109 年特種考試地方政府公務人員考試試題

等　別：三等考試
類　科：統計
科　目：迴歸分析

參考之查表值：F 分布 α=0.05 臨界值 F0.05(df1, df2)

		df1	
		1	2
df2	27	4.2100	3.3541
	28	4.1960	3.3404
	402	3.8647	3.0185
	403	3.8646	3.0181

$t_{0.025}(28)=-2.0484$, $t_{0.025}(30)=-2.0422$

一、一位主管欲知道碩士級分析師的月薪是否可以用年資來預測，以作為未來給薪的參考。他收集了 30 個樣本觀察值，資料包含年資（X，以年為單位） 和月薪 （Y，以千元為單位）。請依據下面數據和圖 1 回答問題。

$$\bar{X}=5.34,\ \bar{Y}=76,\ S_{XY}=\sum_{i=1}^{n}(X_i-\bar{X})(Y_i-\bar{Y})=2198,$$

$$S_{XX}=\sum_{i=1}^{n}(X_i-\bar{X})^2=232.072,\ S_{YY}=\sum_{i=1}^{n}(Y_i-\bar{Y})^2=21890$$

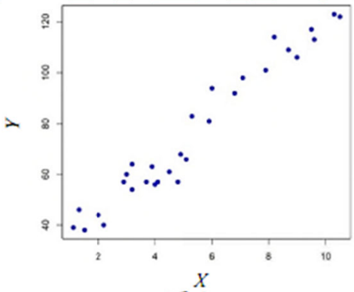

圖 1

㈠在配適 $Y_i=\beta_0+\beta_1X_i+\varepsilon_i$ 的簡單線性迴歸方程式下，請利用最小平方法計算參數 β_0 和 β_1 估計值（estimates）。如果將模型改為 $Y_i=\alpha+\beta_1(X_i-\bar{X})+\varepsilon_i$，請寫出參數 α 和 β_1 最小平方估計式（least-squares estimators）及其估計標準誤（standard errors）。（12 分）

㈡假設 $Y_i\sim N(\beta_0+\beta_1X_i,\sigma^2)$，請在顯著水準 $\alpha=0.05$ 下，檢定 $H_0:\beta_1=0$。請試述檢定統計量之值、決策法則和結論。請寫出在應用最大概似估計（Maximumlikelihood estimation）法， σ^2 的估計值。請寫出利用最小平方法， σ^2 的估計值。（10 分）

㈢請問年資是 5 年的碩士級分析師之平均薪資的 95%信賴區間。（4 分）

【解題關鍵】
1.《考題難易》：★
2.《破題關鍵》：最小平方法與最大概似估計屬課內基本內容，100 年高考有類似考題，可參考迴歸分析課本 P.2-43 頁相同範例。而自變數中心化之估計量與標準誤，在 106 年高考幾乎是完全相同的考法，可迴歸分析課本 P.3-24 頁相同範例。而平均值信賴區間則可雷同 99 年地特試題，可參考迴歸分析課本 P.2-47 頁相同範例。

【擬答】

(一) 1. 令 $Q=\sum \varepsilon_i^2=\sum (Y_i-\beta_0-\beta_1 X_i)^2$

$$\frac{\partial Q}{\partial \beta_0}=-2\sum (Y_i-\beta_0-\beta_1 X_i)=0 \Rightarrow n\beta_0+\sum X_i\beta_1=\sum Y_i$$

$$\frac{\partial Q}{\partial \beta_1}=-2\sum (Y_i-\beta_0-\beta_1 X_i)X_i=0 \Rightarrow \sum X_i\beta_0+\sum X_i^2\beta_1=\sum X_iY_i$$

得 $\hat{\beta}_1=\dfrac{S_{XY}}{S_{XX}}=\dfrac{2198}{232.072}=9.4712$

$$\hat{\beta}_0=\bar{Y}-\hat{\beta}_1\bar{X}=76-9.4712\times 5.34=25.4238$$

2. 令 $\because X_i^*=X_i-\bar{X}\Rightarrow \bar{X}^*=0$

$$\hat{\beta}_1=\frac{S_{X^*Y}}{S_{XX^*}}=\frac{\sum X_i^{*2}Y_i-0}{\sum X_i^{*2}-0}=\frac{\sum (X_i-\bar{X})Y}{\sum (X_i-\bar{X})^2}=9.4712$$

$$\hat{\alpha}=\bar{Y}-\hat{\beta}_1\bar{X}^*=\bar{Y}-\hat{\beta}_1\cdot 0=\bar{Y}=76$$

$$SSR=\hat{\beta}_1^2\cdot S_{XX}=9.4712^2\times 232.072=20817.7007$$

$$SSE=SSTO-SSR=21890-20817.7007=1072.2993$$

$$MSE=\frac{1072.2993}{30-2}=38.2964$$

$$S(\hat{\beta}_1)=\sqrt{Var(\frac{S_{XY}}{S_{XX}})}=\sqrt{\frac{MSE}{S_{XX}}}=\sqrt{\frac{38.2964}{232.072}}=0.4062$$

$$S(\hat{\alpha})=\sqrt{Var(\bar{Y})}=\sqrt{\frac{MSE}{n}}=\sqrt{\frac{38.2964}{30}}=1.1298$$

(二) 1. $H_0:\beta_1=0 \quad H_1:\beta_1\neq 0$

$\alpha=0.05$

$$T^*=\frac{\hat{\beta}_1}{\sqrt{\frac{MSE}{S_{XX}}}}=\frac{9.4712}{\sqrt{\frac{38.2964}{232.072}}}=23.315\in C$$

$C:\{|T^*|>t_{0.975}(28)=2.0484\}$

拒絕 H_0，有顯著證據說 $\beta_1\neq 0$，所以年資可顯著預測月薪。

2. $f(y_i)=(2\pi\sigma^2)^{-\frac{1}{2}}e^{-\frac{(Y_i-\beta_0-\beta_1X_i)^2}{2\sigma^2}}$

$$L(\theta)=L(\beta_0,\beta_1,\sigma^2)=f(y_1,y_2,\ldots,y_n)=(2\pi\sigma^2)^{-\frac{n}{2}}e^{-\frac{\sum(Y_i-\beta_0-\beta_1X_i)^2}{2\sigma^2}}$$

$$\frac{\partial q}{\partial \beta_0}=\frac{1}{\sigma^2}\sum (Y_i-\beta_0-\beta_1X_i)=0 \quad ①$$

$$\frac{\partial q}{\partial \beta_1}=\frac{1}{\sigma^2}\sum (Y_i-\beta_0-\beta_1X_i)X_i=0 \quad ②$$

$$\frac{\partial q}{\partial \sigma^2}=-\frac{n}{2}\cdot\frac{1}{\sigma^2}+\frac{\sum (Y_i-\beta_0-\beta_1X_i)^2}{2\sigma^4} \quad ③$$

①②兩式整理之後恰為最小平方法中之正規方程式，所以 β_0,β_1 之 MLE 與 LSE 相同，代入③

$$\hat{\sigma}^2_{MLE}=\frac{\sum(Y_i-\hat{\beta}_0-\hat{\beta}_1X_i)^2}{n}=\frac{SSE}{n}=\frac{1072.2993}{30}=35.7433$$

而最小平方法並無限定，所以採不偏估計

$$\hat{\sigma}^2_{LSE}=\frac{\sum(Y_i-\hat{\beta}_0-\hat{\beta}_1X_i)^2}{n-2}=MSE=38.2964$$

㈢迴歸方程式為：$\hat{Y}=25.4238+9.4712X$

$E(Y\mid X=5)$ 之 95%信賴區間為

$$\hat{Y}_h\pm t_{\frac{\alpha}{2}}(n-2)\sqrt{(\frac{1}{n}+\frac{(x_h-\overline{X})^2}{SS_X})MSE}$$

$$\Rightarrow(25.4238+9.4712\times5)\pm t_{0.975}(28)\sqrt{(\frac{1}{30}+\frac{(5-5.34)^2}{232.072})38.2964}$$

$$\Rightarrow[70.4482,75.1114]$$

二、㈠一位分析師受託分析一組資料。資料來自於 20 位 25 歲至 34 歲的健康女性，其中包括反應變數 Y（身體脂肪）和三個解釋變數（X_1：皮褶厚度，X_2：大腿圓周和 X_3：中臂圓周）用作預測身體脂肪。該分析師初步配適一個迴歸模型如下：

$$Y_i=\beta_0+\beta_1X_{1i}+\beta_2X_{2i}+\beta_3X_{3i}+\varepsilon_i\qquad i=1,\dots,20$$

另外，表 1 計算解釋變數之間的解釋能力。

表1

反應變數	解釋變數	判定係數 R^2
X_1	X_2,X_3	99.86%
X_2	X_1,X_3	99.82%
X_3	X_1,X_2	99.04%

請由表 1 計算變異數膨脹因子（variance inflation factor, VIF）評論該分析師所配適的迴歸模型 1 是否合適？如果不合適，請詳細說明原因和解決方法。（8 分）

㈡一位分析師受託分析影響縣市首長滿意度的重要因素。滿意度分數 Y（以 1～10 為評分範圍，分數愈高代表愈滿意）作為反應變數。該分析師找到一些重要的解釋變數。依據他所配適的複迴歸模型，有些預測值有超過 10 的情況。請問該分析師所配適的複迴歸模型是否合適？如果不合適，請詳細說明原因和解決的方法。（6 分）

㈢一位分析師分析 2017 年 1 月至 2019 年 12 月的旅遊人數月資料。該分析師配適的迴歸模型如下：

模型 2

$$\ln(y_t)=\beta_0+\beta_1 t+\beta_2 M_1+\beta_3 M_2+\ldots+\beta_{12}M_{11}+\varepsilon_t,\ \varepsilon_t\sim N(0,\sigma^2)$$

此處 t 是時間，ε_t 為獨立且具有共同分配其平均數為 0 變異數 σ^2 的常態分配，M_i 是虛擬變數，第 i 個月為 1，其他月份為 0，i=1, 2, ⋯, 11。請說明在線性迴歸模型下，如何檢查誤差項的所有假設是否有違反。圖 2 是模型 2 的標準化殘差值（studentized residual）對應時間的殘差圖。請問該分析師所配適的複迴歸模型是否合適？如果不合適，請詳細說明原因和解決的方法。（10 分）

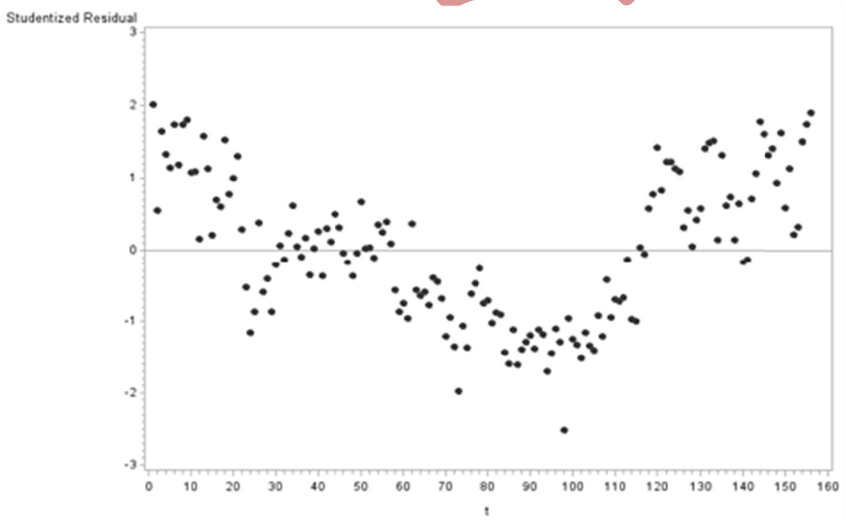

圖2

【解題關鍵】

1. 《考題難易》：★★

2. 《破題關鍵》：變異數膨脹因子的計算在 101 年高考與 103 年高考有類似考題，屬課內基本內容，可參考迴歸分析課本 P.7-96 頁相同範例。預測值超出範圍可能需要有分析經驗才有辦法聯想到是與前提假設有關，但可由第三小題的概念猜想到出題老師的想法。殘差分析的內容，乃至於線性模型檢驗與非線性模型的修正，在 107 年地特有相似命題，可參考迴歸分析課本 P.4-23 頁相同範例。

【擬答】

㈠ $VIF_{X_i}=\dfrac{1}{1-R^2(X_i\text{ 對其他自變數作複迴歸})}$

$$VIF_{X_1}=\frac{1}{1-0.9986}=714.29>10$$

$$VIF_{X_2}=\frac{1}{1-0.9982}=555.56>10$$

$$VIF_{X_3}=\frac{1}{1-0.9904}=104.17>10$$

VIF 值大於 10，則此 X_i 可被其他自變數解釋或取代。此例 VIF 值皆大於 10，則這些解釋變數之間存在共線性問題。

解決的辦法可去除相依的自變數，或採用脊迴歸方法來消除共線性問題。

㈡配適預測值會超過範圍，常見的原因是資料違背前提假設，或者是資料並非線性的分布。以本例而言 Y 的範圍是 1~10 分，主要的原因可能是分布並非線性的情況，所以該分析師所配適的複迴歸模型可能不適合。初步可透過散布圖觀察資料是否符合線性，進一步嘗試一些非線性的模型進行配適。

㈢ *1.* 常態性檢驗

⑴初步圖形檢驗可採用 qq-plot 圖與 pp-plot 圖，若服從常態分配，則散布的點會非常靠近 x＝y 線上。

⑵統計檢定可採用卡方適合度檢定、Kolmogorov-Smirnov 檢定或 Lilliefors 檢定。

2. 同質性檢驗

⑴初步圖形檢驗可採用殘差圖，觀察殘差是否會隨自變數變化而隨之變化的趨勢。

⑵統計檢定可採用 White Test 或 Bartlett Test。

3. 獨立性檢驗

⑴初步圖形檢驗可採用殘差圖或散布圖，可以直接觀察資料是否有二次或更高次的相關。

⑵統計檢定可採用適缺度檢定。

圖 2 的標準化殘差值可以明顯發現存在時間的相依性，隨著時間呈現非線性的變化，可以可考慮將變項 t 轉換為非線性，例如加入二次方項 t^2 至模型中。

三、一位數據分析師受託分析於 33（n=33）位男學生，其腳長（Y，以公分為單位）和 X 身高（以英吋為單位）的關係。所建立的簡單線性模型如下：

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

請使用表 2 部分電腦輸出報表來回答以下問題。表 2 第一欄是觀察值的順序，第二欄是殘差值。

㈠請說明何謂異常點（outlier）和高槓桿觀察值（high leverage observation），及其之間的區別。（8 分）

㈡表 2 第三欄是標準化的殘差值（studentized residual）。請以此判斷是否有異常點存在？請說明判斷準則。

表 2 第五欄是 Student 化刪除殘差（Studentized deleted residuals，以 R-Student 表示）。第 i 個 R-Student 殘差是在假定將資料中的第 i 個觀察值刪除，然後以剩下的 n-1 個觀察值來建立新的估計迴歸方程式而標準化獲得的 R-Student 殘差值。請以此判斷是否有異常點存在？請說明判斷準則。（8 分）

㈢表 2 第六欄是 h_{ii}（hat value），其公式為 $h_{ii} = \dfrac{1}{n} + \dfrac{(X_i - \bar{X})^2}{\sum_{j=1}^{n}(X_j - \bar{X})^2}$，請問 $\sum_{i=1}^{n} h_{ii}$ 的值為何？請以此判斷是否有可能的高槓桿觀察值存在？請說明判斷準則。表 2 的最後一欄，第八欄是 DFFITS（Difference in Fits）值。請以此判斷是否有可能的影響點（influential observation）存在？請說明判斷準則。（8 分）

表2

Obs	Residual	Student Residual	Cook's D	R-Student	Hat Diag	Cov Ratio	DFFITS
1	0.541	0.443	0.011	0.438	0.101	1.173	0.147
2	0.906	0.718	0.009	0.712	0.035	1.070	0.136
3	-1.777	-1.410	0.041	-1.434	0.040	0.974	-0.293
4	0.390	0.308	0.002	0.304	0.033	1.097	0.056
5	-0.977	-0.772	0.010	-0.767	0.032	1.061	-0.140
6	-1.510	-1.194	0.024	-1.203	0.033	1.005	-0.222
7	1.490	1.179	0.024	1.186	0.033	1.007	0.219
8	-0.160	-0.127	0.000	-0.125	0.045	1.117	-0.027
9	1.023	0.809	0.011	0.804	0.032	1.057	0.147
10	-0.510	-0.403	0.003	-0.398	0.033	1.093	-0.073
11	1.957	1.563	0.067	1.602	0.052	0.956	0.374
12	0.157	0.125	0.000	0.123	0.052	1.125	0.029
13	1.023	0.809	0.011	0.804	0.032	1.057	0.147
14	0.556	0.444	0.005	0.438	0.050	1.110	0.101
15	-0.777	-0.614	0.006	-0.608	0.032	1.077	-0.111
16	-0.243	-0.192	0.001	-0.189	0.030	1.099	-0.034
17	-2.043	-1.632	0.073	-1.679	0.052	0.941	-0.392
18	-1.810	-1.458	0.078	-1.486	0.068	0.994	-0.402
19	0.140	0.110	0.000	0.109	0.031	1.101	0.019
20	2.356	1.944	0.236	2.041	0.111	0.926	0.721

21	0.623	0.522	0.022	0.516	0.141	1.221	0.209
22	0.490	0.388	0.003	0.382	0.033	1.093	0.071
23	0.790	0.627	0.008	0.620	0.039	1.083	0.125
24	-0.843	-0.697	0.031	-0.691	0.114	1.168	-0.248
25	-0.810	-0.641	0.007	-0.635	0.033	1.075	-0.117
26	1.490	1.179	0.024	1.186	0.033	1.007	0.219
27	0.490	0.388	0.003	0.382	0.033	1.093	0.071
28	-3.545	-3.437	3.274	-4.299	0.357	0.636	-3.200
29	0.089	0.073	0.000	0.072	0.086	1.168	0.022
30	0.257	0.203	0.001	0.200	0.030	1.098	0.035
31	-1.277	-1.013	0.021	-1.014	0.040	1.040	-0.207
32	1.323	1.065	0.040	1.067	0.066	1.061	0.283
33	0.190	0.153	0.001	0.151	0.068	1.144	0.041

【解題關鍵】

1.《考題難易》：★★★

2.《破題關鍵》：異常點和高槓桿觀察值乃至於影響點的判斷，在近年考試常常出現，106 年高考考了異常點與影響點的差別，而本題則是考了異常點和高槓桿觀察值的差異與用途，算是考得非常細的問題，不易作答，可參考迴歸分析課本 P.4-37 頁類似範例。至於影響點的判斷準則，108 年高考出過類似判讀報表考題，可參考迴歸分析課本 P.4-35 頁相同範例。本題相對困難的槓桿值總和與槓桿值平均數，則是首次出現在考題中，觀念雖簡單，但需要有準備過才有辦法完整作答。

【擬答】

㈠ *1.* 異常點：某些觀察個案與其他資料間有明顯的區隔，離群值可能是對 Y 異常或是對 X 異常，亦可能同時對 X 與 Y 均異常。可採用標準化殘差 $d_i=\dfrac{e_i}{\sqrt{MSE}}$ 的絕對值 $|d_i|$ 是否大於 2；或採用 t 化殘差 $r_i=\dfrac{e_i}{\sqrt{MSE(1-h_{ii})}}$ 的絕對值 $|r_i|$ 是否大於 3 來判斷是否為異常點。所以異常點為殘差很大的點。

2. 高槓桿觀察值：當一個槓桿值之大小超過平均槓桿值的兩倍時，通常會被視為高槓桿觀察值，即 $h_{ii}>2\dfrac{p}{n}$ 可能是潛在成為對模型的影響點。所以高槓桿點為遠離樣本空間中心的點。

㈡標準化殘差的絕對值大於 2 可視為異常點，t 化殘差的絕對值大於 3 可視為異常點，所以表 2 第三欄中，有一個觀察值的絕對值超過 3，即第 28 筆觀察值為異常值。

註：標準化殘差稱為為 standarized residual，t 化殘差才稱為 studentized residual。

㈢ *1.* $\sum_{i=1}^{n} h_{ii}=trace(H)$

$=trace(X(X'X)^{-1}X')=trace(X'X(X'X)^{-1})=trace(I_p)=p$

2. 若要判斷高槓桿觀察值，可觀察槓桿值是否超過平均的兩倍，即 $h_{ii}>2\dfrac{p}{n}=2\dfrac{2}{33}=0.1212$。所以表 2 第六欄中，有兩個觀察值得絕對值超過 3，即第 21 筆與第 28 筆觀察值為高槓桿觀察值。

3. 當 $(DFFITS)_i > 2\sqrt{p/n} = 2\sqrt{2/33} = 0.4924$，則該樣本可以視為具有影響力。所以表 2 第八欄中，有兩個觀察值得絕對值超過 0.4924，即第 20 筆與第 28 筆觀察值為影響點。

四、一位統計分析師受託預測單位面積房價，欲了解房價受到那些因素所影響。收集了 408 筆有關於單位面積房價，屋齡（X_1，以年為單位），到最近的地鐵站的距離（X_2），便利商店數量（X_3），房屋座落的緯度（X_4）和經度（X_5）。擬考慮的模型如下：

模型 1　$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \beta_5 X_{5i} + \varepsilon_i,\quad i = 1,\ldots,n$

模型 2　$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \varepsilon_i,\quad i = 1,\ldots,n$

模型 3　$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_4 X_{4i} + \beta_5 X_{5i} + \varepsilon_i,\quad i = 1,\ldots,n$

請使用表 3 部分電腦輸出三個模型的變異數分析表（ANOVA, Analysis of Variance）報表來回答以下問題。

表3 模型1 ANOVA表

Response：Y	DF	Sum of squares	Mean square	F value	P-value
Model	5	44260	8852.03227	134.46	<.0001
Error	402	26465	65.83443		
Corrected Total	407	70726			

模型2 ANOVA表

Response：Y	DF	Sum of squares	Mean square	F value	P-value
Model	3	41703	13901	193.50	<.0001
Error	404	29023	71.83833		
Corrected Total	407	70726			

模型3 ANOVA表

Response：Y	DF	Sum of squares	Mean square	F value	P-value
Model	4	41879	10470	146.27	<.0001
Error	403	28847	71.57982		
Corrected Total	407	70726			

㈠在考慮模型 1 之下，請檢定便利商店數量（X_3）這個解釋變數是否可以從給定模型 1 中刪除。請用顯著水準 $\alpha = 0.05$ 檢定並敘述對立假設、檢定統計量之值、決策法則和結論。（8 分）

㈡在考慮模型 1 之下，請檢定房屋座落的緯度（X_4）和經度（X_5）這兩個解釋變數是否在模型 1 對預測單位面積房價有影響。亦即請用 $\alpha = 0.05$ 檢定 $H_0: \beta_4 = \beta_5 = 0$，並請敘述對立假設、檢定統計量之值、決策法則和結論。（8 分）

㈢採用 F 檢定法，說明向後消去法 （Backward elimination, stay level=0.05）準則的選模過程，並列出所選取之模式。（10 分）

㈣請計算模型 1，2 和 3 的調整的複判定係數 R^2（the adjusted R-squared）並試述其意義。請敘述(1)(2)檢定，模型誤差項所需要的假設，並綜合(1)(2)檢定結果，請說明在模型 1，2 和 3 中，何者模式為最佳模型。（10 分）

【解題關鍵】

1.《考題難易》：★

2.《破題關鍵》：模型選擇之偏 F 檢定屬萬年必考之考古題，此題在 107 年地特有類似命題，可參考迴歸分析課本 P.7-44 頁相同範例。而誤差所需假設同樣在 107 年地特有出題。而利用調整後判定係數進行模型選擇，則在 106 年地特有類似命題，可參考迴歸分析課本 P.7-57 頁相同範例。

【擬答】

㈠ $H_0: \beta_3 = 0 \qquad H_1: \beta_3 \neq 0$

$\alpha = 0.05$

$$F^* = \frac{SSR(X_3 \mid X_1, X_2, X_4, X_5)/1}{MSE(X_1, X_2, X_3, X_4, X_5)} \sim F(1, 402)$$

$$= \frac{44260 - 41879}{65.83443} = 36.166 \in C$$

$C: \{F^* > F_{0.05}(1, 402) = 3.8647\}$

拒絕 H_0，有顯著證據說在有其他變項下，需要加入 X_3

所以便利商店數量不需從模型 1 中刪除

㈡ $H_0: \beta_4 = \beta_5 = 0 \qquad H_1: \beta_4, \beta_5$ 不全為 0

$\alpha = 0.05$

$$F^* = \frac{SSR(X_4, X_5 \mid X_1, X_2, X_3)/2}{MSE(X_1, X_2, X_3, X_4, X_5)} \sim F(2, 402)$$

$$= \frac{44260 - 41703}{65.83443} = 38.84 \in C$$

$C : \{F^* > F_{0.05}(2, 402) = 3.0182\}$

拒絕 H_0，有顯著證據說在有其他變項下，需要加入 X_4 與 X_5

註：此題變異數分析表數字多有瑕疵，並不精確。

㈢ 1. 當模式引入一新的解釋變數時，只要此一解釋變數稍具解釋力，模式之複判定係數必然提高。然而，解釋變數每增加一個，代表誤差自由度隨之減少一個，代表模式之解釋能力可能降低，複判定係數未能顯示此情形，實務上可計算調整後判定係數 $R^2_{adj} = 1 - \frac{SSE}{SSTO}\frac{n-1}{n-p}$ 來評估多元迴歸模式的解釋度，亦可作為模型選擇的準則。

模型 1 調整的複判定係數

$$R^2_{adj} = 1 - \frac{26465}{70726}\frac{407}{402} = 62.12\%$$

模型 2 調整的複判定係數

$$R^2_{adj} = 1 - \frac{29023}{70726}\frac{407}{404} = 58.66\%$$

模型 3 調整的複判定係數

$$R^2_{adj} = 1 - \frac{28847}{70726}\frac{407}{403} = 58.81\%$$

因為模型 1 調整的複判定係數最大，所以模型 1 為最佳模型。

2. 前述模型選擇之偏 F 檢定之前題假設為：

$\varepsilon_i \sim N(0, \sigma^2)$

(1) $E(\varepsilon_i) = 0$ 、 $Var(\varepsilon_i) = \sigma^2$

(2) $Cov(\varepsilon_i, \varepsilon_j) = 0, \forall i \neq j$

(3) $Cov(\varepsilon_i, X_i) = 0, \forall i$