

等 別：三等考試

類 科：統計

科 目：迴歸分析

考試時間：2 小時

座號：_____

※注意：(一)可以使用電子計算器。

(二)不必抄題，作答時請將試題題號及答案依照順序寫在試卷上，於本試題上作答者，不予計分。

(三)本科目除專門名詞或數理公式外，應使用本國文字作答。

- 一、考慮一多元線性迴歸模型，其反應變數為 Y ，解釋變數為 X_1, X_2, \dots, X_k ，有 n 個觀測值，線性迴歸模型為 $Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik} + \varepsilon_i, i = 1, \dots, n$ ，其中誤差項 ε_i 之期望值為 0，變異數為 σ^2 ，且兩兩獨立，此模型以向量及矩陣方式表示為 $Y = X\beta + \varepsilon$ (*)，其中

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}_{n \times 1}, X = \begin{bmatrix} 1 & X_{11} & \cdots & X_{1k} \\ 1 & X_{21} & \cdots & X_{2k} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & X_{n1} & \cdots & X_{nk} \end{bmatrix}_{n \times (k+1)}, \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix}_{(k+1) \times 1}, \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}_{n \times 1}$$

請回答下列問題：(每小題 5 分，共 30 分)

- (一)以向量及矩陣方式，試求出參數向量 β 之最小平方估計量向量 b 。
 (二)承題(一)，令 A 為一個 $2 \times (k+1)$ 的矩陣，求 Ab 之變異數-共變異數矩陣。
 (三)配適值向量表為 $\hat{Y} = HY$ ，寫出矩陣 H 。
 (四)求出殘差向量 $e = Y - \hat{Y}$ 之變異數-共變異數矩陣。
 (五)令 A 為對稱矩陣，則 $Y'AY$ 稱為 Y 之二次式，將此模型之 SSE (error sum of square) = $e'e$ 表成二次式，其中 Y' 和 e' 分別是 Y 和 e 之轉置矩陣。
 (六)求出 β 之最大概似估計量，對誤差項向量需要什麼假設。

- 二、某無人車研發公司欲預測它的行車時間 Y ，考慮了三個高度相關的解釋變數分別是行駛里數 X_1 ，車種 X_2 ，車齡 X_3 ，收集過去 20 個月資料得到

$$SSR(X_1, X_2, X_3) = 4000,$$

$$SSR(X_1) = 1000,$$

$$SSR(X_1|X_2) = 600,$$

$$SSE(X_1, X_2, X_3) = 800,$$

請回答下列問題：(每小題 5 分，共 15 分)

(一)求偏判定係數 $r_{YX_3.X_1X_2}^2$ 。(二)檢定偏相關係數 $\rho_{YX_3.X_1X_2}$ 是否為 0，請求出 F 檢定統計量的值。(三)求偏判定係數 $r_{YX_1X_2.X_3}^2$ (以最簡分數表示)。

(請接背面)

等 別：三等考試
類 科：統計
科 目：迴歸分析

三、一國內規模最大的律師事務所專門辦理職災案件，總經理想了解他們在捷運上的廣告有沒有增加他們的業務量，根據過去隨機抽取 11 月的資料，利用兩個解釋變數：一為單月廣告費用 X_1 （單位為百萬元，平均值為 1 單位，標準差為 2 單位），另一個為主要競爭對手單月廣告費用 X_2 （單位為百萬元，平均值為 1 單位，標準差為 2 單位）來預測職災案件的每月增加件數 Y （單位為件，平均值為 3 件，標準差為 2 件），下表是以不同解釋變數配適每月增加件數 Y 之迴歸模型的參數最小平方估計量和誤差平方和 SSE。

迴歸模型代號	迴歸模型中的解釋變數	參數最小平方估計量	SSE
LM1	X_1	$b_1 = 0.3$	5
LM2	X_2	$b_2 = -0.1$	8
LM3	X_1, X_2	$b_1 = 0.2, b_2 = -0.2$	4

下面所有小題的計算若除不盡，一律四捨五入到小數第二位，否則不給分。

- (一)分別求出三個迴歸模型 LM1~LM3 之判定係數。(6分)
 (二)分別求出三個迴歸模型 LM1~LM3 之修正判定係數。(6分)
 (三)使用題(二)的結果求出 Y 和 X_1 之相關係數以及 Y 和 X_2 之相關係數。(4分)
 (四)針對迴歸模型 LM3 於試卷上依序填入下列 ANOVA 表中(1)~(8)之 8 個空格內容。(8分)

Analysis of Variance				
Source	DF	Sum of Squares	Mean Square	F Value
Model	(1)	(3)	(6)	(8)
Error	(2)	(4)	(7)	
Corrected Total	10	(5)		

- (五)針對迴歸模型 LM3 欲檢定 X_2 的係數是否為 0，求出偏 F 檢定的計算值。(6分)
 (六)針對迴歸模型 LM3，當迴歸係數在什麼條件下，MSR 的期望值為 σ^2 ？(5分)
 (七)若三個變數之變異數-共變異數矩陣為：

$$V = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 4 & 1 \\ 0 & 1 & 16 \end{bmatrix}$$

為了求其相關係數矩陣，需在 V 前後乘一個對角矩陣，寫出此對角矩陣。(5分)

四、某大數據資料分析公司以 Y 對三個變數 X_1, X_2, X_3 所做的複迴歸分析中，樣本大小 $n = 14$ ，得到複判定係數 $R^2 = 90\%$ ，又將 Y 改對變數 X_2 做迴歸分析時，得到判定係數 $R^2 = 70\%$ ，請回答下列問題：(每小題 5 分，共 15 分)

- (一)此公司欲檢定 X_1, X_3 的迴歸係數是否為 0，請求出 F 檢定統計量的值。
 (二)設速食店營收為 Y （單位為百萬元），廣告費為 X_1 ，有得來速 (drive-through) 服務時 $X_2 = 1$ ，沒有則 $X_2 = 0$ ，做迴歸分析得到迴歸平面為 $Y = 1 + 1.5X_1 + 2X_2$ ，因有得來速服務多出來的平均營收為多少元？
 (三)承題(二)，令 R_j^2 代表解釋變數 X_j 對另一個解釋變數做迴歸分析得到的判定係數， $j = 1, 2$ ，且 $R_1^2 = 0.65, R_2^2 = 0.95$ ，求出 X_2 之變異數膨脹因子 (VIF)，若 VIF 大於 10 代表模式有何問題？