

※ 注意：請於試卷內之「非選擇題作答區」依序作答，並應註明作答之大題及小題題號。

1. 以 A、B 兩種養液配方處理下的水稻幼苗根系，每種處理分別隨機抽取 10 株植株，並測量每株的主根長，得到以下資料 (單位:公分)。

養液配方 A:

1.7 1.9 2.0 2.0 2.1 2.1 2.2 2.2 2.4 2.7

養液配方 B:

1.8 2.0 2.2 2.2 2.2 2.4 2.6 2.7 2.9 3.1

- (1) 計算兩組資料的樣品中位數 (sample median) 及樣品四分位距 (sample inter-quartile range, IQR)。(10 分)
- (2) 計算兩組資料的樣品平均值 (sample mean) 與樣品標準差 (sample standard deviation)。(10 分)
- (3) 根據前兩題的計算結果比較兩組資料的中心位置 (location) 及分散程度 (dispersion) 是否有差異。(10 分)
2. 以人工接種病毒感染的方式研究某作物特定品系的抗病程度，並觀察該品系 25 株植株罹病情況後，分別以 0 到 4 的數字代表每株抗病程度 (0: 最抗病; 4: 最不抗病)，將 25 株的抗病程度在 0 到 4 等級所佔的比例整理如下表：

抗病程度	0	1	2	3	4	總和
株數佔比	12%	36%	28%	12%	12%	100%

- (1) 依據上述資料，計算該品系 25 株抗病程度之樣品平均值與樣品標準差。(10 分)
- (2) 假設國際標準定義：所有植株觀測之抗病程度小於或等於 1 者若佔 90% 以上，則該品系稱為高度抗病 (highly resistant) 品系。那麼上述研究觀察 25 株植株，且已知該品系為高度抗病品系(假設該品系抗病程度小於或等於 1 者佔所有品系植株中的 90%)，則實際觀察到 25 株植株中抗病程度小於或等於 1 者小於 90% (22.5 株)的機率為何？(10 分)

提示：二項分布的機率分布密度函數為

$$P_r(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 1, 2, \dots, n$$

3. 是非題 (每題答對得 2 分，未答得 0 分，答錯倒扣 1 分，最低總得分為 0 分) (40 分)

- (1) 卡方獨立檢定統計量是檢定 H_0 : 兩變數不獨立 H_1 : 兩變數獨立。
- (2) 利用卡方適合度檢定統計量檢定隨機變數 X 是否服從常態分佈。若將樣本分成 10 組，檢定統計量服從卡方自由度

見背面

為 9 的分佈。

- (3) 根據相同的樣本，進行兩個假設檢定問題：(A) 顯著水準 α ，檢定 $H_0: \mu = 2$ vs. $H_A: \mu = 4$; (B) 顯著水準 α ，檢定 $H_0: \mu = 2$ vs. $H_A: \mu = 5$ 。(A)和(B)的檢定方法會不一樣。
- (4) 檢定 10 個母體均值是否有差異: $H_0: \mu_1 = \mu_2 = \dots = \mu_{10}$ vs $H_A: \text{not } H_0$ 。假設透過 t 檢定，得出其中兩母體的均值有顯著差異的結論，則 one-way ANOVA 的分析也會得到拒絕 H_0 的結論。
- (5) 在簡單線性迴歸中，利用最小平方法求解截距項和斜率的估計式時，不需任何分布假設。
- (6) 簡單線性迴歸中，採用最小平方法估計參數，若解釋變數的單位改變（如：公斤變公克），則反應變數的估計值 $(\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n)$ 也會跟著改變。
- (7) 簡單線性迴歸中，採用最小平方法估計參數，若反應變數的單位改變（如：公尺變公分），則決定係數(R^2)也會跟著改變。
- (8) 簡單線性迴歸中，採用最小平方法估計參數，若新的解釋變數為原解釋變數取指數(e^x)，則估計值 $(\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n)$ 也會跟著改變。
- (9) 簡單線性迴歸中，採用最小平方法估計參數，若新的解釋變數為原解釋變數加上某固定值後(ex: $X + 10$)，則估計值 $(\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n)$ 也會跟著改變。
- (10) 簡單線性迴歸中，如果殘差項分佈的範圍會隨解釋變數的大小而有差異，表示誤差項的同質變方的模型假設是錯的。
- (11) 簡單線性迴歸： $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$ ，當 β_1 估計值越接近 1，表示 Y 和 X 的線性關係越明顯。
- (12) 簡單線性迴歸： $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$ ，當 β_1 估計值越小，表示有越低的 R^2 (決定係數) 值。
- (13) 簡單線性迴歸： $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$ ，當 β_1 估計值大於零，不代表 X 和 Y 有正相關。
- 問題(14)~(15): 假設一隨機樣本 $(X_1, X_2, \dots, X_{10})$ ，根據顯著水準 $\alpha = 0.05$ 檢定 $H_0: E[X] = 0$ vs. $H_A: E[X] > 0$ ，分別採用兩種統計量(A) $\frac{1}{10} \sum_{i=1}^{10} X_i$ (B) $\frac{X_1 + X_2 + X_3 + X_4}{4}$ 進行檢定。
- (14) 統計量(A)的檢定，其發生型一誤差的機率較低。
- (15) 統計量(A)的檢定，其發生型二誤差的機率較高。

問題(16)~(17): 利用 one-way ANOVA 檢定不同肥料的使用對稻米的產量是有差異。假設 μ_1, μ_2 和 μ_3 分別是使用三種肥料的稻米產量的期望值。

(16) 其假設檢定的描述是 $H_0: \mu_1 = \mu_2 = \mu_3$ vs. $H_1: \mu_1 \neq \mu_2 \neq \mu_3$ 。

(17) 利用檢定統計量 $F = \frac{\frac{SS_t}{df_1}}{\frac{SS_E}{df_2}}$ 檢定上述問題，當 F 夠小或夠大時，拒絕 H_0 。

問題(18)~(20): 欲檢定台灣成人的教育水平(分成四個級別)與政黨傾向(三個政黨)是否有關係。

(18) 當教育水準與政黨傾向獨立，是指教育水準的高低不會影響政黨的傾向，也等同於政黨傾向不受教育水平的影響。

(19) 當教育水準與政黨傾向獨立，表示三個政黨的支持者，其教育水平的分布都一樣，也表示四個教育水平的人，其政黨傾向的分布都一樣。

(20) 卡方獨立檢定統計量的服從卡方自由度是 7 的分布。

4. 填空題 (1 題 1 分，答錯不扣分)(10 分)

學生甲利用簡單線性回歸分析兩隨機變數 Y 和 X 的線性關係。其分析結果如下表:

Call: lm(formula = Y ~ X)				
Coefficients:				
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	9.94	0.19	51.58	$<2 * 10^{-16}$
X	1.65	0.34	4.90	$6.1 * 10^{-6}$
Residual standard error: 0.96 on 23 degrees of freedom				

其變異數分析表格(ANOVA table)如下表:

Analysis of Variance Table				
Response: Y				
	Df	Sum Sq	F value	Pr(>F)
X	(1)	21.96	(4)	(5)
Residuals.	(2)	(3)		

(1) _____ (2) _____ (3) _____ (4) _____ (5) _____

根據相同樣本資料，學生乙先將資料進行標準化，再進行簡單回歸分析，其分析結果如下表

Call: lm(formula = Y ~ X)				
Coefficients:				
	Estimate	Std. Error.	t value	Pr(> t)
(Intercept).	$2.28 * 10^{-16}$	0.143	0.000	1
X	0.714	0.146	(6)	(7)

見背面

題號： 213

國立臺灣大學 113 學年度碩士班招生考試試題

科目： 生物統計學

題號： 213

節次： 2

共 4 頁之第 4 頁

其變異數分析表格(ANOVA table)如下：

Analysis of Variance Table				
Response: Y				
	Df	Sum Sq	F value	Pr(>F)
X	1	12.346	(9)	(10)
Residuals.	23	(8)		

(6) _____ (7) _____ (8) _____ (9) _____ (10) _____

試題隨卷繳回