

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。

本試題可能使用之查表值如下：

$t_{9,0.05} = 1.833$ ,  $t_{9,0.025} = 2.262$ ,  $t_{10,0.05} = 1.812$ ,  $t_{10,0.025} = 2.228$ ,  $t_{18,0.01} = 2.552$ ,  $t_{18,0.02} = 2.214$ ,  
 $t_{19,0.01} = 2.539$ ,  $t_{19,0.02} = 2.205$ ,  $t_{20,0.01} = 2.528$ ,  $t_{20,0.02} = 2.197$   
 $F_{1,5,0.05} = 6.61$ ,  $F_{1,6,0.05} = 5.99$ ,  $F_{2,5,0.05} = 5.79$ ,  $F_{2,6,0.05} = 5.14$

一、若  $n=10$ ,  $\sum X_i=70$ ,  $\sum Y_i=185$ ,  $\sum X_i^2=652$ ,  $\sum Y_i^2=3793$ ,  $\sum X_i Y_i=1537$ ，請計算表中所列(1)~(4)的簡單線性迴歸的參數估計值與標準誤。(16分)

參數	估計值	標準誤
$\beta_0$	$b_0 =$ (1)	(2)
$\beta_1$	$b_1 =$ (3)	(4)

二、下表所列為四個自變數 $X_1, X_2, X_3, X_4$ 的複迴歸問題，分別以其中一個為因變數，而以其他三個為自變數時，所得到的判定係數 $R^2$ 。請計算(1)~(4)變異數膨脹因子 (Variance Inflation Factor, VIF)。(8分)

因變數	判定係數 $R^2$	變異數膨脹因子 VIF
$X_1$	0.0842	(1)
$X_2$	0.9642	(2)
$X_3$	0.1141	(3)
$X_4$	0.5687	(4)

三、某公司受委託進行飲料公司飲料自動販賣機的定期維修。該公司發現自動販賣機的每半個月維修費支出 (Y：以元為單位)，與自動販賣機之銷售量 (X：以千元為單位) 密切相關，為進一步瞭解二者間之關係，隨機抽取若干台飲料自動販賣機，記錄其每半個月維修費支出及銷售量金額，使用簡單線性迴歸模式進行分析。假設取得之樣本為：

i	1	2	3	4	5	6	7	8	9	10	11
$X_i$	2.5	2.7	3.2	3.8	4.0	4.4	4.8	5.4	6.9	7.9	8.8
$Y_i$	63	60	64	69	70	76	93	102	111	128	133

(一)請利用最小平方法計算參數 $\beta_0$ 與 $\beta_1$ 之估計值。(10分)(二)請計算判定係數 $R^2$ 之值。(5分)(三)請計算參數 $\beta_1$ 估計值 $b_1$ 之估計標準誤 (standard error)。(5分)

(請接第二頁)

類 科：統計  
科 目：迴歸分析

四、若因變數Y和三個自變數 $X_1$ ， $X_2$ ，和 $X_3$ 的複迴歸所得之ANOVA表如下：

Source	df	SS
Reg	3	1029.9
Res	18	575.8
Total	21	1605.7

且參數估計值等表列如下：

參數	估計值	標準誤	p-value	型一 SS
$\beta_0$	48.63	18.56	0.0173	113654
$\beta_1$	-0.0258	0.172	0.8826	597.56
$\beta_2$	-0.507	0.235	0.0445	1.57
$\beta_3$	1.237	0.337	0.0018	430.76

(一)請計算檢定 $H_0: E[Y]=\beta_0 + \beta_1 X_1 + \beta_3 X_3$  vs.  $H_a: E[Y]=\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3$ 之檢定統計值。(3分)

(二)請計算檢定 $H_0: E[Y]=\beta_0 + \beta_1 X_1$  vs.  $H_a: E[Y]=\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3$ 之檢定統計值。(3分)

次若Y對 $X_3$ 之迴歸的ANOVA表如下：

Source	df	SS
Reg	1	849.9
Res	20	(1)
Total	21	(2)

(三)請計算(1)、(2)及檢定 $H_0: E[Y]=\beta_0 + \beta_3 X_3$  vs.  $H_a: E[Y]=\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3$ 之檢定統計值。(6分)

又若樣本數 $n=22$ ，且因變數Y與自變數 $X_1$ 、 $X_2$ 、 $X_3$ 之複迴歸所得的相關係數矩陣如下：

	Y	$X_1$	$X_2$	$X_3$
Y	1.0	-0.610	0.019	0.728
$X_1$		1.0	-0.082	-0.703
$X_2$			1.0	0.439
$X_3$				1.0

(四)請計算偏相關係數 (partial correlation coefficients)  $r_{Y1.2}$  (3分) 和  $r_{Y2.13}$  (6分) 之值。

五、若已知  $n=23$ ，且算出  $(X'X)^{-1} = \begin{bmatrix} 18.93 & 1.28 & 1.70 & 2.78 \\ 1.28 & 0.33 & 0.47 & 0.34 \\ 1.70 & 0.47 & 1.02 & 0.76 \\ 2.78 & 0.34 & 0.76 & 0.74 \end{bmatrix}$ ， $b = \begin{bmatrix} 6.73 \\ -1.92 \\ 5.94 \\ 3.91 \end{bmatrix}$ ，與

MSE=3.64 之線性複迴歸問題，則：

(一)請以 bonferroni 法計算 $\beta_1$ ， $\beta_2$ 與 $\beta_3$ 之 94% 聯合信賴區間 (simultaneous confidence intervals)。(12分)

(二)請計算 $2\beta_1 + \beta_2 + \beta_3$ 的 96%信賴區間。(8分)

(請接第三頁)

類 科：統計  
科 目：迴歸分析

六、去年一年中，某市區裝有冷氣機的四層樓公寓頂樓住戶，平均月付電費（Y：單位元）及每月平均溫度，但每月平均溫度分為低（10.0-15.0°C）、中（15.1-25.0°C）和高（25.1°C以上）三級，得到紀錄如下表：

月份	1	2	3	4	5	6
月平均溫度	低	低	中	中	中	高
月電費：Y	1088	1107	1178	1288	1455	1981
月份	7	8	9	10	11	12
月平均溫度	高	高	高	中	中	低
月電費：Y	2370	2659	2577	1937	1553	1170

研究人員設法探究月平均溫度與月電費 Y 之間的關係，考慮利用虛擬變數：

$$D_{i1} = \begin{cases} 1, & \text{若月平均溫度=高,} \\ 0, & \text{otherwise,} \end{cases} \quad D_{i2} = \begin{cases} 1, & \text{若月平均溫度=中,} \\ 0, & \text{otherwise,} \end{cases}$$

建構並適配複迴歸模式：

$$Y_i = \beta_0 + \beta_1 D_{i1} + \beta_2 D_{i2} + \varepsilon_i \quad \forall i=1, 2, \dots, 12$$

但在編寫電腦程式時，誤使用以月平均溫度為因子之單因子變異數分析模式，得到下列之變異數分析表：

Source	df	SS
處理	2	3,182,320.7
誤差	9	621,360.2
總和	11	3,803,680.9

(一)請計算檢定  $H_0: \beta_1 = \beta_2 = 0$  vs.  $H_a: \beta_1$  與  $\beta_2$  不全為 0 的檢定統計值。(6分)

(二)請計算  $\beta_0$ ,  $\beta_1$  與  $\beta_2$  之估計值。(9分)