應用統計 試題 （限用答案本作答）

可使用計算機

一、以下資料是來自一實驗，研究餵食四種食物之老鼠的體重增加量，這四種食物是依蛋白質含量(高與低)與蛋白質來源(牛肉與穀物)給予分類。此實驗是完全隨機設計，每一處理有十隻老鼠。

老鼠體重增加量(公克)

食物種類			
低蛋白質牛肉	高蛋白質牛肉	低蛋白質穀物	高蛋白質穀物
90	73	107	98
76	102	95	74
90	118	97	56
64	104	80	111
86	81	98	95
51	107	74	88
72	100	74	82
90	87	67	77
95	117	89	86
78	111	58	92

資料摘要

組別	個數	總和	平均	變異數	SS
低蛋白質牛肉	10	792	79.2	192.8444	1735.6
高蛋白質牛肉	10	1000	100	229.1111	2062.0
低蛋白質穀物	10	839	83.9	246.7667	2220.9
高蛋白質穀物	10	859	85.9	225.6556	2030.9
全體樣本	40	3490	87.25	268.0385	10453.5

(a) 如何執行此項實驗，此實驗方稱為完全隨機設計？(5 分)

(b) 利用單因子變異數分析，研究食物種類對於老鼠體重平均增加量之影響。(顯著水準取 5%，$F_{0.05;3,36} = 2.866$) (5 分)

(c) 試估計餵食高蛋白質牛肉之老鼠之體重平均增加量的 95%信賴區間。($t_{0.025;36} = 2.028$) (5 分)

(d) 試以 Tukey 的方法進行多重比較。(聯合顯著水準取 5%，critical value of the studentized range: $Q_{0.05;4,36} = 3.809$) (5 分)

(e) 設在實驗之前研究者就計劃要研究蛋白質含量、蛋白質來源對於老鼠體重平均增加量之影響；試設計出一組合乎研究目的之直交對比(三個對比)，並解釋每一對比之意義。(5 分)

(f) 試進行上述直交對比之變異數分析並檢定各對比之顯著性。(顯著水準取 5%，$F_{0.05;1,36} = 4.113$) (5 分)

二、詳細探討下列觀念問題：(每小題 5 分)

(a) 殘差分析對於建立迴歸模型之重要性；

(b) 因素轉軸之目的；

(c) 多元尺度分析之意義與目的。

本試題兩面印刷

應用統計 試題 （限用答案本作答）

三、考慮下列實驗資料：

X	-5	-4	-3	-2	-1	0	1	2	3	4	5
Y	-8	-4	-5	-2	1	-1	0	4	5	4	9

(a) 假設下列直線迴歸模型：

$$Y = \beta_0 + \beta_1 X + \varepsilon$$

用最小平方法估計 β_0、β_1 並寫出所估計的迴歸方程式。(5 分)

(b) 建立變異數分析表，並以 5%顯著水準檢定此直線迴歸模型之顯著性。($F_{0.05;1,9} = 5.12$)(5 分)

(c) 估計 β_0 之 95% CI。($t_{0.025;9} = 2.262$)(5 分)

(d) 試估計 $(E[Y \mid X=3] - E[Y \mid X=-2])$ 之 95% CI。(5 分)

(e) 設模型 $Y = \alpha_1 X + \varepsilon$ 為正確的模型。

(1) 求出 α_1 之估計式。(5 分)

(2) 試證 β_1 之估計式之期望值為 α_1。(5 分)

(3) 試比較 α_1 之估計式之變異數與 β_1 之估計式之變異數。(5 分)

四、考慮以下兩群組的區別分析(discriminant analysis)，有兩個自變數 X_1 與 X_2，六十筆觀察值自母體中隨機抽取，得到以下摘要統計量：

$$n_1 = 20, \quad n_2 = 40$$

$$\bar{\mathbf{X}}'_{(1)} = (-0.5, 0.0), \quad \bar{\mathbf{X}}'_{(2)} = (0.5, 0.0)$$

$$\mathbf{C}_w = \begin{bmatrix} 0.80 & 0.50 \\ 0.50 & 1.00 \end{bmatrix}$$ ($\mathbf{C}_w$為合併樣本共變異數矩陣)

(a) 試建立 Fisher 線性區別規則。(5 分)

(b) 試依 Fisher 線性區別規則將新觀察值 $\mathbf{X}' = (0, 0.5)$ 給予歸類。(5 分)

五、令 $(Y_1, X_1), (Y_2, X_2), \cdots, (Y_n, X_n)$ 為一隨機樣本，其中 Y_i 的可能值為 0 或 1，而 X_i 為計量變數。考慮下列邏輯模式(logit model)：

$$\ln\left(\frac{P(Y_i = 1 \mid X_i)}{P(Y_i = 0 \mid X_i)}\right) = \alpha + \beta X_i, \quad i = 1, 2, \cdots, n$$

(a) 試求出 $\pi(X_i) = P(Y_i = 1 \mid X_i)$；(5 分)

(b) 在 $(X_1, \cdots, X_n)$ 已知條件下，$(Y_1, \cdots, Y_n)$ 之概似函數(likelihood function)為

$$L(\alpha, \beta) = \prod_{i=1}^{n} \pi(X_i)^{Y_i} [1 - \pi(X_i)]^{1-Y_i}$$

試證 (α, β) 之最概似估計(MLE)滿足下列聯立方程式：(5 分)

$$\sum_{i=1}^{n} [Y_i - \pi(X_i)] = 0$$

$$\sum_{i=1}^{n} X_i [Y_i - \pi(X_i)] = 0$$

試題完

本試題兩面印刷