統計方法

- 敘述性統計量(Descriptive Statistics)

  →例如平均數、標準差、所佔比例(圖表)等

- 相關性分析

  → 比較平均數(Z 或 t 檢定)、變異數

  → 線性相關係數(Correlation Coefficient)

- 卡方檢定(Chi-Square Test)

  →交叉分析(獨立性檢定)

- 因素分析(Factor Analysis)

  →資料精簡(Data Reduction)與詮釋(Interpretation)

- 其他方法

  → 迴歸(Regression)

  → 時間數列(Time Series)

  → 存活分析(Survival Analysis)

  → 類別資料分析(Categorical Data Analysis)

常見的敘述性統計量(Statistics):

1. 集中趨勢量數(Central Tendency)
2. 差異量數(Dispersion)

● 集中趨勢量數:

1. 平均數(Mean;期望值)
   (a) 算術平均數(Average)
   (b) 加權平均數(Weighted Average)
   e.g.(a)全班的平均身高
   (b)本學期的平均成績
   (c)其他(幾何平均數、調合平均數)

2. 中位數(Median): 一半的數值比中位數大，一半的數值比中位數小。

e.g. (a)員工薪資為 25, 30, 30, 30, 35, 43, 70, 80, 85→中位數是 35
(b) 25, 30, 30, 30, 35, 43, 70, 80, 85, 90
→中位數是 $\frac{35+43}{2}=39$

3. 眾數(Mode):出現次數最多的數值

e.g.(a)員工薪資為 25, 30, 30, 30, 35, 43, 70, 80, 85→眾數是 30
(b)員工薪資為 25, 30, 30, 35, 38, 43, 43, 80, 85→眾數是 30 及 43(眾數不唯一)

- **差異量數:**

1.全距(Range):

最大與最小數值之差(Range = Max－Min)

2.四分位差(Quartile Deviation):

(a)四分位數(Quartile; Q_1): 3/4 的數值比 Q_1 大，1/4 的數值比 Q_1 小。

(b)四分位差 $= Q_3 - Q_1$

3. 變異數(Variance; σ^2)與標準差(Standard Deviation; σ):

母體變異數: $$\sigma^2 = \frac{\sum_{i=1}^{N}(X_i - \mu)^2}{N}$$

樣本變異數: $$s^2 = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n-1}$$

4. 變異係數(Coefficient of Variance; CV):

$$CV = \frac{\sigma}{\mu} \times 100\%$$

- **假設檢定(Testing Statistical Hypothesis)**

對母體特性建立一個正面及一個反面的敘述，藉由樣本資料判斷假設對錯的過程。

〔註〕正面的敘述為虛無假設(Null Hypothesis; H_0)，而反面的敘述為對立假設(Alternative Hypothesis; H_1)。H_0 通常為與期望結果相反的敘述。

e.g. (a)行政院長的執政能力遭立法委員質疑，據悉已有超過一半的委員不滿意。以 p 代表立法委員不滿意的比例，則假設檢定可定為

$$H_0: p \le \frac{1}{2} \quad \rightarrow \quad \text{單尾檢定（複合假設）}$$

$$H_1: p > \frac{1}{2}$$

(b)某減肥食品公司宣稱其產品 A 較另一公司之產品 B 更有效果。以 μ 代表服用產品 A 與 B 所減輕重量之差，可定義

$$H_0: \mu = 0 \quad \rightarrow \quad \text{簡單假設}$$

$$H_1: \mu > 0$$

(c)美國加州政府宣稱在加州無種族歧視問題，黑人與白人學齡小孩的就學率無差異。根據官方說

法，以 p 代表黑人與白人的就學率之差值，可定義為

$$\begin{cases} H_0: p = 0 \\ H_1: p \neq 0 \end{cases} \quad \rightarrow \quad \text{雙尾檢定}$$

Note：檢定以拒絕 H_0 (Reject H_0)與不拒絕 H_0 為結論。

● 假設檢定的步驟

(1)建立假設 H_0 及 H_1

(2)選擇適當的檢定統計量 $\hat{\theta}$

(3)決定顯著水準 α 並確定 C

(4)抽樣、檢定與結論

● 平均數 μ 的檢定

(1) 大樣本檢定或是 $n \geq 30$（若 $H_0: \mu = \mu_0$）

因為在大樣本下，$\overline{X}_n \sim N(\mu, \frac{\sigma^2}{n})$，也就是說

$$\frac{\overline{X}_n - \mu}{\sigma / \sqrt{n}} \sim N(0,1)$$

在左尾檢定下，即 $H_0: \mu < \mu_0$ ，直觀上 $\overline{X}_n$ 愈小，愈有可能是 H_1 ，也就是

$$C = \left\{ Z < -Z_\alpha \right\}$$

同理， 右尾檢定（若 $H_1: \mu > \mu_0$）時 $C = \left\{ Z > Z_\alpha \right\}$

雙尾檢定（若 $H_1: \mu \neq \mu_0$）時 $C = \left\{ |Z| > Z_{\alpha/2} \right\}$

(2) 小樣本檢定（ $n < 30$ ）

與上例相似，但 Z 值以 t 檢定代替。

e.g. (a) 對台北縣某候選人作支持度調查，若抽樣數為1000 人及 $\alpha = 0.05$，

$H_0: p = \frac{1}{2}$ v.s. $H_1: p > \frac{1}{2}$，則拒絕域應滿足 $C = \{x | x > x_0\}$

其中

$$\mathrm{P}\left(x > x_0 \mid p = \frac{1}{2}\right) = \mathrm{P}\left(\frac{x - 1000 \times \frac{1}{2}}{\sqrt{1000 \times \frac{1}{2} \times \frac{1}{2}}} > \frac{x_0 - 1000 \times \frac{1}{2}}{\sqrt{1000 \times \frac{1}{2} \times \frac{1}{2}}} \mid P = \frac{1}{2}\right)$$

$$= \mathrm{P}\left(Z > \frac{x_0 - 500}{15.81}\right) = 0.05$$

或是 $\frac{x_0 - 500}{15.81} = 1.645 \Rightarrow x_0 = 526$

(b) 延續(a)的敘述，計算對應 P=0.55 的 β 值

$$\beta = \mathrm{P}\left(x \le 256 \mid P = 0.55\right)$$

$$= \mathrm{P}\left(\frac{x - 1000 \times 0.55}{\sqrt{1000 \times 0.55 \times 0.45}} \le \frac{526 - 1000 \times 0.55}{\sqrt{1000 \times 0.55 \times 0.45}} \mid P = 0.55\right)$$

$$= \mathrm{P}\left(Z \le -1.526\right) \approx 0.0636$$

- **P-值（P-Value）法**

P 值法是依樣本觀察值作成拒絕域，進而求出的機率值。例如在左尾檢定時（$H_0: \theta = \theta_0$ vs. $H_1: \theta < \theta_0$）若樣本統計量為 $\hat{\theta}_0$，則

$$P - Value = P\left(\hat{\theta} \le \hat{\theta}_0 \mid H_0\text{為真}\right) = P\left(\hat{\theta} \le \hat{\theta}_0 \mid \theta = \theta_0\right)$$

〔註〕：若 $P - Value < \alpha \Rightarrow$ 拒絕 H_0

（反之，$P - Value \ge \alpha \Rightarrow$ 接受 H_0）

e.g.(c)某工廠工程師為了控制產品品質,抽出重量應為180 公克的 10 罐產品,得出平均數 158 公克及標準差20 公克。

首先定義 $H_0: \mu = 180$ vs. $H_1: \mu \neq 180$

$$而\ P-Value = 2P\left(\frac{\overline{X}_n - 180}{20} \geq \left|\frac{158-180}{20}\right| \mid \mu = 180\right)$$

$$= 2P\big(t(9) \geq 1.1\big) = 2 \times 0.15 = 0.30$$

在 $\alpha = 0.05$ 的狀況下不拒絕 H_0 。

(d) 同上例,但計算 $\overline{X}_n = 158$ 的顯著水準

$$t_0 = \frac{\overline{X}_n - 180}{20} = \frac{158-180}{20} = -1.1$$

$|t_0|$ 小於 $t_{0.975,(9)} = 2.262$,因此不拒絕 H_0 。

- **決策法則:**

①當檢定量落入拒絕域,則拒絕 H_0 。

②當檢定量未落入拒絕域,則不拒絕 H_0 。

其中拒絕域(Critical Region 或 Rejection Region)與接受域(Acceptance Region)相對,與檢定假設相關。另外,臨界值(Critical Value)則為兩個區域的界點。

- **型 I 與型 II 誤差(Type I and Type II Errors)**

型 I 誤差(顯著水準一般記為 α)是 H_0 為真時卻作出拒絕 H_0 的決策;而型 II 誤差(記為 β)則是在 H_1 為真時卻不拒絕 H_0 。

	H_0 為真	H_1 為真
拒絕 H_0	α 誤差	正確決策
接受 H_0	正確決策	β 誤差

→ α 與 β 和拒絕域有絕對關係

α =P(拒絕 H_0 | H_0 為真)

β =P(接受 H_0 | H_1 為真)

〔註〕①固定樣本數，α 與 β 呈現反比關係

②欲使 α 及 β 同時下降，可提高樣本數

{

卡方檢定(Chi-Square Test)

- 卡方檢定主要在處理類別資料的檢定，依序為

(1)**Goodness of Fit Test (適合度檢定)**

(2)**Tests of Independence (獨立性檢定)**

(3)**Tests of Homogeneity(齊一性檢定)**

e.g. (a)調查台北對今年治安的看法，如下表:

看法	人數
比去年好	245
差不多	126
比去年差	713

(b)調查大學生中性別與選系的關係，得出

學院 / 性別	商	工	藝術	公衛
男	21	16	145	8
女	14	4	175	17

(c)調查年紀與買車習性的關係

年齡	大型車	中型車	小型車
20~39	90	18	92
40~59	40	60	100

● **適合度檢定：**

通常用來檢定母體是否為某一特定分配．

e.g. (d)投擲一硬幣 500 次，得出 230 個正面，欲檢定此硬幣為公平硬幣

	正面	反面
觀察值	230	270
理論值	250	250

$(i)Z$ 檢定： $P(|Z| \geq \dfrac{|230-250|}{\sqrt{500\times\frac{1}{2}\times\frac{1}{2}}}=1.789) \cong 0.0736$

$(ii)\chi^2$ 檢定： $\chi^2=\sum_{i=1}^{2}\dfrac{(o_i-e_i)^2}{e_i}=\dfrac{400}{250}+\dfrac{400}{250}=3.2$

其中 o_i, e_i 為第 i 個值的觀察值與理論值。查表得出 $\chi^2_{0.10}(1)=2.706, \chi^2_{0.05}(1)=3.841$ 。

(e)調查 400 個家中有兩個子女的家庭以研究男女嬰出生的機會是否相同．

	男男	男女	女男	女女
觀察值	92	94	110	104
理論值	100	100	100	100

$$\chi^2=\sum_{i=1}^{4}\frac{(o_i-e_i)^2}{e_i}=\frac{64}{100}+\frac{16}{100}+\frac{100}{100}+\frac{16}{100}=1.96<\chi^2_{0.05}(3)=7.815$$

$\Rightarrow \alpha=0.05$ 下我們認為男女嬰出生機會相同．

(f)統計歷年台灣核能電廠的出事率，以確定其發出次數是否為平均每年兩次的布阿松分配（假設資料，共30 年資料）

每年次數	0~1 次	2 次	3 次	4 次	5 次
觀察值	3	5	10	8	4
理論值	5.97	6.72	6.72	5.04	5.55

$$\chi^2=\sum_{i=1}^{5}\frac{\left(o_i-e_i\right)^2}{e_i}\cong 5.69<\chi^2_{0.05}(4)=9.488$$

因此我們不拒絕核能電廠每年出事次數為 Poisson(3) 的假設．

[註]：χ^2檢定要求每個理論值至少為 5，以避免分母太小產生誤判。

●獨立性檢定：

用來檢定母體中的各項特性間是否會互相影響．

e.g. (b)

學院 / 性別	商	工	藝術	公衛	列合計
男	21	16	145	8	190
女	14	4	175	17	210
行合計	35	20	320	25	400

$$\chi^2 = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{(o_{ij}-e_{ij})^2}{e_{ij}}$$

其中o_{ij}與e_{ij}為第 i 列第 j 行觀察值與理論值，$i=1,\ldots,r,\ j=1,\ldots,c$,而自由度為$(r-1)(c-1)$。假設檢定$H_0$:行與列的類別互相獨立

$$\longrightarrow H_0:\ p_{ij}=p_{i.}\,p_{.j},\ i=1,\ldots,r,\ j=1,\ldots,c$$

其中$p_{i.}$為第 i 列平均，$p_{.j}$為第 j 列平均，行列互相獨立的理論值 $p_{ij}=p_{i.}\times p_{.j}$

理論值：

	商	工	藝術	公衛	列
男	0.0416 (16.625)	0.02375 (9.5)	0.38 (152)	0.0297 (11.875)	0.475
女	0.0459 (18.375)	0.02625 (10.5)	0.42 (168)	0.0328 (13.125)	0.525
行	0.0875	0.05	0.8	0.0625	1

χ^2＝13.675 ＞ $\chi^2_{0.05}(3)$＝7.815

因此我們認為性別與選系有關聯。

(g) 隨機抽樣調查 200 位大學生，獲得其就讀年級與吸煙習慣的資料如下：

年級 \ 習慣	一年級	二年級	三年級	四年級
吸煙	21	33	25	20
不吸煙	47	26	19	9

在 $\alpha=0.01$ 下檢定大學生之就讀年級與吸煙習慣是否獨立，可得

$$\chi^2=15.7438 \ > \chi^2_{0.05}(3)=11.34$$

因此我們認為抽煙與就讀的年齡有關聯。

● 齊一性檢定：

與獨立性檢定類似，但列聯表(Contingency Table)中的列（或行）代表不同的母體，而且檢定主旨在於比較各母體的異同。

e.g.(c)欲知不同年齡層的買車習性是否有差異

$(\alpha=0.01)$

車別／年齡	大型車	中型車	小型車
20～39	90	18	92
40～59	40	60	100

$$\chi^2=\frac{625}{65}+\frac{441}{39}+\frac{16}{96}+\frac{625}{65}+\frac{441}{39}+\frac{16}{96}$$

$$=42.179 \ > \ \chi^2_{0.01}(2)=9.210$$

迴歸分析與相關分析

● 迴歸(Regression)

若兩變數的關係屬於線性相關,可以直線方程式表達.

e.g.(a)標準體重=(身高－100)×0.9, 單位:公斤及公分
(b)標準體重=25×(身高)^2, 單位:公斤及公尺

[註]: 1.只要滿足 y = a + bx 的方程式即可視為迴歸,上例中標準體重為 y (因變數; Dependent Variable),身高或(身高)^2 為x (自變數;Independent Variable).
2.若迴歸方程式中只有一個自變數,則為簡單迴歸(Simple Regression) ; 兩個或兩個以上的自變數,為複迴歸(Multiple Regression).

● 模型:

$$Y_i = \mu_i + \varepsilon_i = \beta_0 + \beta_1 X_i + \varepsilon_i \,,$$

$$\varepsilon_i \sim N(0, \sigma^2), \quad i = 1, 2, \ldots, n$$

因此 $E(Y_i) = \beta_0 + \beta_1 X_i$

$Var(Y_i) = Var(\varepsilon_i) = \sigma^2$

其中β_0及β_1參數的求取，以最小平方法(Method of Least Squares)配適， 即

$$\min_{\beta_o, \beta_1} (Y_i - \beta_0 - \beta_1 X_i)^2$$

其中 (X_i, Y_i)為第i組的觀察值.

● 誤差項平方法

$$Q=\sum_{i=1}^{n}\varepsilon_i^2=\sum_{i=1}^{n}\ [Y_i-(\beta_0+\beta_1X_i)]^2$$

依 $\dfrac{\partial Q}{\partial \beta_0}=0$ 及 $\dfrac{\partial Q}{\partial \beta_1}=0$,可求得

$$\hat{\beta}_1=\frac{\sum_{i=1}^{n}(X_i-\overline{X})(Y_i-\overline{Y})}{\sum_{i=1}^{n}\ (X_i-\overline{X})^2}=\frac{S_{xy}}{S_x^2}$$

$$\hat{\beta}_0=\overline{Y}-\hat{\beta}_1\overline{X}$$

e.g.(c)下表為八名售貨員的年資與月銷售金額資料

售貨員	A	B	C	D	E	F	G	H
銷售金額(Y)	9	6	4	3	3	5	8	2
年　　資 (X)	6	5	3	1	4	3	6	2

$$\hat{\beta}_1=\frac{n\sum X_iY_i-(\sum X_i)(\sum Y_i)}{n\sum X_i{}^2-(\sum X_i)^2}=\frac{8\times178-30\times40}{8\times136-(30)^2}=\frac{56}{47}$$

$$\hat{\beta}_0=\frac{40}{8}-(\frac{56}{47})(\frac{30}{8})=\frac{25}{47}$$

(d)下表為1988年紐約市一至七月的地下鐵資料

	一月	二月	三月	四月	五月	六　月	七月
(y)搶劫案	475	465	470	500	550	600	602
(z)破案數	180	155	160	190	225	220	223

y=419+26.1*月份，$R^2=87.1\%$

z=147+11.6*月份，$R^2=69.9\%$

- **判定係數(Coefficient of Determination)**

由$\hat{y}=\alpha+\beta$x的模型預測真正的y值

總變異(SST)

= 迴歸變異(SSR) + 無法解釋之變異(SSE)

或

$$\sum(Y_i-\overline{Y})^2=\sum(\hat{Y}_i-\overline{Y})^2+\sum(Y_i-\hat{Y}_i)^2$$

→判定係數即為由迴歸可解釋之部份

$$R^2=\frac{SSR}{SST}=1-\frac{SSE}{SST},\quad 0\le R^2\le 1$$

- **相關係數(Correlation Coefficient)**

相關係數用來判定兩個變數間線性關係之強度

$$\rho=\frac{E(X_i-\mu_x)(Y_i-\mu_y)}{\sigma_x\sigma_y}=\frac{\sigma_{xy}}{\sigma_x\sigma_y}\rightarrow\text{母體相關係數}$$

$$r=\frac{\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})(Y_i-\overline{Y})}{\sqrt{\sum_{i=1}^{n}(X_i-\overline{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i-\overline{Y})^2}}=\frac{S_{xy}}{S_x^2}\longrightarrow\text{樣本相關係數}$$

無母數統計

(Nonparametric Statistical Methods)

不以母體的母數進行統計推論，也不需要母體的分配
→無母數的推論方法(Distribution-free Methods)

(1)單一樣本的推論方法

I.單一樣本符號檢定(One-sample Sign test)

→檢定中位數(使用二項分配)

e.g.(a)紀錄台北市今年十月及十一月每週三之日平均溫，得到

30,24,27,25,20,23,24,18,22

欲檢定日平均溫之中位數是否為 20 度。

$H_0: p = 0.5$ v.s. $H_1: p \neq 0.5$

→大於 20 為＋，而小於 20 為－，等於 20 則不考慮。

＋＋＋＋＋＋＋－＋ ⇒8 個中有 7 個大於 20

$\therefore P(X \geq 7 及 X \leq 1 | p = 0.5)$

$$= \left[\binom{8}{0} + \binom{8}{1} + \binom{8}{7} + \binom{8}{8}\right] \cdot \left(\frac{1}{2}\right)^8 = \frac{18}{256} = 0.07$$

[註]若總樣本數夠大時(若 np>5 且 n(1-p)>5)

$$用 Z = \frac{X - np}{\sqrt{np(1-p)}} \sim N(0,1)$$

II.Wilcoxon 符號等級檢定

(Wilcoxon Signed-rank test)

與 Sign test 相似，但必須排序處理

e.g.(a)	30	24	27	25	20	23	24	18	22
差別	10	4	7	5		3	4	-2	2
級數	8	4.5	7	6		3	4.5	1.5	1.5

$\therefore T^{+}=34.5$ 及 $T^{-}=1.5 \Rightarrow T=Min(T^{+},T^{-})=1.5$

查表得出在 n=8 時，α =0.05 及雙尾檢定時，

拒絕域為 T<4 $\Rightarrow$拒絕H_0

[註]大樣本之下，

$$E(T)=\frac{n(n+1)}{4}\ ,\ Var(T)=\frac{n(n+1)(2n+1)}{4}$$

再以 $Z=\dfrac{T-E(T)}{\sqrt{Var(T)}}\sim N(0,1)$ 作檢定

III.連檢定(Run test)

檢定一組樣本是否合乎隨機性(Randomness)

e.g.(b)某是非題之標準答案如下，欲檢定題目編排是否為隨機：

<u>○</u> <u>× ×</u> <u>○</u> <u>×</u> <u>○</u> <u>×</u> <u>○ ○</u> <u>×</u> <u>○</u> <u>× ×</u> <u>○</u> <u>×</u> <u>○</u> <u>×</u> <u>○ ○</u> <u>×</u>

以上○者有 $n_1=10$ 及×者有 $n_2=10$，連的個數有 16 個

查表得 $P(R\le 6$ 或 $R\ge 16)$

$=P(R\le 6)+P(R\ge 16)=0.038$

也就是拒絕答案編排為隨機之假設。

[註]←n_1及n_2各為小於及大於中位數之個數

$$當\begin{cases} R_0 \geq \dfrac{2n_1n_2}{n_1+n_2}+1 \Rightarrow p-value = P(R \geq R_0 \mid n_1, n_2) \\ R_0 < \dfrac{2n_1n_2}{n_1+n_2}+1 \Rightarrow p-value = P(R \leq R_0 \mid n_1, n_2) \end{cases}$$

$$\uparrow E(R) = \frac{2n_1n_2}{n_1+n_2}+1$$

$$Var(R) = \frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1+n_2)^2(n_1+n_2-1)}$$

(2)兩組樣本的推論方法

I.兩組樣本是否相同→Wilcoxon Signed-rank test

[註]兩兩作比較，因此兩組之樣本數必須相同

e.g.(c)兩組學生各 12 人作教材訓練之成效比較

A	77	86	43	72	95	79	93	70	54	90	61	69
B	55	46	40	57	92	63	53	70	29	70	80	22
△	22	40	3	15	31	6	40	0	25	20	-19	47

$\Rightarrow T = T^- = 5$

拒絕域 T<11，因此拒絕 A 與 B 相同的假設

II.Mann-Whitney U 檢定

→與 Wlicoxon sined-rank test 類似，但兩個樣本之樣本數不必相同

e.g.(d)以下為某公司男女職員平均午餐消費金額

男	300	150	45	120	80	120	40	
女	70	100	140	150	50	200	60	30

30	40	45	50	60	70	80	100	120	120	140	150	150	200	300
1	2*	3*	4	5	6	7*	8	9.5*	9.5*	11	12.5*	12.5	14	15*

$$\therefore T_1 = 58.5 \Rightarrow U_1 = n_1 n_2 + \frac{n_1(n_1+1)}{2} - T_1 = 25.5$$

$$T_2 = 61.5 \Rightarrow U_2 = n_1 n_2 + \frac{n_2(n_2+1)}{2} - T_2 = 30.5$$

因此$U = \min(U_1, U_2) = 25.5$

查表$(n_1 = 7, n_2 = 8)$得$U \le 13$為拒絕域⇒不拒絕H_0

[註]大樣本之下，Mann-Whitrey U test 有

$$E(U) = \frac{n_1 n_2}{2} \quad , \quad Var(U) = \frac{n_1 n_2 (n_1 + n_2 + 1)}{12}$$

III.Spearman 等級相關係數

參閱迴歸分析

(3)多組樣本的推論方法

I.Kruskal-Wallis 檢定法

與 Mann-Whitney U test 相似，將所有觀察值一起排序，再分別求出各樣本之排序和T_i，代入

$$H=\frac{12}{n(n+1)}\sum_{i=1}^{k}\frac{T_i^2}{n_i}-3(n+1)$$

(n_i為第 i 樣本總數，$n=n_1+n_2+\cdots+n_k$)

在”H_0：k 組樣本之母體分配相同”為真下，
$H\sim\chi^2(k-1)$

II.Friedman 檢定法

與 Kruskal-Wallis 檢定法類似，但適用於隨機區集實驗設計(Random Block Design)，且排序為分集(Block)而定

檢定量 $$F_r=\frac{12}{bk(k+1)}\sum_{i=1}^{k}T_i^2-3b(k+1)$$

在H_0之下，$F_r\sim\chi^2(k-1)$

e.g.(e)

	第一組	第二組	第三組
	59(6)	52(1)	58(4.5)
	64(9.5)	58(4.5)	65(11)
	57(3)	54(2)	71(12)
	62(7)		63(8)
			64(7.5)
T_i	(25.5)	(7.5)	(4.5)

用 Kruskal-Wallis 檢定法得出

$H = 6.101 > \chi^2_{0.05}(2) = 5.99$

(f)

區集	第一組	第二組	第三組
1	25(3)	24(2)	22(1)
2	18(1)	27(3)	24(2)
3	15(1)	24(2)	27(3)
4	21(2)	19(1)	25(3)
5	22(2)	17(1)	30(3)
T_i	(9)	(9)	(12)

得 Friedman 檢定值 $F_r = 1.2 < \chi^2_{0.05}(2) = 5.99$