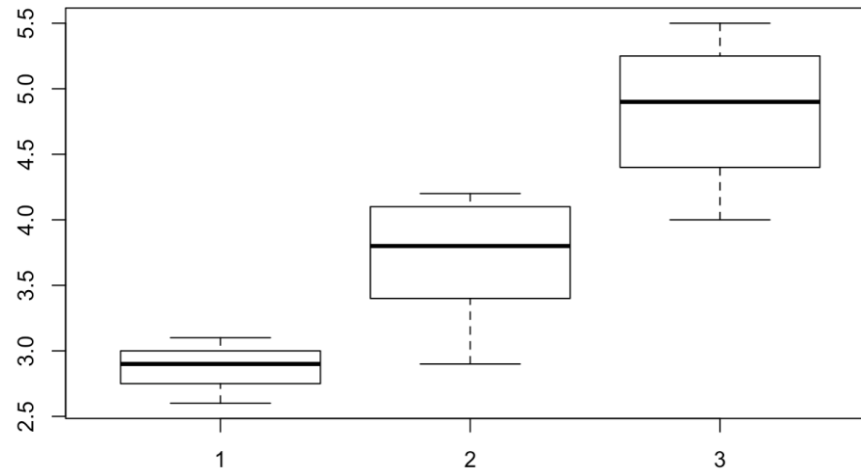
- 三種不同飲食的受試者血清白蛋白水準是否存在差異？
 - ▣ 參考文獻，正常白蛋白水準一般在3.4 ~ 5.4 g/dL之間
- 結果顯示，15%蛋白飲食的參與者比5%蛋白飲食的參與者有更高的白蛋白水準。問題是觀察到的差異是否顯著

繪製箱型圖

```
albumin <- c(3.1, 2.6, 2.9, 3.8, 4.1, 2.9, 3.4, 4.2, 4.5, 5.5, 5, 4.8)
```

```
group <- c(1, 1, 1, 2, 2, 2, 2, 2, 3, 3, 3, 3)
```

```
boxplot(albumin ~ group)
```



The Kruskal-Wallis Test

```
kruskal.test(albumin~ group)
```

Kruskal-Wallis rank sum test

data: albumin by group

Kruskal-Wallis chi-squared = 7.5495, df = 2, p-value = 0.02294

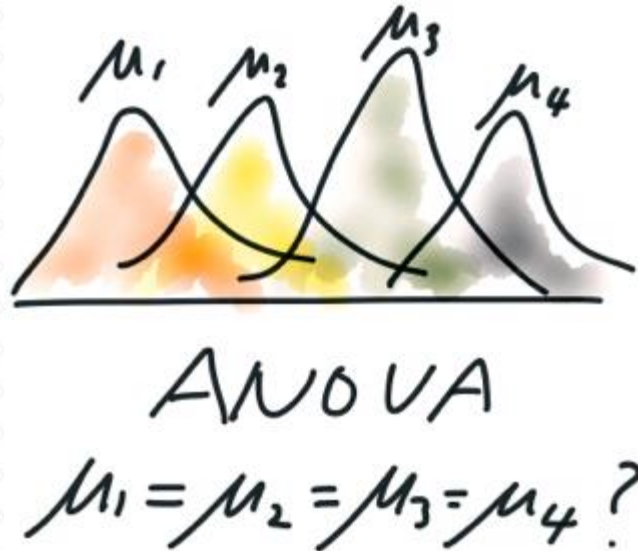
拒絕均值相等的原假設

變異數分析

變異數分析 (ANOVA)

■ ANOVA (Analysis of Variance)

- 檢視類別的自變數與連續依變數之間的關係，用以檢視不同組(適用於多組資料)之間的平均是否相同？



變異數分析 (ANOVA)

- 單因子變異數
- 雙因子變異數
 - 存在交互影響
 - 不存在交互影響
- 多因子變因數
 - 存在交互影響
 - 不存在交互影響

變異數分析 (ANOVA)

- 假設有四組資料(A, B, C, D)，我們現在要比較他們血壓的收縮壓是否有統計上的顯著差異

	Group 1	Group 2	Group 3	Group 4
Sample Size	n_1	n_2	n_3	n_4
Sample Mean	\bar{X}_1	\bar{X}_2	\bar{X}_3	\bar{X}_4
Sample Standard Deviation	s_1	s_2	s_3	s_4

- $H_0: \mu_1 = \mu_2 = \mu_3 \dots = \mu_k$
- H_1 : 組別間的均值不等.

單因子變異數分析

■ 計算步驟

- 計算各組平均數與總平均數
- 計算組內離差平方和 S_e ，組間離差平方和 S_B ，總離差平方和 S_T
 - $S_T = S_e + S_B$ ， S_e 與 S_B 互相獨立
 - 自由度滿足 $f_T = f_e + f_B$
- 將各離差平方和除以對應自由度得到平均離差平方和(均方)
 - $MS_B = S_B / f_B$
 - $MS_e = S_e / f_e$
- 計算F 統計量
 - $F = MS_B / MS_e$
- 使用F 推導P值

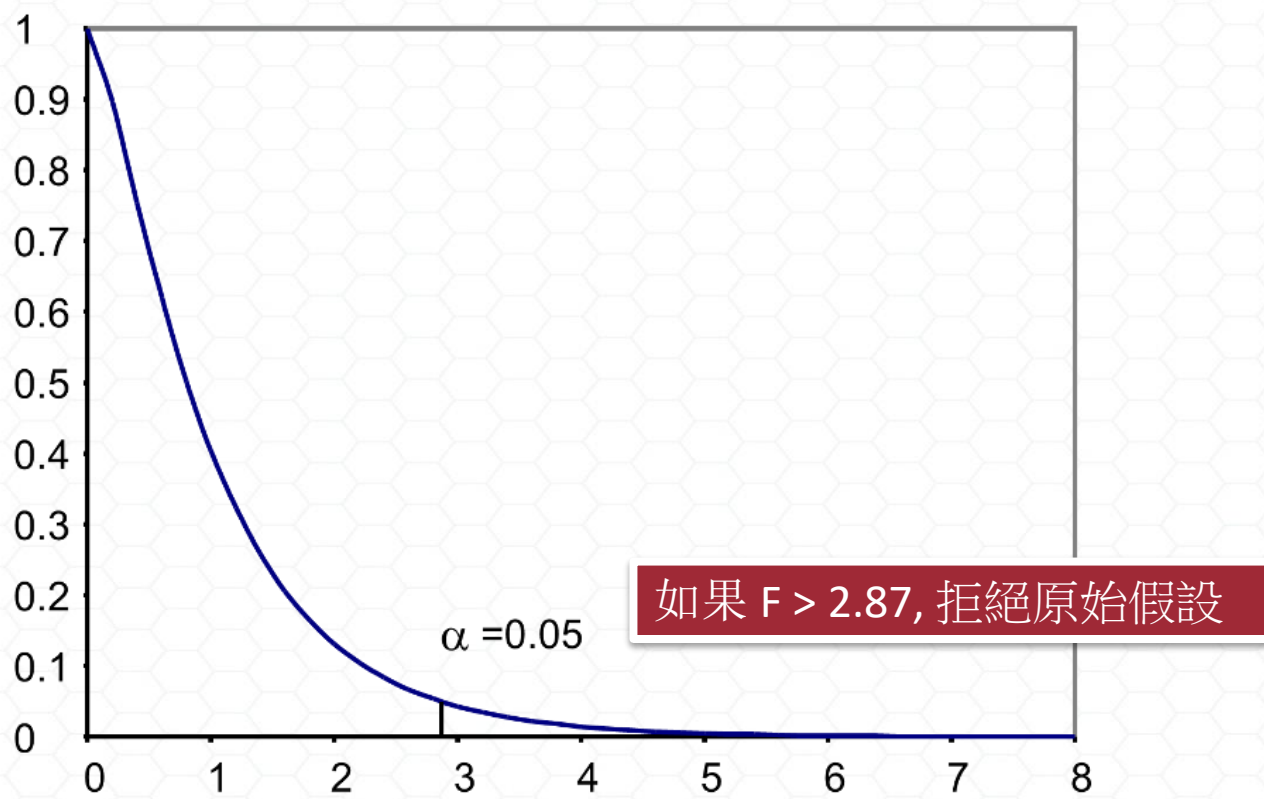
變異數分析計算步驟

Source of Variation	Sums of Squares (SS)	Degrees of Freedom (df)	Mean Squares (MS)	F
Between Treatments	$SSB = \sum n_j (\bar{X}_j - \bar{X})^2$	k-1	$MSB = \frac{SSB}{k-1}$	$F = \frac{MSB}{MSE}$
Error (or Residual)	$SSE = \sum \sum (X - \bar{X}_j)^2$	N-k	$MSE = \frac{SSE}{N-k}$	
Total	$SST = \sum \sum (X - \bar{X})^2$	N-1		

- X = individual observation,
- \bar{X}_j = sample mean of the jth treatment (or group),
- \bar{X} = overall sample mean,
- k = the number of treatments or independent comparison groups, and
- N = total number of observations or total sample size.

F統計圖

■ F Test with $\alpha = 0.05$, $df_1 = 3$ and $df_2 = 36$ ($k = 4$, $N = 40$)



分析實例

- 比較不同減肥計畫帶來的成效，參與者被隨機分配到其中一個實驗組別，並就指定計劃的細節進行減肥諮詢
- 研究測量減肥前的體重(基線)和研究結束後體重(8周)的差異
- 研究計畫分三組:
 - 首先是低熱量的飲食
 - 第二種是低脂肪飲食
 - 第三種是低碳水化合物飲食
 - 第四組被視為對照組

Low Calorie	Low Fat	Low Carbohydrate	Control
8	2	3	2
9	4	5	2
6	3	4	-1
7	5	2	0
3	1	3	3

檢視不同組別的減重結果

```
wt1 <- c(8,9,6,7,3)
```

```
wt2 <- c(2,4,3,5,1)
```

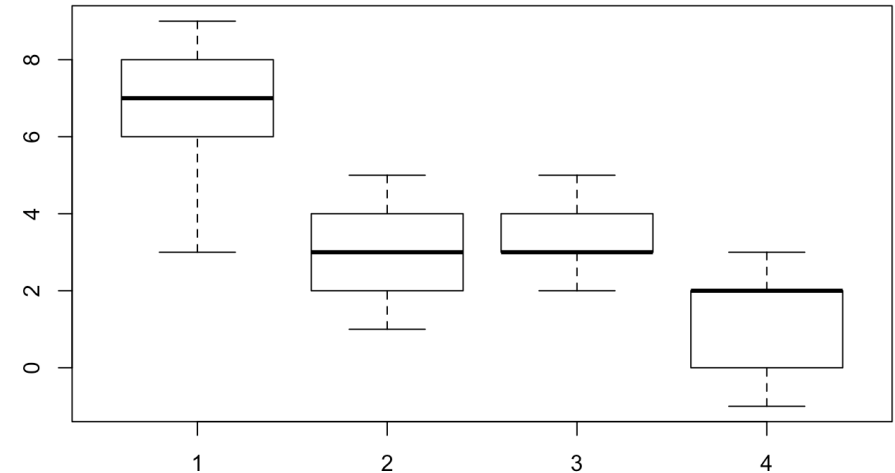
```
wt3 <- c(3,5,4,2,3)
```

```
wt4 <- c(2,2,-1,0,3)
```

```
weights <- c(wt1, wt2, wt3, wt4)
```

```
group <- c(rep(1,5),rep(2,5),rep(3,5),rep(4,5))
```

```
boxplot(weights ~ group )
```



執行ANOVA 檢驗

```
oneway.test(weights ~ group , var.equal = TRUE)
```

One-way analysis of means

data: weights and group

F = 8.5593, num df = 3, denom df = 16, p-value = 0.001278

拒絕原假設
證明不同減肥方法在效果上有差異

雙因子變異分析

- 比較三種治療關節疼痛的治療方法在骨關節炎患者的疼痛緩解時間。研究人員推測，男性和女性的疼痛緩解時間可能存在差異，於是他們隨機分配15名參與男性，與15名女性參與其中一項治療
- 參與試驗的男性和女性不知道他們接受了哪種治療。他們被要求在關節疼痛時服用指定的藥物，並記錄時間，以分鐘為單位，直到疼痛消退

讀取不同組別數據

```
MaleA <- c(22,25,26,27,24)
```

```
FemaleA <- c(21,19,18,24,25)
```

```
MaleB <- c(14, 17, 19, 20, 17)
```

```
FemaleB <- c(21, 20 , 23, 27, 25)
```

```
MaleC <- c(15, 16, 19, 14, 12)
```

```
FemaleC <- c(37, 34, 36, 26, 29)
```

```
timespan <- c(MaleA, FemaleA, MaleB, FemaleB, MaleC, FemaleC)
```

```
sex <- c(rep('Male', 5),rep('Female', 5),rep('Male', 5),  
        rep('Female', 5),rep('Male', 5),rep('Female', 5))
```

```
treatment <- c(rep('A',10),rep('B',10),rep('C',10))
```

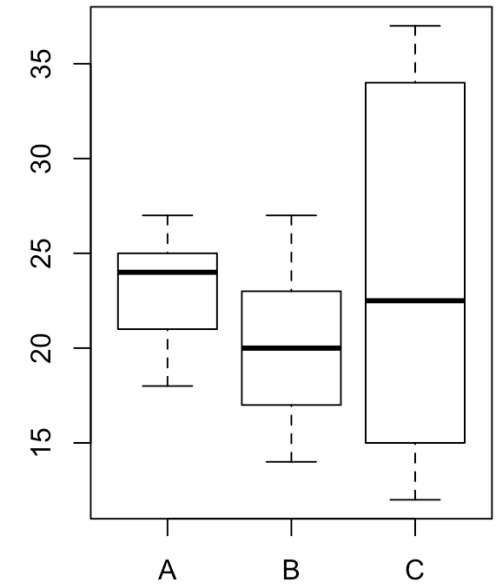
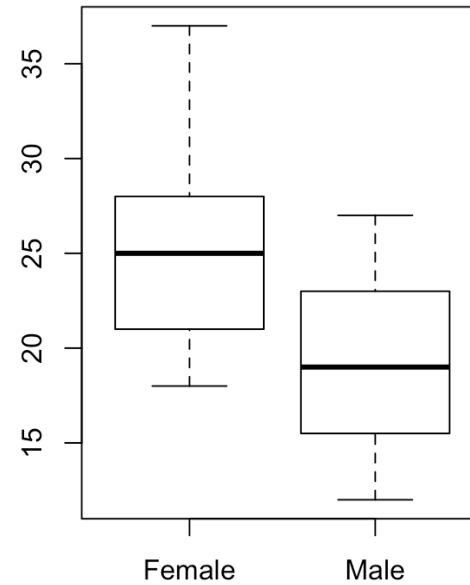
Treatment	Male	Female
A	12	21
	15	19
	16	18
	17	24
	14	25
B	14	21
	17	20
	19	23
	20	27
	17	25
C	25	37
	27	34
	29	36
	24	26
	22	29

比較不同組的數據

```
par(mfrow=c(1,2))
```

```
boxplot(timespan ~ sex)
```

```
boxplot(timespan ~ treatment)
```



Two-Way ANOVA

```
anova_res <- aov(timespan~factor(sex)*factor(treatment))  
summary(anova_res)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
factor (sex)	1	320.1	320.1	34.609	4.55e-06	***
factor (treatment)	2	68.6	34.3	3.708	0.0395	*
factor (sex) : factor (treatment)	2	532.5	266.2	28.782	4.21e-07	***
Residuals	24	222.0	9.3			

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

The background features a light gray hexagonal grid pattern. Overlaid on this is a series of concentric, semi-transparent circles in shades of light blue and white. The circles have a slightly irregular, hand-drawn appearance. A solid dark blue horizontal line runs across the top of the image, and a darker, textured blue horizontal band runs across the bottom.

THANK YOU