

※注意：請於答案卷上依序作答，並註明作答之題號。

1. 假設 X 服從標準常態分配 $N(0,1)$ 。請問以下敘述何者錯誤？(5分)
 - (A) X 落在 $[-2,2]$ 之間的機率約為 0.95
 - (B) $X+5$ 會服從 $N(5,1)$
 - (C) X 的期望值會比中位數來的大一點
 - (D) X^2 會服從自由度為 1 的卡方分配
 - (E) $X > 1$ 與 $X < -1$ 的機率相等

2. 下列關於相關係數 (correlation coefficient) 的敘述，何者錯誤？(5分)
 - (A) 其數值介於 -1 和 1 之間
 - (B) 改變某個變數的測量單位不會影響相關係數的數值
 - (C) 對每筆資料某變數的數值都加上 5，則相關係數會變大
 - (D) 兩個獨立變數的相關係數必定為 0
 - (E) 用來測量兩變數之間線性相關的程度

3. 下列敘述何者正確？(5分)
 - (A) 型一誤差和型二誤差的和是固定的
 - (B) 型一誤差越大，則型二誤差會越小
 - (C) 當 p-value 小於型一誤差，則虛無假說描述了一個不可能發生的事件
 - (D) 當 p-value 小於型一誤差，則對立假說描述了一個事實
 - (E) 當 p-value 小於型二誤差，則虛無假說描述了一個不可能發生的事件

4. 考慮一研究， X_1 代表某新藥的劑量 (1:標準劑量, 2:兩倍劑量, 3:placebo)， X_2 代表病人居住的地區 (1:北, 2:中, 3:南)， Y 代表治療的成效。假設 X_1 和 X_2 之間沒有交互作用，吾人欲建立迴歸模型以分析之。(10分)
 - (1) 寫出迴歸模型 (4分)
 - (2) 根據你的模型，寫出<新藥沒有療效>的虛無假說 (3分)
 - (3) 根據你的模型，寫出<居住地區對療效沒有影響>的虛無假說 (3分)

見背面

是非簡答題 (本單元共 25 分，每題 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 (理由佔該題 4 分)：

5. 「統計檢定的結果若 p 值 (p-value) > 0.05，則代表虛無假設為真」。
6. 「變異數分析 (Analysis of variance) 的目的是為了要檢定獨立的三組或三組以上連續型資料之母體變異數是否相等」。
7. 「皮爾森相關係數 (Pearson's correlation coefficient) 可衡量兩個連續型變項之間的線性與曲線相關程度，其範圍介於 -1 至 1 之間」。
8. 「若藉由統計檢定發現簡單線性迴歸模式 (simple linear regression model) 裡的迴歸係數斜率項顯著異於 0，意指解釋變數和反應變數之間存在著因果關係」。
9. 「在生物統計學裡，若反應變項為個案的病況：分有病與沒病兩種，研究者可以使用羅吉斯迴歸 (logistic regression) 予以分析之」。

問答題：

10. 兒童肥胖盛行率日益上升，若對於兒童肥胖發生的原因 (包含環境因子、基因標記等等) 進行研究，可能的研究設計 (研究族群、暴露測量、結果測量、資料分析) 有哪些？請根據 Hill's criteria 說明如何判定暴露指標與疾病指標的因果關係？並請說明該注意的偏差 (bias) 有哪些。(15 分)
11. 某人欲探討「吃魚是否較不容易得到心血管疾病」，對一群吃魚較多的人進行心血管疾病篩檢，並同時對一群吃魚較少的人進行心血管疾病篩檢做比較，結果發現多吃魚可以預防心血管疾病。請問在什麼情況下這個結論是選擇偏差 (selection bias) 的結果？在什麼情況下這個結論是訊息偏差 (information bias) 的結果？在什麼情況下這個結論是干擾偏差 (confounding bias) 的結果？(10 分)
12. 進行流行病學研究一開始的課題即是找出研究主題，研究主題的“重要性”以及過去“研究之不足 (research gap)”常需要細心的探索，請問要如何做才能找出研究的重要性及 research gap (4 分)？如果你即將要進行一項流行病學研究，你會探討什麼主題？其重要性以及 research gap 分別為何 (6 分)？
13. 混淆因子 (confounding factor) 在流行病學研究中需要加以控制，才能獲得較正確的估計結果，請列舉出三種控制混淆因子的方法 (6 分)。針對每一種方法，請舉出一個“實例”來說明如何控制混淆因子 (9 分)。