

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。

一、兩位科學家試圖研究某一個解釋變數 X 與某一個反應變數 Y 之間的直線關係。數據如下，請根據提供的數據回答以下問題：

X	10	9	8	7	6	5	4	3
Y	45	20	34	58	70	57	55	44

- (一)兩位科學家委託一位統計學家協助計算他們有興趣的直線關係。為了徹底了解 X 與 Y 之間的直線關係，統計學家建議先試試「 $Y=A$ 」，也就是說， X 與 Y 之間沒關係。請問根據以上數據， A 的估計值等於多少？(5分)
- (二)接下來，統計學家假設的直線關係是「 $Y=A+B \times X$ 」，其中「 $B \times X$ 」意味著「『直線斜率』乘以『 X 』」。也就是說，統計學家第二次研究具有「直線截距」與「直線斜率」的直線關係。請問根據以上數據，直線斜率 B 的估計值等於多少？(10分)
- (三)最後，兩位科學家在收到上述分析報告之後，決定更動直線關係為「 $Y=B \times X$ 」。也就是說，兩位科學家第三次研究的議題是「無直線截距」的直線關係。請再一次根據以上數據，回答直線斜率 B 的估計值等於多少？(10分)

二、兩位科學家試圖研究某一個解釋變數 X 與某一個反應變數 Y 之間的直線關係。數據如下：

X	10	9	8	7	6	5	4	3
Y	45	20	34	58	70	57	55	44

這兩位科學家決定取得數據在「 $Y=A+B \times X+\epsilon$ ， ϵ 為誤差項」這一項假設下的變異數分析表。請問：

- (一)總平方和 (total sum of squares) 等於多少？(10分)
- (二)迴歸平方和 (regression sum of squares) 等於多少？(10分)
- (三)誤差平方和 (residual sum of squares) 等於多少？(5分)

三、假設一組 38 個樣本、三個變數的數據集，其中三個變數分別是一個反應變數、一個解釋變數、加上一個源自前述解釋變數的「兩水準虛擬變數 (dummy variable)」。也就是說，這一個虛擬變數只會出現兩種數字，假設不是「0」就是「1」。請回答以下問題：

- (一)如果數據科學家提出一個這樣的複迴歸模型「反應變數 = $A+B \times$ 解釋變數 + $C \times$ 虛擬變數 + $D \times$ 虛擬變數 \times 解釋變數」。請寫下「虛擬變數等於 1」的數學方程式？(5分)
- (二)請說明(一)的迴歸係數 A, B, C, D 的數值以描述在「『虛擬變數等於 0』與『虛擬變數等於 1』」下的迴歸直線是同一條直線。」(10分)
- (三)請說明(一)的迴歸係數 A, B, C, D 的數值以描述在「『虛擬變數等於 0』與『虛擬變數等於 1』」下的迴歸直線是兩平行直線。」(10分)

(請接背面)

等 別：三等考試
類 科：統計
科 目：迴歸分析

四、統計學家為一位有興趣採用「迴歸分析」實踐「檢定三個處理平均數是否相等」的科學家寫下這樣的設計矩陣 X ，

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

請回答以下問題：

(一)請計算 H 矩陣 (hat matrix)。(5 分)

(二)請用(一)結果證明這時候 $(I-H)(I-H) = (I-H)$ ，其中 I 是一個對角線上是 1、其他都是 0 的方陣。(10 分)

(三)根據迴歸分析理論的計算，統計學家發現檢定「三個處理平均數相等」這一項虛無假設的檢定統計量是兩項「迴歸平方和」之間的「差」。請問是那兩項迴歸平方和？回答問題時，請根據題意自行定義相關的符號。(10 分)