【統計】

《迴歸分析》

試題評析

今年高考迴歸分析考題偏難，往年少見的模型選取判斷準則，如 Mallows' C_p 法、Forward Selection Method、AIC、BIC 等都出現了；此外，迴歸係數的加權最小平方估計量、殘差分析原理等往年考題亦不最常見，考生不易拿取高分。試題內容涵蓋了迴歸模型係數的估計、檢定、ANOVA、殘差分析原理、無截距項的最小平方估計量、加權最小平方估計量、各種模型選取判斷準則。一般考生有可能拿不到 60 分，程度好的考生在考題的作答上恐怕也要花上不少時間。

一、迴歸（regression）一詞是由英國Galton爵士所提出，更早他所用的字為”相反”（reversion）。在當時，他調查一組父母的加權平均身高$X_1, X_2, \cdots X_n$及其子女的身高$Y_1, Y_2, \cdots Y_n$時，其所得到的估計值（fitted values）類似像

$$\hat{Y}=\bar{Y}+\frac{2}{3}(X_i-\bar{X}),\ i=1,2,...,n$$

其中 $\bar{Y}=\frac{1}{n}\sum_{i=1}^{n}Y_i$ 以及 $\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$ 。試詳加說明為何根據上述關係式，我們可得到父母身高若都很高，其小孩身高雖會比一般人高，但不會像其父母那麼高的結論；亦即很高父母的小孩之身高有退化的趨勢。（10分）

答：我們知道簡單迴歸模型是假設 $y=\beta_0+\beta_1 x+\varepsilon$ ，亦即 y 和 x 有線性關係，基於此假設下可得 y 的平均數的估計式爲 $\hat{y}=\hat{\beta}_0+\hat{\beta}_1 x$ ，其中迴歸係數 β_1,β_0 之 LSE 分別爲：

$$\hat{\beta}_1=\frac{SSXY}{SSX}=\frac{\sum_{i=1}^{n}(X_i-\bar{X})(Y_i-\bar{Y})}{\sum_{i=1}^{n}(X_i-\bar{X})^2}=\frac{\sum_{i=1}^{n}(X_i)(Y_i)-n(\bar{X})(\bar{Y})}{\sum_{i=1}^{n}(X_i)^2-n(\bar{X})^2}$$

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

$$\therefore\ \hat{y}=\hat{\beta}_0+\hat{\beta}_1 x=\bar{Y}-\hat{\beta}_1\bar{X}+\hat{\beta}_1 x=\bar{Y}+\hat{\beta}_1(x-\bar{X})$$

對照上面的迴歸線估計式可知子女身高（Y）與父母加權平均身高（X）之間的迴歸方程式的 $\hat{\beta}_1=\frac{2}{3}$

此值爲迴歸直線斜率的估計值，其含意爲「X 每變動一單位對 Y 的平均數造成的影響」，亦即「父母的加權平均身高每增加一單位，子女的平均身高會增加 2/3 單位」。由 $\hat{\beta}_1$ 的值爲「正值」，我們知道 Y 和 X 之間爲正相關，亦即父母身高若都很高，其小孩身高也會比一般人來得高，反之則否；而由 $\hat{\beta}_1$ 的值小於 1，我們知道父母身高若都很高，其小孩身高也會比一般人來得高，但絕對不會像其父母身高那麼高，因爲父母的

加權平均身高每增加一單位，子女的平均身高只會增加 2/3 單位，比一單位還來得小。

二、對資料（$Y_i, Y_{i1}, \cdots Y_{i2}$），i=1, …, 5，考慮線性迴歸模式

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \varepsilon_i, i = 1, 2, ..., 5$$

其中 $\varepsilon_1, \varepsilon_2, ..., \varepsilon_5$ 為獨立且平均數為0，變異數為 σ_2 之常態分配（normally distributed）隨機變數。一些能會用到的統計量彙整如下：

$\bar{Y} = \frac{1}{5}\sum_{i=1}^{5} Y_i = 20$ ，$\hat{Y} = 20 + 6X_{i1}$ ，$\sum_{i=1}^{5}(\hat{Y}_i - \bar{Y})^2 = 360$ ，$R^2 = 0.9$（R-Square）

迴歸參數	最小平方（least squares）估計值	估計值標準誤差（standard error）
β_0	$\hat{\beta}_0 = 20$	$s.e(\hat{\beta}_0) = 2$
β_1	$\hat{\beta}_1 = 6$	$s.e(\hat{\beta}_1) = 1.4142$
β_2	$\hat{\beta}_3 = 0$	$s.e(\hat{\beta}_2) = 2.23607$

$P(F(3,2)>19.16)=0.05$　　$P(F(2,3)>9.55)=0.05$
$P(T(2)>1.886)=0.1$　　$P(T(3)>1.638)=0.1$
$P(T(2)>2.920)=0.05$　　$P(T(3)>2.353)=0.05$

其中 $F(n_1, n_2)$ 為一分子自由度為 n_1，分母自由度為 n_2 的 F 分配，而 $T(n)$ 為一自由度為 n 的 t 分配。

(一)請完成以下的變異數分析表，並將變異數分析表畫在試卷上。(15分)

變異來源	平方和（Sum of Squares）	自由度（Degree of Freedom）	均方和（Mean Sum of Squares）	F檢定
迴歸	(a)	2	(f)	(h)
殘差	(b)	(d)	(g)	
總和	(c)	(e)		

(二)在顯著水準（level of significance）為0.1下，請利用 t 統計量檢定 $H_0:\beta_1=5$ 。請判斷檢定結果。(5分)

(三)在顯著水準為0.05下，請利用 F 統計量檢定 $H_0:\beta_0=\beta_1=\beta_2=0$ 。請判斷檢定結果。(5分)

答：

(一)若 X 服從常態分配 $N(\mu,\sigma^2)$,則其 p.d.f 為：

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}, x、\mu \in R, \sigma > 0$$

(a) = $SSR = \sum_{i=1}^{5}(\hat{Y}_i - \bar{Y})^2 = 360$

(b) = $SSE = SST - SSR$ = (c)-(a) = 400-360 = 40

(c) = $SST = \frac{SSR}{R^2} = \frac{360}{0.9} = 400$

(d) = n-k = 5-3 = 2

(e) = n –1 = 5-1 = 4

(f) = $MSR = SSR/(\,k\,-1)$ = 360/2 = 180

(g) = $MSE = SSE/(\,n\,-k)$ = 40/2 = 20

(h) = $F = \dfrac{MSR}{MSE}$ = 180/20 = 9

(二) ① $H_0:\beta_1$=5

② $H_1:\beta_1 \neq 5$

③ α = 0.1

④C={t| $|t| > t_{0.05}(2) = 2.92$ }其中 $P(t > t_\alpha) = \alpha$

⑤檢定統計量值

$$t = \frac{\hat{\beta}_1 - 5}{s.e.(\hat{\beta}_1)} = \frac{6-5}{1.4142} \cong 0.7071 \notin C$$

⑥結論：無法拒絕 H_0 ，亦即在顯著水準 α = 0.1 下，無足夠的證據顯示 $\beta_1 \neq 5$。

(三) ① $H_0:\beta_0 = \beta_1 = \beta_2$ =0

② $H_1:\beta_i$ 不全爲 0

③ α = 0.05

④C={F| $F > F_{0.05}(3,2) = 19.16$ }其中 $P(F > F_\alpha) = \alpha$

⑤ $F = \dfrac{[SSE(R) - SSE(F)]/(df_R - df_F)}{SSE(F)/df_F} = \dfrac{[400-40]/[(5-0)-(5-3)]}{40/(5-3)} = 6 \notin C$

⑥結論：無法拒絕 H_0 ，亦即在顯著水準 α = 0.05 下，無足夠的證據顯示 β_i 不全爲 0。

三、對以下的資料

i	1	2	3	4	5	6
Y_i	0	1	0	1	0	1
X_i	1	1	2	2	3	3

(一)考慮配適線性迴歸模式 $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i, i = 1, 2, ..., 6$ 來分析上述資料，其中 $\varepsilon_1, \varepsilon_2, ..., \varepsilon_6$ 為獨立且平均數為0，變異數為 σ^2 之隨機變數。請寫出估計迴歸式（estimated regression equation）及其判定係數（coefficient of determination）R^2。（6分）

(二)考慮配適線性迴歸模式 $Y_i = \beta_1 X_i + \varepsilon_i, i = 1, 2, ..., 6$ 來分析上述資料，其中 $\varepsilon_1, \varepsilon_2, ..., \varepsilon_6$ 為獨立且平均數為0，變異數為 σ^2 之隨機變數。請寫出估計迴歸式以及 R^2。（6分）

(87)(三)上述(一)與(二)中 $\varepsilon_1, \varepsilon_2, ..., \varepsilon_6$ 的平均數假設為0，根據此假設請評論 $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$ 及 $Y_i = \beta_1 X_i + \varepsilon_i$ 何者為較佳之模式？請詳述理由。（6分）

(四)考慮配適線性迴歸模式 $Y_i=\beta_1 X_i+\varepsilon_i, i=1,2,...,6$，其中 $\varepsilon_1,\varepsilon_2,...,\varepsilon_6$ 為獨立且平均數為0，變異數為 $Var(\varepsilon_i)=0.5X_i$ 之隨機變數。試計算迴歸參數 β_1 的加權最小平方（weighted least squares）估計值。（7分）

答：

(一) $$\hat{\beta}_1=\frac{SSXY}{SSX}=\frac{\sum_{i=1}^{n}(X_i-\overline{X})(Y_i-\overline{Y})}{\sum_{i=1}^{n}(X_i-\overline{X})^2}=\frac{\sum_{i=1}^{n}(X_i)(Y_i)-n(\overline{X})(\overline{Y})}{\sum_{i=1}^{n}(X_i)^2-n(\overline{X})^2}=\frac{6-6(2)(0.5)}{28-6(2)^2}=0$$

$$\hat{\beta}_0=\overline{Y}-\hat{\beta}_1\overline{X}=0.5-0\cdot 2=0.5$$

$$\therefore \hat{y}=\hat{\beta}_0+\hat{\beta}_1 x=0.5$$

$$R^2=\frac{SSR}{SST}=\frac{\hat{\beta}_1^2 SSX}{SST}=0$$

(二) $$\hat{\beta}_1=\frac{\sum_{i=1}^{n}X_iY_i}{\sum_{i=1}^{n}(X_i)^2}=\frac{6}{28}=0.2143$$

$$\therefore \hat{y}=\hat{\beta}_1 x=0.2143x$$

$$R^2=\frac{SSR}{SST}==\frac{\sum_{i=1}^{n}(\hat{Y}_i-\overline{Y})^2}{\sum_{i=1}^{n}(Y_i-\overline{Y})^2}=\frac{\sum_{i=1}^{n}(\hat{\beta}_1 X_i-\overline{Y})^2}{\sum_{i=1}^{n}(Y_i-\overline{Y})^2}=\frac{0.2143}{1.5}\cong 0.1429$$

(三)由(一)、(二)的結果知兩個模式皆符合迴歸模式 $E(\varepsilon_i)=0$， $Var(\varepsilon_i)=\sigma^2$（定值）的假設，但第一個模式的 $R^2=0$，相較之下，(二)中的模式會比(一)中的模式來得好得多。

(四)1.∵ $Var(\varepsilon_i)=0.5x_i$ 不滿足傳統迴歸的基本假設，

$$\therefore 令 \varepsilon_i^*=\frac{\varepsilon_i}{\sqrt{X_i}}$$

$$\Rightarrow E(\varepsilon_i^*)=E(\frac{\varepsilon_i}{\sqrt{X_i}})=0$$

$$Var(\varepsilon_i^*)=Var(\frac{\varepsilon_i}{\sqrt{X_i}})=\frac{1}{X_i}Var(\varepsilon_i)=\frac{1}{X_i}\cdot 0.5X_i=0.5$$

$Cov(\varepsilon_i^*, \varepsilon_j^*) = 0$

⇒符合傳統迴歸的基本假設，故可將模式改寫爲

$$Y_i^* = \frac{Y_i}{\sqrt{X_i}} = \frac{\beta_1}{\sqrt{X_i}} X_i + \frac{\varepsilon_i}{\sqrt{X_i}} = \beta_1\sqrt{X_i} + \varepsilon_i^*，$$

由 Gauss-Markov 定理可知β_1的 BLUE 即爲 LSE，

2.令$L = \sum_{i=1}^{n}\left(\frac{Y_i}{\sqrt{X_i}} - \beta_1\sqrt{X_i}\right)^2$

$$\Rightarrow 0 = \frac{\partial L}{\partial \beta_1} = -2\sum_{i=1}^{n}\left(Y_i - \beta_1 X_i\right)$$

⇒β_1之 B.L.U.E 爲$\hat{\beta}_1 = \dfrac{\sum Y_i}{\sum X_i}$，其亦爲$\beta_1$的加權最小平方估計量。

四、對資料（Y_1, X_1），（Y_2, X_2），⋯（Y_n, X_n），考慮以簡單線性迴歸模式

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i,\ i = 1, 2, ..., n$$

來分析資料，其中$\varepsilon_1, \varepsilon_2, ..., \varepsilon_n$為獨立且平均數為0，變異數為$\sigma^2$之隨機變數。一些可能會用到的統計量及關係式整理如下：

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i,\quad \hat{\varepsilon}_i = Y_i - \hat{Y}_i,\quad \sum_{i=1}^{n}\hat{\varepsilon}_i = 0,$$

$$\sum_{i=1}^{n}(Y_i - \bar{Y})^2 = \sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2 = \sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2,\ \bar{Y} = \frac{\sum_{i=1}^{n} Y_i}{n}，$$

其中$\hat{\beta}_0$及$\hat{\beta}_1$為迴歸參數的最小平方估計值。

（一）$\hat{\varepsilon}_1, \hat{\varepsilon}_2, ..., \hat{\varepsilon}_n$與$Y_1, Y_2, ..., Y_n$的樣本相關係數（sample correlation coefficient）為$(1-R^2)^{1/2}$，試推導之。請詳列推導過程。（10分）

（二）試解釋為何畫殘差圖時，一般用的是($\hat{Y}_1,\hat{\varepsilon}_1$)，（$\hat{Y}_2,\hat{\varepsilon}_2$），⋯，（$\hat{Y}_n,\hat{\varepsilon}_n$），而非($Y_1,\hat{\varepsilon}_1$)，（$Y_2,\hat{\varepsilon}_2$），⋯，（$Y_n,\hat{\varepsilon}_n$）。（5分）

答：

(一)1. $\hat{\varepsilon}_1, \cdots, \hat{\varepsilon}_n$ 與$Y_1, \cdots, Y_n$的樣本相關係數爲

$$r_{Y,\hat{\varepsilon}} = \frac{\sum_{i=1}^{n}(Y_i - \bar{Y})(\hat{\varepsilon}_i - \bar{\hat{\varepsilon}})}{\sqrt{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}\sqrt{\sum_{i=1}^{n}(\hat{\varepsilon}_i - \bar{\hat{\varepsilon}})^2}} = \frac{\sum_{i=1}^{n}(Y_i - \bar{Y})\hat{\varepsilon}_i}{\sqrt{SST}\sqrt{SSE}}\ (\because \bar{\hat{\varepsilon}} = 0)$$

2. $\left(1-R^2\right)^{1/2}=\sqrt{\frac{SSE}{SST}}=\frac{\sqrt{\sum_{i=1}^{n}(\hat{\varepsilon}_i-\bar{\hat{\varepsilon}})^2}}{\sqrt{\sum_{i=1}^{n}(Y_i-\bar{Y})^2}}$

3. $\sum_{i=1}^{n}(Y_i-\bar{Y})\hat{\varepsilon}_i=\sum_{i=1}^{n}Y_i\hat{\varepsilon}_i-\bar{Y}\sum_{i=1}^{n}\hat{\varepsilon}_i=\sum_{i=1}^{n}Y_i\hat{\varepsilon}_i=\sum_{i=1}^{n}Y_i(Y_i-\hat{Y}_i)=\sum_{i=1}^{n}(Y_i-\hat{Y}_i)^2 \quad \sum_{i=1}^{n}\hat{\varepsilon}_i{}^2=SSE$ ($\because$ 由

Normal Equation 可得 $\sum_{i=1}^{n}\hat{Y}_i(Y_i-\hat{Y}_i)=\sum_{i=1}^{n}\hat{Y}_i\hat{\varepsilon}_i=0$)

4.由 1.,2.,3.可得

$$r_{Y,\hat{\varepsilon}}=\frac{\sum_{i=1}^{n}(Y_i-\bar{Y})\hat{\varepsilon}_i}{\sqrt{SST}\sqrt{SSE}}=\frac{SSE}{\sqrt{SST}\sqrt{SSE}}=\sqrt{\frac{SSE}{SST}}=\sqrt{1-R^2}$$

故得證

(二)$\because \sum_{i=1}^{n}\hat{Y}_i\hat{\varepsilon}_i=\sum_{i=1}^{n}(\hat{\beta}_0+\hat{\beta}_1X_i)\hat{\varepsilon}_i=\hat{\beta}_0\sum_{i=1}^{n}\hat{\varepsilon}_i+\hat{\beta}_1\sum_{i=1}^{n}X_i\hat{\varepsilon}_i=0$（其中$\sum_{i=1}^{n}\hat{\varepsilon}_i=0$，$\sum_{i=1}^{n}X_i\hat{\varepsilon}_i=0$ 可由 Normal Equation 推導過程而得）

$\therefore \quad \hat{\varepsilon}_1,\cdots,\hat{\varepsilon}_n$ 與 $\hat{Y}_1,\cdots,\hat{Y}_n$ 的樣本相關係數為

$$r_{\hat{Y},\hat{\varepsilon}}=\frac{\sum_{i=1}^{n}(\hat{Y}_i-\bar{\hat{Y}})(\hat{\varepsilon}_i-\bar{\hat{\varepsilon}})}{\sqrt{\sum_{i=1}^{n}(\hat{Y}_i-\bar{\hat{Y}})^2}\sqrt{\sum_{i=1}^{n}(\hat{\varepsilon}_i-\bar{\hat{\varepsilon}})}}=\frac{\sum_{i=1}^{n}\hat{Y}_i\hat{\varepsilon}_i-n(\bar{\hat{Y}})(\bar{\hat{\varepsilon}})}{\sqrt{\sum_{i=1}^{n}(\hat{Y}_i-\bar{\hat{Y}})^2}\sqrt{\sum_{i=1}^{n}(\hat{\varepsilon}_i-\bar{\hat{\varepsilon}})}}=0$$

亦即兩者無關，故可用以繪製殘差圖並進一步判斷誤差的基本假設，如平均數是否為 0，變異數是否齊一等。

五、考慮以下資料（Y_i, X_{i1}, X_{i2}），i=1,…,5，我們試著以不同的迴歸模式來分析，並希望利用不同的方法選取適合的迴歸模式。以下為可能用到的迴歸模式以及統計量：

迴歸模式	殘差平方和（Residual Sum of Squares）	AIC	BIC
$Y_i=\beta_0+\beta_1X_{i1}+\beta_2X_{i2}+\varepsilon_i$	0.107	-13.2218	-6.7218
$Y_i=\beta_0+\beta_1X_{i1}+\varepsilon_i$	0.6695	-6.0533	-7.1762
$Y_i=\beta_0+\beta_2X_{i2}+\varepsilon_i$	8.927	6.8982	3.1361

$P(F(1,2)>18.51)=0.05$　　$P(F(1,3)>10.13)=0.05$

$P(F(1,4)>7.71)=0.05$　　$P(F(1,5)>6.61)=0.05$

其中 $\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_5$ 為獨立且平均數為0，變異數為 σ^2 之常態分配（normally distributed）隨機變數。

（一）請利用AIC（Akaike's Information Criterion）來選取最好的迴歸模式。（3分）

（二）請利用BIC（Bayesian Information Criterion）來選取最好的迴歸模式。（3分）

（三）若迴歸模式 $Y_i=\beta_0+\beta_1X_{i1}+\beta_2X_{i2}+\varepsilon_i$ 的Mallows' C_P值為3，請計算其他兩個迴歸模式的C_P值，並評論若以C_P值選取模式，這兩個迴歸模式 $Y_i=\beta_0+\beta_1X_{i1}+\varepsilon_i$ 以及 $Y_i=\beta_0+\beta_2X_{i2}+\varepsilon_i$ 何者為較合適的模式？（9分）

（四）若選入一變數所設之顯著水準固定為0.05，且 $\sum_{i=1}^{5}(Y_i-\bar{Y})^2=25.172$，其中 $\bar{Y}=\frac{1}{5}\sum_{i=1}^{5}Y_i$。請利用向前選取法（forward selection method）選取一合適模式。請詳列解答過程。（10分）

答：

(一)因爲 AIC 值最小的爲最適模式，而三個迴歸模式中以 $Y_i=\beta_0+\beta_1X_{i1}+\beta_2X_{i2}+\varepsilon_i$ 的 AIC=-13.2218 最小，所以選此模式爲最適模式。

(二)因爲 BIC 值最小的爲最適模式，而三個迴歸模式中以 $Y_i=\beta_0+\beta_1X_{i1}+\varepsilon_i$ 的 BIC=-7.1762 最小，所以選此模式爲最適模式。

(三) $C_p=\dfrac{SSE_p}{MSE_F}-(n-2p)=(n-p)\dfrac{SSE_p}{SSE_F}-(n-2p)=(n-p)\dfrac{1-R_p^2}{1-R_F^2}-(n-2p)$，其中 p 爲迴歸參數個數。

當 p=2 時，模式 $Y_i=\beta_0+\beta_1X_{i1}+\varepsilon_i$ 的 $C_2=(5-2)\dfrac{0.6695}{0.107}-(5-2\cdot2)$=17.7710；

而模式 $Y_i=\beta_0+\beta_1X_{i2}+\varepsilon_i$ 的 $C_2=(5-2)\dfrac{8.927}{0.107}-(5-2\cdot2)$=249.2897

因爲 C_p 值愈小，且愈接近 p 者，爲最佳自變數組合。若僅就 p=2 的這兩個模式來比較，則 $Y_i=\beta_0+\beta_1X_{i1}+\varepsilon_i$ 應爲較合適的模式。

(四)1.向前選取法的步驟如下：

(1)令 $F_i=\dfrac{MSR(X_i)}{MSE(X_i)}$ 爲只將一個自變數 X_i 引進迴歸模型的 F 檢定統計量值，自 F_i 中選取最大的（假設爲 F_1），並與 $F_\alpha(1,n-2)$ 作比較，

若 $F_1 < F_\alpha(1, n-2)$ ，代表所有自變數相對應變數而言都不是重要的變數；若 $F_1 > F_\alpha(1, n-2)$ ，則將自變數 X_1 納入迴歸模型中。

(2)在已選入的自變數外，再增加一個自變數引入迴歸模型中，並計算其偏檢定值 $F = \dfrac{MSR(X_2 \mid X_1)}{MSE(X_1, X_2)}$ ，

並與 $F_\alpha(1, n-3)$ 作比較，

若 $F < F_\alpha(1, n-3)$ ，則自變數 X_2 相對應變數而言不是重要的變數；

若 $F > F_\alpha(1, n-3)$ ，則將自變數 X_2 納入迴歸模型中。

2.就此題而言：

(1)因爲

$$F_1 = \frac{MSR(X_1)}{MSE(X_1)} = \frac{SSR(X_1)}{SSE(X_1)} \cdot \frac{n-k}{k-1} = \frac{R_1^2}{1-R_1^2} \cdot \frac{n-k}{k-1} = \frac{0.9734}{1-0.9734} \cdot \frac{5-2}{2-1} = 109.7946 \text{ ，}$$

$$F_2 = \frac{MSR(X_2)}{MSE(X_2)} = \frac{SSR(X_2)}{SSE(X_2)} \cdot \frac{n-k}{k-1} = \frac{R_2^2}{1-R_2^2} \cdot \frac{n-k}{k-1} = \frac{0.6454}{1-0.6454} \cdot \frac{5-2}{2-1} = 5.4593$$

因爲 $F_1 > F_{0.05}(1,3) = 10.13$ ，所以將自變數 X_1 納入迴歸模型中。

（其中 $R_i^2 = 1 - \dfrac{SSE(X_i)}{SST(X_i)}$ ， $R_1^2 = 1 - \dfrac{SSE(X_1)}{SST(X_1)} = 1 - \dfrac{0.6695}{25.172} = 0.9734$ ，

$R_2^2 = 1 - \dfrac{SSE(X_2)}{SST(X_2)} = 1 - \dfrac{8.927}{25.172} = 0.6454$ ）

(2)將另一個自變數 X_2 引入迴歸模型中，並計算其偏檢定值

$$F = \frac{MSR(X_2 \mid X_1)}{MSE(X_1, X_2)} = \frac{[SSE(X_1) - SSE(X_2)]/1}{SSE(X_1, X_2)/(n-3)} = \frac{[0.6695 - 0.107]/1}{0.107/(5-3)} = 10.5140 \text{ 因爲}$$

$F < F_{0.05}(1,2) = 18.51$ ，所以自變數 X_2 相對應變數而言不是重要的變數；

所以由向前選取法知，最適模型爲

$Y_i = \beta_0 + \beta_1 X_{i1} + \varepsilon_i$