多變量變異數分析 **MANOVA**

$\boxed{\text{意義}}$

應變數超過 2 個時，所採用的變異數分析方法。

$\boxed{\text{公式}}$

MANOVA 必須計算下列數據：

變異	SSCP 矩陣	S 矩陣	特徵值 λ	統計量
組間	$SSCP_b$	$SSCP_w^{-1}$ x $SSCP_b$	S 矩陣的 λ	Pillai's trace Hotelling's trace Wilk's lambda Roy's largest root
組內	$SSCP_w$			
總變異	$SSCP_t$			

總變異 $SSCP_{total}$ (Sum of Squares & Cross Products Matrix for Total)

$$\begin{bmatrix} \sum_{i=1,k=1}^{n}(x_{ik}-\bar{\bar{X}})^2 & \sum_{i=1,k=1}^{n}(x_{ik}-\bar{\bar{X}})(y_{ik}-\bar{\bar{Y}}) \\ \sum_{i=1,k=1}^{n}(x_{ik}-\bar{\bar{X}})(y_{ik}-\bar{\bar{Y}}) & \sum_{i=1,k=1}^{n}(y_{ik}-\bar{\bar{Y}})^2 \end{bmatrix} \text{................公式①}$$

$$\begin{bmatrix} x_{11}-\bar{\bar{X}} & x_{21}-\bar{\bar{X}} & \cdots & x_{ik}-\bar{\bar{X}} \\ y_{12}-\bar{\bar{Y}} & y_{22}-\bar{\bar{Y}} & \cdots & y_{ik}-\bar{\bar{Y}} \end{bmatrix} \times \begin{bmatrix} x_{11}-\bar{\bar{X}} & y_{12}-\bar{\bar{Y}} \\ x_{21}-\bar{\bar{X}} & y_{22}-\bar{\bar{Y}} \\ \vdots & \vdots \\ x_{ik}-\bar{\bar{X}} & y_{ik}-\bar{\bar{Y}} \end{bmatrix} \text{..................公式②}$$

組內變異 $SSCP_{within}$ (Sum of Squares & Cross Products Matrix for Within Groups)

$$\begin{bmatrix} \sum_{i=1,k=1}^{n}(x_{ik}-\bar{x}_k)^2 & \sum_{i=1,k=1}^{n}(x_{ik}-\bar{x}_k)(y_{ik}-\bar{y}_k) \\ \sum_{i=1,k=1}^{n}(x_{ik}-\bar{x}_k)(y_{ik}-\bar{y}_k) & \sum_{i=1,k=1}^{n}(y_{ik}-\bar{y}_k)^2 \end{bmatrix} \text{........公式①}$$

$$\begin{bmatrix} x_{11}-\bar{x}_1 & x_{21}-\bar{x}_1 & \cdots & x_{i1}-\bar{x}_1 \\ y_{12}-\bar{y}_2 & y_{22}-\bar{y}_2 & \cdots & y_{i2}-\bar{y}_2 \end{bmatrix} \times \begin{bmatrix} x_{11}-\bar{x}_1 & y_{12}-\bar{y}_2 \\ x_{21}-\bar{x}_1 & y_{22}-\bar{y}_2 \\ \vdots & \vdots \\ x_{i1}-\bar{x}_1 & y_{i2}-\bar{y}_2 \end{bmatrix} \text{...........公式②}$$

組間變異 $SSCP_{between}$ (Sum of Squares & Cross Products Matrix for Between Groups)

$$SSCP_{between} = SSCP_{total} - SSCP_{within}$$

i: 第 i 個樣本觀察值。

k: 第 k 組。

x_{ik}: x 變數中第 k 組中的第 i 個觀察值，例x_{21}為第 1 組中第 2 個觀察值。

y_{ik}: y 變數中第 k 組中的第 i 個觀察值，例y_{12}為第 2 組中第 1 個觀察值。

$\bar{x}_k$: 第 k 組中 x 變數的算術平均數。

$\bar{y}_k$: 第 k 組中 y 變數的算術平均數。
$\bar{\bar{X}}, \bar{\bar{Y}}$: X, Y 變數的算術平均數。

$\boxed{\text{範例}}$
下表為感染呼吸道融合病毒及幽門螺旋桿菌病患，身上所測得的血液生物標記及體溫。以 MANOVA 來驗證能否從抽血及體溫，來區別不同的感染源。

Disease	Biomarker (mg/l)	Temperature (℃)
RSV	40.0	36.0
RSV	11.1	37.2
RSV	30.0	36.5
RSV	21.4	39.4
RSV	10.7	39.6
RSV	3.4	40.7
pylori	42.0	37.6
pylori	31.1	42.2
pylori	50.0	38.5
pylori	60.4	39.4
pylori	45.7	38.6
pylori	17.3	42.7

資料取自：https://youtu.be/CBXYxs9pLW8

設 X =血液生物標記、Y= 體溫 、k=1=RSV 組、k=2=pylori 組，分別計算 X、Y 的總平均，以及分組的平均數：

$$\bar{\bar{X}} = \frac{40 + 11.1 + 30 + 21.4 + 10.7 + 3.4 + 42 + 31.1 + 50 + 60.4 + 45.7 + 17.3}{12}$$
$$= 30.3$$

$$\bar{\bar{Y}} = \frac{36 + 37.2 + 36.5 + 39.4 + 39.6 + 40.7 + 37.6 + 42.2 + 38.5 + 39.4 + 38.6 + 42.7}{12}$$
$$= 39$$

$$\overline{x_{RSV}} = \frac{40 + 11.1 + 30 + 21.4 + 10.7 + 3.4}{6} = 19.4$$

$$\overline{x_{pylori}} = \frac{42 + 31.1 + 50 + 60.4 + 45.7 + 17.3}{6} = 41.1$$

$$\overline{y_{RSV}} = \frac{36 + 37.2 + 36.5 + 39.4 + 39.6 + 40.7}{6} = 38.2$$

$$\overline{y_{pylori}} = \frac{37.6 + 42.2 + 38.5 + 39.4 + 38.6 + 42.7}{6} = 39.8$$

1. SSCP 矩陣：總變異

首先計算總變異矩陣，先分別計算 X 與 Y 的離差平方和、X 與 Y 的離差交乘積：

$$\sum_{i=1,k=1}^{n} (x_{ik} - \bar{\bar{X}})^2$$

$=(40 - 30.3)^2 + (11.1 - 30.3)^2 + (30 - 30.3)^2 + (21.4 - 30.3)^2 + (10.7 - 30.3)^2 + (3.4 - 30.3)^2 + (42 - 30.3)^2 + (31.1 - 30.3)^2 + (50 - 30.3)^2 + (60.4 - 30.3)^2 + (45.7 - 30.3)^2 + (17.3 - 30.3)^2$= 3488

$$\sum_{i=1,k=1}^{n} (y_{ik} - \bar{\bar{Y}})^2$$

$=(36 - 39)^2 + (37.2 - 39)^2 + (36.5 - 39)^2 + (39.4 - 39)^2 + (39.6 - 39)^2 + (40.7 - 39)^2 + (37.6 - 39)^2 + (42.2 - 39)^2 + (38.5 - 39)^2 + (39.4 - 39)^2 + (38.6 - 39)^2 + (42.7 - 39)^2$= 48

$$\sum_{i=1,k=1}^{n} (x_{ik} - \bar{\bar{X}})(y_{ik} - \bar{\bar{Y}})$$

$$\begin{aligned}
&= \{(40 - 30.3) \times (36 - 39)\} + \{(11.1 - 30.3) \times (37.2 - 39)\} \\
&+ \{(30 - 30.3) \times (36.5 - 39)\} + \{(21.4 - 30.3) \times (39.4 - 39)\} \\
&+ \{(10.7 - 30.3) \times (39.6 - 39)\} + \{(3.4 - 30.3) \times (40.7 - 39)\} \\
&+ \{(42 - 30.3) \times (37.6 - 39)\} + \{(31.1 - 30.3) \times (42.2 - 39)\} \\
&+ \{(50 - 30.3) \times (38.5 - 39)\} + \{(60.4 - 30.3) \times (39.4 - 39)\} \\
&+ \{(45.7 - 30.3) \times (38.6 - 39)\} + \{(17.3 - 30.3) \times (42.7 - 39)\} \\
&= -121
\end{aligned}$$

故 **SSCPtotal** $= \begin{bmatrix} 3488 & -121 \\ -121 & 48 \end{bmatrix}$

2. SSCP 矩陣：組內變異

接下計算組內變異矩陣。

$SSCP_{within} = SSCP_{RSV} + SSCP_{pylori}$，故需要分別計算$SSCP_{RSV}$與$SSCP_{pylori}$。

首先計算 k=1=RSV 組：

$$\sum_{i=1,k=1}^{n}(x_{ik}-\bar{x}_k)^2\ for\ Biomarker$$

$$= (40-19.4)^2 + (11.1-19.4)^2 + (30-19.4)^2 + (21.4-19.4)^2 + (10.7-19.4)^2 + (3.4-19.4)^2 = 941.3$$

$$\sum_{i=1,k=1}^{n}(x_{ik}-\bar{x}_k)^2\ for\ temperature$$

$$= (36-38.2)^2 + (37.2-38.2)^2 + (36.5-38.2)^2 + (39.4-38.2)^2 + (39.6-38.2)^2 + (40.7-38.2)^2 = 18.4$$

$$\sum_{i=1,k=1}^{n}(x_{ik}-\bar{x}_k)(y_{ik}-\bar{y}_k)$$

$$= (20.6\times -2.2) + (-8.3\times -1) + (10.6\times -1.7) + (2\times 1.2) + (-8.7\times 1.4) + (-16\times 2.5) = -104.8$$

故$SSCP_{RSV} = \begin{bmatrix} 941.3 & -104.8 \\ -104.8 & 18.4 \end{bmatrix}$

接下來計算 k=2=pylori 組：

$$\sum_{i=1,k=2}^{n}(y_{ik}-\bar{y}_k)^2\ for\ Biomarker$$

$$= (42-41.1)^2 + (31.1-41.1)^2 + (50-41.1)^2 + (60-41.1)^2 + (45.7-41.1)^2 + (17.3-41.1)^2 = 1140.1$$

$$\sum_{i=1,k=2}^{n}(y_{ik}-\bar{y}_k)^2 \; for\ temperature = (37.6-39.8)^2+(42.2-39.8)^2 + (38.5-39.8)^2+(39.4-39.8)^2+(38.6-39.8)^2 + (42.7-39.8)^2 = 22.3$$

$$\sum_{i=1,k=2}^{n}(x_{ik}-\bar{x}_k)(y_{ik}-\bar{y}_k) = (0.9\times -2.2)+(10\times 2.4)+(8.9\times -1.3)+(19.3\times -0.4) + (-4.6\times -1.2)+(-23.8\times 2.9) = -119.8$$

故$SSCP_{pylori} = \begin{bmatrix} 1140.1 & -119.8 \\ -119.8 & 22.3 \end{bmatrix}$

故$SSCP_{within}$

$$= \begin{bmatrix} 941.3 & -104.8 \\ -104.8 & 18.4 \end{bmatrix} + \begin{bmatrix} 1140.1 & -119.8 \\ -119.8 & 22.3 \end{bmatrix}$$
$$= \begin{bmatrix} 2081 & -225 \\ -225 & 41 \end{bmatrix}$$

3. SSCP 矩陣：組間變異

有了總變異跟組內變異後，根據公式 $SSCP_{total} = SSCP_{within} + SSCP_{between}$計算組間變異矩陣。

故$SSCP_{between}$

$$= \begin{bmatrix} 3488 & -121 \\ -121 & 48 \end{bmatrix} - \begin{bmatrix} 2081 & -225 \\ -225 & 41 \end{bmatrix} = \begin{bmatrix} 1407 & 104 \\ 104 & 7 \end{bmatrix}$$

4. S 矩陣

由 3 個 SSCP 可以計算矩陣積 $S = SSCP_{within}^{-1} \times SSCP_{between}$。

首先先求得組內變異矩陣的反矩陣：

$SSCP_{within}^{-1}$

$$=\frac{1}{2081\times 41-(-225)\times(-225)}\times\begin{bmatrix}41 & 225\\ 225 & 2081\end{bmatrix}$$

$$=\begin{bmatrix}0.00118 & 0.00648\\ 0.00648 & 0.05998\end{bmatrix}$$

故矩陣積 S= $\begin{bmatrix}\mathbf{0.00118} & \mathbf{0.00648}\\ \mathbf{0.00648} & \mathbf{0.05998}\end{bmatrix}\times\begin{bmatrix}\mathbf{1407} & \mathbf{104}\\ \mathbf{104} & \mathbf{7}\end{bmatrix}=\begin{bmatrix}\mathbf{2.34} & \mathbf{0.17}\\ \mathbf{15.36} & \mathbf{1.09}\end{bmatrix}$

5. 特徵值 λ

設 A 為矩陣、$\vec{v}$為特徵向量、λ 為特徵值、I 為單位矩陣

$\because\ \mathrm{A}\vec{v}=\lambda\vec{v}$

$\therefore\ \mathrm{A}\vec{v}-\lambda\vec{v}=0$

$(\mathrm{A}-\lambda\mathrm{I})\vec{v}=0$

$det|A-\lambda I|=0$

根據上述定義，求$\begin{bmatrix}2.34 & 0.17\\ 15.36 & 1.09\end{bmatrix}$的$eigenvalue$ λ

$$\begin{bmatrix}2.34 & 0.17\\ 15.36 & 1.09\end{bmatrix}-\lambda\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}=\begin{bmatrix}2.34 & 0.17\\ 15.36 & 1.09\end{bmatrix}-\begin{bmatrix}\lambda & 0\\ 0 & \lambda\end{bmatrix}=\begin{bmatrix}2.34-\lambda & 0.17-0\\ 15.36-0 & 1.09-\lambda\end{bmatrix}$$

$det\begin{vmatrix}2.34-\lambda & 0.17-0\\ 15.36-0 & 1.09-\lambda\end{vmatrix}$=0

$(2.34-\lambda)(1.09-\lambda)-0.17\times 15.36=0$

$\lambda^2-3.43\lambda-0.03=0$

故 λ= $\mathbf{3.45}$ **或** $\mathbf{-0.01}$

6. 統計量

Pillai's trace= $\sum_{k=1}^{k}\frac{\boldsymbol{\lambda}_k}{1+\boldsymbol{\lambda}_k}=\frac{3.45}{1+3.45}+\frac{-0.01}{1+(-0.01)}=0.78$

Hotelling's trace= $\sum_{k=1}^{k}\lambda_k=3.45+(-0.01)=3.44$

Wilk's lambda= $\prod_{k=1}^{k}\frac{1}{1+\lambda_k}=\frac{1}{1+3.45}\times\frac{1}{1+(-0.01)}=\ 0.22$

Roy's largest root= λ_{max}= 3.45

附表 SSCP 計算過程

Disease (k)	**Bio (x)**	**Temp (y)**	$x-\bar{\bar{X}}$ ①	$y-\bar{\bar{Y}}$ ②	$(x-\bar{\bar{X}})^2$	$(y-\bar{\bar{Y}})^2$	①×②	**Bio (x)**		**Temp (y)**		③²	④²	⑤²	⑥²	③×⑤	④×⑥
								$x-\bar{x}$ ③	$y-\bar{y}$ ④	$x-\bar{x}$ ⑤	$y-\bar{y}$ ⑥						
RSV	40	36	9.7	-3.0	94.9	9.2	-29.5	20.6		-2.2		423.0		5.0		-45.9	
RSV	11.1	37.2	-19.2	-1.8	367.0	3.4	35.1	-8.3		-1.0		69.4		1.1		8.6	
RSV	30	36.5	-0.3	-2.5	0.1	6.4	0.7	10.6		-1.7		111.7		3.0		-18.3	
RSV	21.4	39.4	-8.9	0.4	78.5	0.1	-3.2	2.0		1.2		3.9		1.4		2.3	
RSV	10.7	39.6	-19.6	0.6	382.5	0.3	-11.1	-8.7		1.4		76.3		1.9		-11.9	
RSV	3.4	40.7	-26.9	1.7	721.4	2.8	-44.8	-16.0		2.5		257.1		6.1		-39.5	
pylori	42	37.6	11.7	-1.4	137.9	2.1	-16.8		0.9		-2.2		0.8		5.0		-2.0
pylori	31.1	42.2	0.8	3.2	0.7	10.0	2.7		-10.0		2.4		99.7		5.6		-23.6
pylori	50	38.5	19.7	-0.5	389.7	0.3	-10.5		8.9		-1.3		79.5		1.8		-11.9
pylori	60.4	39.4	30.1	0.4	908.5	0.1	11.1		19.3		-0.4		373.1		0.2		-8.4
pylori	45.7	38.6	15.4	-0.4	238.4	0.2	-6.7		4.6		-1.2		21.3		1.5		-5.7
pylori	17.3	42.7	-13.0	3.7	167.9	13.4	-47.5		-23.8		2.9		565.6		8.2		-68.2
Σ	363.1	468.4			3488	48	-121					941.3	1140.1	18.4	22.3	-104.8	-119.8
μ_all	30.3	39.0															
μ_RSV	19.4	38.2															
μ_pylori	41.1	39.8															

附錄 矩陣計算 SSCP

設離差矩陣為 A，反矩陣為A^{-1}，SSCP 可以直接從$A^{-1} \times A$求得：

$$\begin{bmatrix} x_{11}-\bar{\bar{X}} & x_{21}-\bar{\bar{X}} & \cdots & x_{ik}-\bar{\bar{X}} \\ y_{12}-\bar{\bar{Y}} & y_{22}-\bar{\bar{Y}} & \cdots & y_{ik}-\bar{\bar{Y}} \end{bmatrix} \times \begin{bmatrix} x_{11}-\bar{\bar{X}} & y_{12}-\bar{\bar{Y}} \\ x_{21}-\bar{\bar{X}} & y_{22}-\bar{\bar{Y}} \\ \vdots & \vdots \\ x_{ik}-\bar{\bar{X}} & y_{ik}-\bar{\bar{Y}} \end{bmatrix}$$

$$SSCP_{Total} = \begin{bmatrix} 9.7 & -19.2 & -0.3 & -8.9 & -19.6 & -26.9 & 11.7 & 0.8 & 19.7 & 30.1 & 15.4 & -13 \\ -3 & -1.8 & -2.5 & 0.4 & 0.6 & 1.7 & -1.4 & 3.2 & -0.5 & 0.4 & -0.4 & 3.7 \end{bmatrix} \times \begin{bmatrix} 9.7 & -3 \\ -19.2 & -1.8 \\ -0.3 & -2.5 \\ -8.9 & 0.4 \\ -19.6 & 0.6 \\ -26.9 & 1.7 \\ 11.7 & -1.4 \\ 0.8 & 3.2 \\ 19.7 & -0.5 \\ 30.1 & 0.4 \\ 15.4 & -0.4 \\ -13 & 3.7 \end{bmatrix} = \begin{bmatrix} 3488 & -121 \\ -121 & 48 \end{bmatrix}$$

$$SSCP_{RSV} = \begin{bmatrix} 20.6 & -8.3 & 10.6 & 2 & -8.7 & -16 \\ -2.2 & -1 & -1.7 & 1.2 & 1.4 & 2.5 \end{bmatrix} \times \begin{bmatrix} 20.6 & -2.2 \\ -8.3 & -1 \\ 10.6 & -1.4 \\ 2 & 1.2 \\ -8.7 & 1.4 \\ -16 & 2.5 \end{bmatrix} = \begin{bmatrix} 941.3 & -104.8 \\ -104.8 & 18.4 \end{bmatrix}$$

$$SSCP_{pylori} = \begin{bmatrix} 0.9 & -10 & 8.9 & 19.3 & 4.6 & -23.8 \\ -2.2 & 2.4 & -1.3 & -0.4 & -1.2 & 2.9 \end{bmatrix} \times \begin{bmatrix} 0.9 & -2.2 \\ -10 & 2.4 \\ 8.9 & -1.3 \\ 19.3 & -0.4 \\ 4.6 & -1.2 \\ -23.8 & 2.9 \end{bmatrix} = \begin{bmatrix} 1140.1 & -119.8 \\ -119.8 & 22.3 \end{bmatrix}$$

$$SSCP_{within} = \begin{bmatrix} 941.3 & -104.8 \\ -104.8 & 18 \end{bmatrix} + \begin{bmatrix} 1140.1 & -119.8 \\ -119.8 & 22.3 \end{bmatrix} = \begin{bmatrix} 2081 & -225 \\ -225 & 41 \end{bmatrix}$$