

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。

一、現有一實驗，取得有單一解釋變數與應變數之一組獨立數據 (x_i, y_i) , $i=1, \dots, n$ ，如下：

i	1	2	3	4	5	6	7	8	9	10	11	12
x_i	-3	-2	-1	-1	0	0	0	0	1	1	2	3
y_i	2	6	14	9	15	14	13	17	12	16	16	13

擬以簡單線性迴歸模型 $y_i = \beta_0 + \beta_1 x_i + \epsilon_i$, $i=1, \dots, n$ ，來描述上述應變數與解釋變數間之迴歸關係，其中 β_0, β_1 均為未知，並假設 $\epsilon_i, i=1, \dots, n$ ，為 i.i.d. $N(0, \sigma^2)$, $\sigma^2 > 0$ 亦為未知。已知 $\sum_{i=1}^{12} y_i = 147$, $\sum_{i=1}^{12} y_i^2 = 2021$, $\sum_{i=1}^{12} x_i y_i = 58$ 。

(一)試寫出估計未知參數向量 $\theta = (\beta_0, \beta_1, \sigma^2)^T$ 之概似函數 (likelihood function)。
(10分)

(二)試求 θ 之最大概似估計量 (maximum likelihood estimator)，及 $\beta = (\beta_0, \beta_1)^T$ 最小平方估計量 (least squares estimator)。(10分)

(三)試給出此迴歸分析之變異數分析表，及其 R^2 值。(10分)

(四)試求 $\hat{\beta} = (\hat{\beta}_0, \hat{\beta}_1)^T$ 之 95% 聯合信賴域 (joint confidence region)，並說明 x 變數對解釋 y 之變異是否有幫助？(10分)

($F_{0.025,1,9} = 7.21$, $F_{0.025,1,10} = 6.94$, $F_{0.025,2,10} = 5.48$, $F_{0.05,1,10} = 4.96$, $F_{0.05,1,9} = 5.12$,
 $F_{0.05,2,10} = 4.10$, $z_{0.95} = 1.65$, $z_{0.975} = 1.96$)

二、某一工程師擬探討某機器之有效使用年限(y)與使用頻率(x_1)及其品牌類別間之關係。使用頻率為每週多少小時，品牌類別共有三種，擬建立 y 與二解釋變數間之線性迴歸模型。

(一)由於品牌類別屬於離散型數據，在進行迴歸分析前，三種品牌可以二指標變數 (indicator variables) (x_2, x_3) 表之，如第一種品牌之 $(x_2, x_3) = (0, 0)$ ，試寫出其完整定義 (x_2, x_3) 方式。(8分)

(二)根據 $x = (x_1, x_2, x_3)^T$ 以及線性模型 $y(x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \epsilon$ ，試寫出：(12分)

1. 可比較品牌對有效使用年限是否有差異之假設檢定，i.e. H_0 與 H_1 以 $\beta_i, i=0, \dots, 3$ 表示為何？

2. 可比較兩兩品牌間差異之假設檢定之 H_0 及 H_1 。

3. 模型中不考慮 x_1 與 x_2, x_3 交互作用，其模型有何特性？

(請接背面)

類 科：統計
科 目：迴歸分析

三、在一項有關體脂肪之研究中，收集到 11 位健康女性之資料。擬探討應變數體脂肪 (y)，與解釋變數手臂三頭肌皮摺厚度(x_1)、臀圍(x_2)及臂中圍(x_3)之相關性。根據這 11 筆數據，在 y 服從常態分布假設下，擬建立一 y 與 x_i 間之線性模型

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \epsilon_i$$

首先得到以下 x_1, x_2, x_3 之相關矩陣 (correlation matrix)

$$R = \begin{pmatrix} 1.00 & 0.94 & 0.35 \\ 0.94 & 1.00 & 0.03 \\ 0.35 & 0.03 & 1.00 \end{pmatrix}$$

(一)試求 R^{-1} ，並求各解釋變數之變異膨脹因子 (variance inflation factors, VIF) 之值。
(10 分)

(二)說明 x_i 間是否有共線性 (multicollinearity)？若有共線性，對建立上述模型有何影響？(10 分)

(三)完成下列變異數分析表 (ANOVA table)。(10 分)

來源 (Source)	平方和(SS)	自由度(df)	均方(MS)	F
迴歸 (Regression)	202.59			
x_1	183.92			
$x_2 x_1$	9.42			
$x_3 x_1, x_2$	9.25			
誤差 (Error)	28.79			
總和 (Total)	231.38			

(四)在 $\alpha = 0.05$ 水準下檢定 x_2, x_3 對解釋體脂肪多寡是否有幫助？(10 分)

即檢定 $H_0: \beta_2 = \beta_3 = 0$ v.s. $\beta_2 \neq 0$ or $\beta_3 \neq 0$ 。

($F_{0.025,2,6} = 7.26$, $F_{0.025,2,7} = 6.54$, $F_{0.05,2,6} = 5.14$, $F_{0.05,2,7} = 4.74$, $t_{0.025,2} = 4.30$,
 $t_{0.025,7} = 2.365$, $t_{0.05,2} = 2.92$, $t_{0.05,7} = 1.895$)