第 15 章

多元線性迴歸

定義 ***15-1-1***：多元線性迴歸統計分析模式

$$y_i=\beta_0+\beta_1 x_{i1}+\beta_2 x_{i2}+\cdots+\beta_p x_{ip}+\varepsilon_i \qquad i=1,\cdots,n$$

其中(1) y_i 表示當自變項 $X_1=x_{i1},\cdots,X_p=x_{ip}$ 時的因變項 Y 的反應值。

(2) $\beta_0+\beta_1 x_{i1}+\beta_2 x_{i2}+\cdots+\beta_p x_{ip}$ 表示當自變項 $X_1=x_{i1},\cdots,X_p=x_{ip}$ 時因變項 y_i 的期望值（平均反應）。

(3) ε_i 表示當自變項 $X_1=x_{i1},\cdots,X_p=x_{ip}$ 時，因變項 y_i 的誤差值（或個別差異）。

(4) $\varepsilon_i\sim N(0,\sigma^2_{Y|X_1,\cdots,X_p})$ 表示 ε_i 具有共同的變異數，並假設所有 ε_i 間具相互獨立性。

定義 ***15-2-1***：多元迴歸母群體迴歸式（ population regression equation）

$$E(Y\,|\,X_1=x_{i1},\cdots,X_p=x_{ip})=\beta_0+\beta_1 x_{i1}+\cdots+\beta_p x_{ip}$$

表 15-1-1 多元迴歸觀測值

Y	X_1	X_2	$\cdots$	X_p
y_1	x_{11}	x_{12}	$\cdots$	x_{1p}
y_2	x_{21}	x_{22}	$\cdots$	x_{2p}
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$
y_i	x_{i1}	x_{i2}	$\cdots$	x_{ip}
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$
y_n	x_{n1}	x_{n2}	$\cdots$	x_{np}

定義 *15-2-2*：多元迴歸樣本迴歸式（ sample regression equation）

$$\hat{y}_i = b_0 + b_1 x_{i1} + \cdots + b_p x_{ip}$$

其中(1)$b_0, b_1, \cdots, b_P$ 為 $\beta_0, \beta_1, \cdots, \beta_P$ 的估計值

(2) $\hat{y}_i$ 為 $X_1 = x_{i1}, \cdots, X_P = x_{ip}$ 時母群體迴歸式上 $E\left(Y \middle| X_1 = x_{i1}, \cdots, X_P = x_{ip}\right)$ 的估計值，或以擬合值（fitted value）稱之。

定義 *15-2-3*：殘差（ residual）及殘差平方和（ SSE, residual sum of square）

$$\begin{aligned} e_i &= y_i - \hat{y}_i \\ &= y_i - (b_0 + b_1 x_{i1} + \cdots + b_p x_{ip}) \end{aligned}$$

$$\begin{aligned} \text{SSE} &= \Sigma e_i^2 \\ &= \Sigma (y_i - \hat{y}_i)^2 \\ &= \Sigma (y_i - b_0 - b_1 x_{i1} - \cdots - b_p x_{ip})^2 \end{aligned}$$

定理 *15-2-1*：多元迴歸最小平方迴歸式（ least-square regression equation ）

最小平方法所求得的樣本迴歸式係數以矩陣表示為

$$\boldsymbol{b} = (X'X)^{-1}(X'Y)$$

其中，$\boldsymbol{b}$為（$p+1$）×1 係數矩陣，X為（$p+1$）×（$p+1$）自變項矩陣，Y為（$p+1$）×1 因變項矩陣。

定理 *15-2-2*：*p = 2* 時之最小平方迴歸式

$$\hat{y}_i = b_0 + b_1 x_{i1} + b_2 x_{i2}$$

$$\text{其中}(1) b_1 = \frac{\begin{vmatrix} SS_{1Y} & SS_{12} \\ SS_{2Y} & SS_{22} \end{vmatrix}}{\begin{vmatrix} SS_{11} & SS_{12} \\ SS_{21} & SS_{22} \end{vmatrix}}$$

$$(2) b_2 = \frac{\begin{vmatrix} SS_{11} & SS_{1Y} \\ SS_{21} & SS_{2Y} \end{vmatrix}}{\begin{vmatrix} SS_{11} & SS_{12} \\ SS_{21} & SS_{22} \end{vmatrix}}$$

$$(3) b_0 = \overline{Y} - b_1\overline{X}_1 + b_2\overline{X}_2$$

$$\text{其中} SS_{11} = \Sigma(x_{i1} - \overline{X}_1)^2$$

$$SS_{12} = SS_{21} = \Sigma(x_{i1} - \overline{X}_1)(x_{i2} - \overline{X}_2)$$

$$SS_{22} = \Sigma(x_{i2} - \overline{X}_2)^2$$

$$SS_{1Y} = \Sigma(x_{i1} - \overline{X}_1)(y_i - \overline{Y})$$

$$SS_{2Y} = \Sigma(x_{i2} - \overline{X}_2)(y_i - \overline{Y})$$

定義 *15-2-5*：多元迴歸最小殘差平方和及其自由度

我們定義多元迴歸最小殘差平方和為最小平方迴歸式下之殘差平方和，它的自由度為樣本數減去迴歸式中所估計之參數數目，$n-(p+1)$。

定義 15-2-6：$\sigma^2_{Y|X_1,X_2,\cdots,X_p}$之估計量

$$\hat{\sigma}^2_{Y|X_1,X_2,\cdots,X_p}=S^2_{Y|X_1,X_2,\cdots,X_p}=\frac{\text{最小平方迴歸式殘差平方和}}{n-(p+1)}$$

定理 15-3-1：多元迴歸最小平方迴歸式的統計性質

$$\hat{y}_i=b_0+b_1x_{i1}+\cdots+b_px_{ip}$$

(1)迴歸式通過資料中心$(\bar{X}_1,\cdots,\bar{X}_p,\bar{Y})$

(2)$\Sigma e_i=0$（殘差和等於 0）

(3)$\Sigma e_ix_{ik}=0$，$k=1,\cdots,p$（殘差與每個自變項之乘積和等於 0）

(4)$\Sigma e_i\hat{y}_i=0$（殘差與擬合值乘積和等於 0）

定理 15-3-2：迴歸係數 $b_0, b_1, \cdots, b_p$的抽樣分配

迴歸係數 $b_0, b_1, \cdots, b_p$ 的變異共變異估計矩陣（estimated variance-covariance matrix）為$\hat{Cov}(\text{b})$，且

$$\hat{Cov}(\text{b})=S^2_{Y|X_1,\cdots,X_p}(X'X)^{-1}$$

$$=S^2_{Y|X_1,\cdots,X_p}\left(\begin{bmatrix}1 & 1 & \cdots & 1\\ x_{11} & x_{21} & \cdots & x_{n1}\\ x_{12} & x_{22} & \cdots & x_{n2}\\ \vdots & \vdots & & \vdots\\ x_{1p} & x_{2p} & \cdots & x_{np}\end{bmatrix}\begin{bmatrix}1 & x_{11} & x_{12} & \cdots & x_{1p}\\ 1 & x_{21} & x_{22} & \cdots & x_{2p}\\ \vdots & \vdots & \vdots & & \vdots\\ 1 & x_{n1} & x_{n2} & \cdots & x_{np}\end{bmatrix}\right)^{-1}$$

所以$b_0, b_1, \cdots, b_p$ 的抽樣分配為

$b_i\sim N(\beta_i, \hat{Cov}(\text{b})_{i+1,i+1})$　　$i=0, 1, \cdots, p$

其中$\hat{Cov}(\text{b})_{i+1,i+1}$為$\hat{Cov}(\text{b})$主對角線上第（$i+1$）個元素。

定義 ***15-3-1***：多元迴歸式 SST, SSR 及 SSE

$$(1)\mathrm{SST}=\Sigma y_i^2-\frac{(\Sigma y_i)^2}{n}$$，它的自由度為 $n-1$

$$(2)\mathrm{SSR}=(b_0\Sigma y_i+b_1\Sigma x_{i1}y_i+\cdots+b_p\Sigma x_{ip}y_i)-\frac{(\Sigma y_i)^2}{n}$$它的自由度為 p

$$(3)\mathrm{SSE}=\Sigma y_i^2-(b_0\Sigma y_i+b_1\Sigma x_{i1}y_i+\cdots+b_p\Sigma x_{ip}y_i)$$它的自由度為 $n-(p+1)$

表 15-3-1　多元迴歸 ANOVA 表

來源	平方和	自由度	均方和	F 值
迴歸	SSR	p	$\mathrm{MSR}=\frac{\mathrm{SSR}}{p}$	$\frac{\mathrm{MSR}}{\mathrm{MSE}}$
誤差	SSE	$n-(p+1)$	$\mathrm{MSE}=\frac{\mathrm{SSE}}{n-p-1}$	
總計	SST	$n-1$		

定義 ***15-4-1***：多元迴歸判定係數（ coefficient of multiple determination ）

$$R^2=1-\frac{\mathrm{SSE}}{\mathrm{SST}}=\frac{\mathrm{SSR}}{\mathrm{SST}}$$

定義 ***15-4-2***：多元迴歸校正後的判定係數（ Adjusted R^2 ）

$$Adj\ R^2=1-\frac{\mathrm{SSE}/n-p-1}{\mathrm{SST}/n-1}=1-\left(\frac{n-1}{n-p-1}\right)\frac{\mathrm{SSE}}{\mathrm{SST}}$$

此處所謂的校正是以 SSE 及 SST 的自由度對 R^2 進行校正。

整條迴歸式檢定（ testing the significance of multiple regression equation ）

$$H_0：\beta_1=\beta_2=\cdots=\beta_P=0$$

H_1：至少有一個或一個以上 $\beta_j \neq 0$

顯著水準為 α 之拒絕域 R 為

$$R=\left\{\frac{\text{MSR}}{\text{MSE}}\middle|\frac{\text{MSR}}{\text{MSE}}>F_{\alpha}(p,n-p-1)\right\}$$

迴歸式個別係數 β_j 的檢定

H_0：$\beta_j=0 \qquad j=0,1,\cdots,p$

H_1：$\beta_j \neq 0$

顯著水準為 α 時的拒絕域 R 為

$$R=\left\{b_j\middle|\left|\frac{b_j}{S(b_j)}\right|>t_{\frac{\alpha}{2}}(n-p-1)\right\}$$

其中 b_j 為 β_j 的估計值，S（b_j）為 b_j 抽樣分配標準差估計值。

迴歸式個別係數 β_j 的信賴區間

β_j 的 100（1−α）%信賴區間為

$$\left[b_j-t_{\frac{\alpha}{2}}(n-p-1)S(b_j),\quad b_j+t_{\frac{\alpha}{2}}(n-p-1)S(b_j)\right]$$

15.1 多元迴歸（X_1, X_2, Y）樣本資料如表 E-15-1-1

表 E-15-1-1

X_1	X_2	Y
0.2	5	43
0.1	2	35
0.4	6	47
0.5	3	32
0.3	4	38

(1)計算最小平方多元迴歸式。

(2)計算最小平方多元迴歸式的殘差平方和。

(3)估計迴歸模式中誤差項ε_i的變異數及標準差。

(4)編製 ANOVA 表。

(5)計算迴歸線的判定係數及校正後的判定係數。

(6)驗証$\Sigma e_i = 0$，$\Sigma e_i x_{ik} = 0$及$\Sigma e_i \hat{y}_i = 0$。

(7)估計b_0的變異數及標準差。

(8)估計b_1的變異數及標準差。

(9)估計b_2的變異數及標準差。

(10)顯著水準$\alpha = 0.05$下檢定，$H_0 : \beta_1 = 0$，$H_1 : \beta_1 \neq 0$。

(11)顯著水準$\alpha = 0.05$下檢定，$H_0 : \beta_2 = 0$，$H_1 : \beta_2 \neq 0$。

(12)顯著水準$\alpha = 0.05$下檢定，$H_0 : \beta_0 = 0$，$H_1 : \beta_0 \neq 0$。

(13)解釋(10),(11),(12)檢定的結果所顯示的意義。

(14)計算β_1的 95%信賴區間。

(15)計算β_2的 95%信賴區間。

解：

表 E-15-1-2

x_1	x_2	y	x_1^2	x_2^2	x_1x_2	x_1y	x_2y	y^2
0.2	5	43	0.04	25	1	8.6	215	1849
0.1	2	35	0.01	4	0.2	3.5	70	1225
0.4	6	47	0.16	36	2.4	18.8	282	2209
0.5	3	32	0.25	9	1.5	16	96	1024
0.3	4	38	0.09	16	1.2	11.4	152	1444
1.5	20	195	0.55	90	6.3	58.3	815	7751
0.3	4	39	0.11	18	1.26	11.66	163	1550.2

(1) $SS_{11}=\Sigma x_{i1}^2-n\overline{X}_1^2=0.55-5(0.3)^2=0.1$

$SS_{22}=\Sigma x_{i2}^2-n\overline{X}_2^2=90-5(4)^2=10$

$SS_{12}=\Sigma x_{i1}x_{i2}-n\overline{X}_1\overline{X}_2=6.3-5(0.3)(0.4)=0.3$

$SS_{1Y}=\Sigma x_{i1}y_i-n\overline{X}_1\overline{Y}=58.3-5(0.3)(39)=-0.2$

$SS_{2Y}=\Sigma x_{i2}y_i-n\overline{X}_2\overline{Y}=815-5(4)(39)=35$

$$b_1=\frac{\begin{vmatrix} SS_{1Y} & SS_{12} \\ SS_{2Y} & SS_{22} \end{vmatrix}}{\begin{vmatrix} SS_{11} & SS_{12} \\ SS_{21} & SS_{22} \end{vmatrix}}=\frac{\begin{vmatrix} -0.2 & 0.3 \\ 35 & 10 \end{vmatrix}}{\begin{vmatrix} 0.1 & 0.3 \\ 0.3 & 10 \end{vmatrix}}=-13.736$$

$$b_2=\frac{\begin{vmatrix} SS_{11} & SS_{1Y} \\ SS_{21} & SS_{2Y} \end{vmatrix}}{\begin{vmatrix} SS_{11} & SS_{12} \\ SS_{21} & SS_{22} \end{vmatrix}}=\frac{\begin{vmatrix} 0.1 & -0.2 \\ 0.3 & 35 \end{vmatrix}}{\begin{vmatrix} 0.1 & 0.3 \\ 0.3 & 10 \end{vmatrix}}=3.912$$

$b_0=\overline{Y}-b_1\overline{X}_1-b_2\overline{X}_2=39-(-13.736)(0.3)-(3.912)(4)=27.473$

所以，最小平方多元迴歸式為$\hat{y}_i=27.473-13.736\,x_{i1}+3.912\,x_{i2}$

⑵

表 E-15-1-3

x_1	x_2	y	$\hat{y}_i$	$y_i-\hat{y}_i$	$(y_i-\hat{y}_i)^2$
0.2	5	43	44.286	-1.286	1.653
0.1	2	35	33.923	1.077	1.160
0.4	6	47	45.451	1.549	2.401
0.5	3	32	32.341	-0.341	0.116
0.3	4	38	39	-1	1
					6.330

由試算表 E-15-1-3 得知 $\text{SSE}=6.33$

⑶$\hat{\sigma}^2_{Y1X_1,X_2}=S^2_{Y1X_1,X_2}=\text{MSE}=\dfrac{\text{SSE}}{n-(2+1)}=\dfrac{6.33}{2}=3.165$

⑷$\text{SST}=\Sigma y_i^2-n\overline{Y}^2=7751-5(39)^2=146$

$\text{SSR}=\text{SST}-\text{SSE}=146-6.33=139.67$

所以本題的 ANOVA 表如表 E-15-1-4 所示。

表 E-15-1-4

來源	平方和	自由度	均方和	F 值
迴歸	139.67	2	69.84	22.1
誤差	6.33	2	3.165	
總計	146	4		

⑸$R^2=\dfrac{\text{SSR}}{\text{SST}}=1-\dfrac{\text{SSE}}{\text{SST}}=1-\dfrac{6.33}{146}=0.9566$

$Adj\,R^2=1-(\dfrac{n-1}{n-p-1})\dfrac{\text{SSE}}{\text{SST}}=1-\left(\dfrac{4}{2}\right)\dfrac{6.33}{1.46}=0.9133$

(6)

表 E-15-1-5

x_{i1}	x_{i2}	y_i	$\hat{y}_i$	e_i	$e_i x_{i1}$	$e_i x_{i2}$	$e_i \hat{y}_i$
0.2	5	43	44.286	-1.286	-0.2572	-6.43	-56.952
0.1	2	35	33.923	1.077	0.1077	2.154	36.535
0.4	6	47	45.451	1.549	0.6196	9.294	70.404
0.5	3	32	32.341	-0.341	-0.1705	-0.942	-11.028
0.3	4	38	39	-1	-0.3	-4	-39
				0	0	0	0

由表E-15-1-5的第5欄驗證 $\Sigma e_i = 0$，由表E-15-1-5的第6欄驗證 $\Sigma e_i x_{i1} = 0$，由表E-15-1-5的第7欄驗證 $\Sigma e_i x_{i2} = 0$，由表E-15-1-5的第8欄驗證 $\Sigma e_i \hat{y}_i = 0$

(7)
$$\boldsymbol{X'X} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0.2 & 0.1 & 0.4 & 0.5 & 0.3 \\ 5 & 2 & 6 & 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 0.2 & 5 \\ 1 & 0.1 & 2 \\ 1 & 0.4 & 6 \\ 1 & 0.5 & 3 \\ 1 & 0.3 & 4 \end{bmatrix}$$

$$= \begin{bmatrix} 5 & 1.5 & 20 \\ 1.5 & 0.55 & 6.3 \\ 20 & 6.3 & 90 \end{bmatrix}$$

$$(\boldsymbol{X'X})^{-1} = \begin{bmatrix} 2.156 & -1.978 & -0.341 \\ -1.978 & 10.989 & -0.33 \\ -0.341 & -0.33 & 0.11 \end{bmatrix}$$

所以 X_1, X_2, Y 的變異共變異矩陣的估計矩陣為

$$S^2_{Y1X_1,X_2} \begin{bmatrix} 2.156 & -1.978 & -0.341 \\ -1.978 & 10.989 & -0.33 \\ -0.341 & -0.33 & 0.11 \end{bmatrix} = 3.165 \begin{bmatrix} 2.156 & -1.978 & -0.341 \\ -1.978 & 10.989 & -0.33 \\ -0.341 & -0.33 & 0.11 \end{bmatrix}$$

$$=\begin{bmatrix} 6.824 & -6.26 & -1.079 \\ -6.26 & 34.78 & -1.045 \\ -1.079 & -1.045 & 0.348 \end{bmatrix}$$

b_0的變異數為$S^2_{Y1X_1,X_2}(\boldsymbol{X'X})^{-1}$主對角線上的第一個元素數值，所以$b_0$的變異數及標準差估計值$S^2_{b_0}$及$S_{b_0}$分別為

$$S^2_{b_0}=6.824$$
$$S_{b_0}=\sqrt{6.824}=2.612$$

(8)b_1的變異數為$S^2_{Y1X_1,X_2}(\boldsymbol{X'X})^{-1}$主對角線上第二個元素數值，所以$b_1$的變異數及標準差估計值$S^2_{b_1}$及$S_{b_1}$分別為

$$S^2_{b_1}=34.78$$
$$S_{b_1}=\sqrt{34.78}=5.897$$

(9)b_2的變異數為$S^2_{Y1X_1,X_2}(\boldsymbol{X'X})^{-1}$主對角線上第三個元素數值，所以$b_2$的變異數及標準差估計值$S^2_{b_2}$及$S_{b_2}$分別為

$$S^2_{b_2}=0.348$$
$$S_{b_2}=\sqrt{0.348}=0.59$$

(10)
$$H_0：\beta_1=0$$
$$H_1：\beta_1\neq 0$$

顯著水準$\alpha=0.05$時，$t_{0.025}(2)=4.303$，所以檢定的拒絕域為

$$R=\left\{\frac{b_1}{S_{b_1}}\middle|\left|\frac{b_1}{S_{b_1}}\right|>4.303\right\}$$

根據本題樣本資料計算檢定值

$$\left|\frac{b_1}{S_{b_1}}\right|=\left|\frac{-13.736}{5.897}\right|=2.3294<4.303$$

所以檢定結果為不拒絕H_0，這表示迴歸係數斜率β_1為 0。

⑾
$$H_0: \beta_2 = 0$$
$$H_1: \beta_2 \neq 0$$

顯著水準$\alpha = 0.05$，$t_{0.025}(2) = 4.303$，所以檢定的拒絕域為

$$R = \left\{ \frac{b_2}{S_{b_2}} \middle| \left| \frac{b_2}{S_{b_2}} \right| > 4.303 \right\}$$

根據本題樣本資料計算檢定值

$$\left| \frac{b_2}{S_{b_2}} \right| = \left| \frac{3.912}{0.59} \right| = 6.631 > 4.303$$

所以檢定結果為拒絕H_0，這表示迴歸係數斜率β_2顯著不等於 0。

⑿
$$H_0: \beta_0 = 0$$
$$H_1: \beta_0 \neq 0$$

顯著水準$\alpha = 0.05$，$t_{0.025}(2) = 4.303$，所以檢定的拒絕域為

$$R = \left\{ \frac{b_0}{S_{b_0}} \middle| \left| \frac{b_0}{S_{b_0}} \right| > 4.303 \right\}$$

根據本題樣本資料計算檢定值

$$\left| \frac{b_0}{S_{b_0}} \right| = \left| \frac{27.473}{2.612} \right| = 10.518$$

所以檢定結果為拒絕H_0，這表示迴歸係數β_0顯著不等於 0。

⒀由 ANOVA 表中的$F = 22.1 > F_{0.05}(2,2) = 19$，這表示在顯著水準$\alpha = 0.05$下這條多元迴歸線整體來說是有意義的。但是當我們進一步逐一檢定迴歸係數$\beta_0, \beta_1, \beta_2$時，卻發現$\beta_1 \neq 0$並不顯著。對於這種現象，我們可以解釋為，整條迴歸線是有意義的，我們也可以用它來作預測。但當我們進一步解釋迴歸係數的意義時便要注意了，對於β_2，我們可以說，當其他條件固定時，$b_2 = 3.912$所顯示的意義為每增加一單位X_2對Y的貢獻為 3.912。但對於β_1，我們便不可以說，$b_1 = -13.736$所代表的意義為，每增加一單位X_1時對Y的

貢獻為−13.736。

⒁ $1-\alpha=0.95$，$t_{0.025}(2)=4.303$，所以β_1的 95%信賴區間為

$$\left[b_1-4.303\ S_{b_1}, b_1+4.303\ S_{b_1}\right]$$

根據本題樣本資料計算得到的區間為

$$\begin{aligned}&\left[-13.736-(4.303)(5.897), -13.736+(4.303)(5.897)\right]\\=&\left[-39.111, 11.639\right]\end{aligned}$$

⒂ β_2 的 95%信賴區間為

$$\begin{aligned}&\left[b_2-4.303\ S_{b_2}, b_2+4.303\ S_{b_2}\right]\\=&\left[3.912-(4.303)(0.59), 3.912+(4.303)(0.59)\right]\\=&\left[1.373, 6.451\right]\end{aligned}$$

15.2 多元迴歸（X_1, X_2, Y）樣本資料如表 E-15-2-1

表 E-15-2-1

X_1	X_2	Y
10	5	41
8	2	35
11	6	45
7	3	32
9	4	37

⑴計算最小平方多元迴歸式。

⑵計算最小平方多元迴歸式的殘差平方和。

⑶估計迴歸模式中誤差項ε_i的變異數及標準差。

⑷編製 ANOVA 表。

(5)計算迴歸線的判定係數及校正後的判定係數。

(6)驗証 $\Sigma e_i=0$，$\Sigma e_i x_{ik}=0$ 及 $\Sigma e_i \hat{y}_i=0$。

(7)估計 b_0 的變異數及標準差。

(8)估計 b_1 的變異數及標準差。

(9)估計 b_2 的變異數及標準差。

(10)顯著水準 $\alpha=0.05$ 下檢定，H_0：$\beta_1=0$，H_1：$\beta_1 \neq 0$。

(11)顯著水準 $\alpha=0.05$ 下檢定，H_0：$\beta_2=0$，H_1：$\beta_2 \neq 0$。

(12)顯著水準 $\alpha=0.05$ 下檢定，H_0：$\beta_0=0$，H_1：$\beta_0 \neq 0$。

(13)計算 β_1 的 95%信賴區間。

(14)計算 β_2 的 95%信賴區間。

解：

表 E-15-2-2

x_1	x_2	y	x_1^2	x_2^2	x_1x_2	x_1y	x_2y	y^2
10	5	41	100	25	50	410	205	1681
8	2	35	64	4	16	280	70	1225
11	6	45	121	36	66	495	270	2025
7	3	32	49	9	21	224	96	1024
9	4	37	81	16	36	333	148	1369
45	20	190	415	90	189	1742	789	7324
9	4	38	83	18	37.8	348.4	157.8	1464.8

(1)$SS_{11}=\Sigma x_{i1}^2-n\overline{X}_1^2=415-5(9)^2=10$

$SS_{22}=\Sigma x_{i2}^2-n\overline{X}_2^2=90-5(4)^2=10$

$SS_{12}=\Sigma x_{i1}x_{i2}-n\overline{X}_1\overline{X}_2=189-5(9)(4)=9$

$SS_{1Y}=\Sigma x_{i1}y_i-n\overline{X}_1\overline{Y}=1742-5(9)(38)=32$

$SS_{2Y}=\Sigma x_{i2}y_i-n\overline{X}_2\overline{Y}=789-5(4)(38)=29$

$$b_1=\frac{\begin{vmatrix}SS_{1Y} & SS_{12}\\ SS_{2Y} & SS_{22}\end{vmatrix}}{\begin{vmatrix}SS_{11} & SS_{12}\\ SS_{21} & SS_{22}\end{vmatrix}}=\frac{\begin{vmatrix}32 & 9\\ 29 & 10\end{vmatrix}}{\begin{vmatrix}10 & 9\\ 9 & 10\end{vmatrix}}=3.105$$

$$b_2=\frac{\begin{vmatrix} SS_{11} & SS_{1Y} \\ SS_{21} & SS_{2Y} \end{vmatrix}}{\begin{vmatrix} SS_{11} & SS_{12} \\ SS_{21} & SS_{22} \end{vmatrix}}=\frac{\begin{vmatrix} 10 & 32 \\ 9 & 29 \end{vmatrix}}{\begin{vmatrix} 10 & 9 \\ 9 & 10 \end{vmatrix}}=0.105$$

$$b_0=\overline{Y}-b_1\overline{X}_1-b_2\overline{X}_2=38-(3.105)(9)-(0.105)(4)=9.632$$

(2)

表 E-15-2-3

x_1	x_2	y	$\hat{y}_i$	$y_i-\hat{y}_i$	$(y_i-\hat{y}_i)^2$
10	5	41	41.211	−0.211	0.044
8	2	35	34.684	0.316	0.100
11	6	45	44.421	0.579	0.335
7	3	32	31.684	0.316	0.100
9	4	37	38	−1	1
					1.579

由試算表 E-15-2-3 得知 SSE $=1.579$

(3) $\hat{\sigma}^2_{Y1X_1,X_2}=S^2_{Y1X_1,X_2}=\text{MSE}=\dfrac{\text{SSE}}{n-(2+1)}=\dfrac{1.579}{2}=0.789$

(4) $\text{SST}=\Sigma y_i^2-n\overline{Y}^2=7324-5(38^2)=104$

$\text{SSR}=\text{SST}-\text{SSE}=104-1.579=102.421$

所以本題的 ANOVA 表如表 E-15-2-4 所示。

表 E-15-2-4

來源	平方和	自由度	均方和	F 值
迴歸	102.421	2	51.211	64.867
誤差	1.579	2	0.789	
總計	104	4		

(5)$R^2=\frac{\text{SSR}}{\text{SST}}=1-\frac{\text{SSE}}{\text{SST}}=1-\frac{1.579}{104}=0.9848$

$AdjR^2=1-(\frac{n-1}{n-p-1})\frac{\text{SSE}}{\text{SST}}=1-\left(\frac{4}{2}\right)\frac{1.579}{104}=0.97$

(6)

表 E-15-2-5

x_{i1}	x_{i2}	y_i	$\hat{y}_i$	e_i	$e_i x_{i1}$	$e_i x_{i2}$	$e_i \hat{y}_i$
10	5	41	41.211	−0.211	−2.11	−1.055	−8.696
8	2	35	34.684	0.316	2.528	0.632	10.96
11	6	45	44.421	0.579	6.369	3.474	25.719
7	3	32	31.684	0.316	2.212	0.948	10.012
9	4	37	38	−1	−9	−4	−38
				0	0	0	0

由表 E-15-2-5 第 5 欄驗證 $\Sigma e_i=0$，由表 E-15-2-5 第 6 欄驗證 $\Sigma e_i x_{i1}=0$，由表 E-15-2-5 第 7 欄驗證 $\Sigma e_i x_{i2}=0$，由表 E-15-2-5 第 8 欄驗證 $\Sigma e_i \hat{y}_i=0$

(7) $$\boldsymbol{X'X}=\begin{bmatrix}1 & 1 & 1 & 1 & 1\\ 10 & 8 & 11 & 7 & 9\\ 5 & 2 & 6 & 3 & 4\end{bmatrix}\begin{bmatrix}1 & 10 & 5\\ 1 & 8 & 2\\ 1 & 11 & 6\\ 1 & 7 & 3\\ 1 & 9 & 4\end{bmatrix}$$

$$=\begin{bmatrix}5 & 45 & 20\\ 45 & 415 & 189\\ 20 & 189 & 90\end{bmatrix}$$

$$(\boldsymbol{X'X})^{-1}=\begin{bmatrix}17.147 & -2.842 & 2.158\\ -2.842 & 0.526 & -0.474\\ 2.158 & -0.474 & 0.526\end{bmatrix}$$

所以X_1 , X_2 , Y的變異共變異矩陣的估計矩陣為

1H1515.tpf-17 5/9/2002 14:26:13

$$S^2_{Y1X_1,X_2}\begin{bmatrix}17.147 & -2.842 & 2.158\\ -2.842 & 0.526 & -0.474\\ 2.158 & -0.474 & 0.526\end{bmatrix}$$

$$=0.789\begin{bmatrix}17.147 & -2.842 & 2.158\\ -2.842 & 0.526 & -0.474\\ 2.158 & -0.474 & 0.526\end{bmatrix}$$

$$=\begin{bmatrix}13.529 & -2.242 & 1.703\\ -2.242 & 0.415 & 0.374\\ 1.703 & 0.374 & 0.415\end{bmatrix}$$

b_0 的變異數為 $S^2_{Y1X_1,X_2}(\boldsymbol{X'X})^{-1}$ 主對角線上的第一個元素數值，所以 b_0 的變異數及標準差估計值 $S^2_{b_0}$ 及 S_{b_0} 分別為

$$S^2_{b_0}=13.529$$
$$S_{b_0}=\sqrt{13.529}=3.679$$

(8) b_1 的變異數為 $S^2_{Y1X_1,X_2}(\boldsymbol{X'X})^{-1}$ 主對角線上第二個元素數值，所以 b_1 的變異數 $(S^2_{b_1})$ 及標準差 (S_{b_1}) 估計值分別為

$$S^2_{b_1}=0.415$$
$$S_{b_1}=\sqrt{0.415}=0.645$$

(9) b_2 的變異數為 $S^2_{Y1X_1,X_2}(\boldsymbol{X'X})^{-1}$ 主對角線上第三個元素數值，所以 b_3 的變異數及標準差估計值 $S^2_{b_2}$ 及 $S^2_{b_1}$ 分別為

$$S^2_{b_2}=0.415$$
$$S_{b_2}=\sqrt{0.415}=0.645$$

(10)
$$H_0：\beta_1=0$$
$$H_1：\beta_1\neq 0$$

顯著水準 $\alpha=0.05$ 時，$t_{0.025}(2)=4.303$，所以檢定的拒絕域為

$$R=\left\{\frac{b_1}{S_{b_1}}\middle|\left|\frac{b_1}{S_{b_1}}\right|>4.303\right\}$$

根據本題樣本資料計算檢定值

$$\left|\frac{b_1}{S_{b_1}}\right|=\left|\frac{3.105}{0.645}\right|=4.814>4.303$$

所以檢定結果為拒絕 H_0，這表示迴歸係數 β_1 顯著不等於 0。

(11)
$$H_0：\beta_2=0$$
$$H_1：\beta_2\neq 0$$

顯著水準 $\alpha=0.05$，$t_{0.025}(2)=4.303$，所以檢定的拒絕域為

$$R=\left\{\frac{b_2}{S_{b_2}}\middle|\left|\frac{b_2}{S_{b_2}}\right|>4.303\right\}$$

根據本題樣本資料計算檢定值

$$\left|\frac{b_2}{S_{b_2}}\right|=\left|\frac{0.105}{0.645}\right|=0.163<4.303$$

所以檢定結果為不拒絕 H_0，這表示迴歸係數 β_2 為 0。

(12)
$$H_0：\beta_0=0$$
$$H_1：\beta_0\neq 0$$

顯著水準 $\alpha=0.05$，$t_{0.025}(2)=4.303$，所以檢定的拒絕域為

$$R=\left\{\frac{b_0}{S_{b_0}}\middle|\left|\frac{b_0}{S_{b_0}}\right|>4.303\right\}$$

根據本題樣本資料計算檢定值

$$\left|\frac{b_0}{S_{b_0}}\right|=\left|\frac{9.632}{3.679}\right|=1.803<4.303$$

所以檢定結果為不拒絕 H_0，這表示迴歸係數 β_0 為 0。

⒀ $1-\alpha=0.95$，$t_{0.025}(2)=4.303$。所以 β_1 的 95%信賴區間為

$$\left[b_1-4.303\ S_{b_1}\,,\,b_1+4.303\ S_{b_1}\right]$$

根據本題樣本資料計算得到的區間為

$$\begin{aligned}&\left[3.105-(4.303)(0.645)\,,\,3.105+(4.303)(0.645)\right]\\=&\left[0.33\,,\,5.88\right]\end{aligned}$$

⒁ β_2 的 95%信賴區間為

$$\begin{aligned}&\left[0.105-(4.303)(0.645)\,,\,0.105+(4.303)(0.645)\right]\\=&\left[-2.67\,,\,2.88\right]\end{aligned}$$