

※ 注意：請於試卷內之「非選擇題作答區」作答，並應註明作答之題號。

1. (5分) 欲檢定兩獨立群體之間的平均數是否有差異，可以使用以下哪些方法？

- (1) two sample t-test
- (2) hypothesis test in simple linear regression
- (3) paired t-test
- (4) McNemar test

- (A) 只可用 (1)
- (B) 只可用 (1) (2)
- (C) 只可用 (1) (3)
- (D) 只可用 (1) (4)
- (E) 只可用 (2) (4)

2. (5分) 一個假說檢定 (Hypothesis testing) 的 P 值為 0.0001，下列何者正確

- (A) 實驗 10000 次平均下來會有一次拒絕虛無假設
- (B) 實驗 10000 次平均下來會有一次接受虛無假設
- (C) 虛無假設為真
- (D) 對立假設為真
- (E) 以上皆不正確

3. (5分) 下列關於分佈的敘述：

- (1) t 分佈在自由度為 30 的時候就等於常態分布
- (2) 布阿松 (Poisson) 分佈的數值是依連續正數，常用來描述事件發生的機率
- (3) 常態分佈是以中位數為中心而左右對稱的分布

請選出最適當的選項：

- (A) 只有(1)正確
- (B) 只有(2)正確
- (C) 只有(3)正確
- (D) (1)(2)(3)都正確
- (E) (1)(2)(3)都不正確

4. (10分) 考慮一研究，其中X代表某新藥的劑量，Y代表病人是否有康復。

- (1) 吾人希望用邏輯斯迴歸 (logistic regression) 來分析這筆資料。請詳細寫出此研究的邏輯斯迴歸模型。
- (2) 請詳細解釋在此邏輯斯迴歸模型中，X的係數的意義。

見背面

5. 是非簡答題 (本大題共 25 分, 每小題 5 分):

下列敘述, 對的請打○, 錯的請打×, 若敘述為錯者請說明理由 (理由佔該小題 4 分):

- (1) 「樣本數愈大, 資料將愈呈現常態分布」。
- (2) 「無母數分析方法不需假設資料呈常態分布, 可用於分析有序尺度 (ordinal scale) 的資料」。
- (3) 「兩個連續型變項的皮爾森相關係數 (Pearson's correlation coefficient) 愈大, 代表二者的因果關係愈強」。
- (4) 「變異係數 (coefficient of variation) 與原始資料的測量單位無關, 所以可用來比較單位不同之變項的變異程度」。
- (5) 「以平均數為中心, 加減兩倍標準差的範圍內, 約含有 95% 的資料」。

問答題:

6. 某醫院於一社區進行糞便潛血篩檢, 下表顯示該社區 50 歲以上居民的篩檢結果, 與實際罹患大腸癌之關係, 請以代號說明糞便潛血篩檢的陽性預測值 (positive predictive value) 及陰性預測值 (negative predictive value) 指標, 及其代表之意義。並說明推廣糞便潛血篩檢須考慮的因素有哪些? (12%)

		大腸癌	
		有病	沒病
糞便潛血 篩檢	陽性	A	B
	陰性	C	D

7. 某學者想了解吸菸與心血管等慢性疾病發生的關係, 對一群有吸菸習慣的民眾進行心血管疾病篩檢, 並同時對一群沒有吸菸的民眾進行心血管疾病篩檢做比較, 結果發現吸菸者得到心血管疾病較多。請問在什麼情況下吸菸者得到心血管疾病是選擇偏差 (selection bias) 的結果? 如何預防以上之選擇偏差情形發生? 在什麼情況下吸菸者得到心血管疾病是干擾偏差 (confounding bias) 的結果? 以上干擾偏差情形該如何處理? (13%)
8. 生態模型 (Ecological model) 從 1979 年開始出現, 由於所謂的生態謬誤 (Ecological Fallacy) 導致多數人對此種模型的運用有所遲疑。請解釋何謂生態模型? (4%) 何謂 Ecological Fallacy? (4%) 此種模型可以運用於何種研究, 請舉例。 (4%)
9. 系統誤差 (systematic error) 亦即偏差 (bias), 請舉出系統誤差的三種來源 (9%)。再者隨機誤差 (random error) 可以用什麼統計量測值來表示 (4%)。

試題隨卷繳回