

是非題(每題 2 分)

- 1 若隨機抽選 500 位台大學生測量其每分鐘心跳次數，結果可以利用盒鬚圖(boxplot)呈現，並檢查是否有離群值(outlier)。
- 2 相對風險(relative risk)是兩個機率的比值，這兩個機率的和為 1。
- 3 勝算比值(odds ratio)越接近 0 表示在暴露情況下罹病風險越低。
- 4 敏感度(sensitivity)與特異度(specificity)的和為 1。
- 5 每個人會不會罹患高血壓的事件可以算是二項式分配(Binomial distribution)。
- 6 任何兩個獨立的伯努利分配(Bernoulli distribution)相加必為二項式分配。
- 7 常態分配與 t 分配在三個標準差之內的機率大約都是 99%。
- 8 常態分配與 t 分配都是對稱於 0 的分配。
- 9 台灣各縣市兒童罹患白血病的人數可以使用布阿松(Poisson)分布來描述。
- 10 描述是否飲用含砷井水與是否罹患烏腳病的關係時，可以使用勝算比值(odds ratio)來描述，也可以使用勝算比值來檢定。
- 11 在簡單迴歸分析中，判定係數(R^2)越大，則解釋變數和反應變數間的相關係數越大。
- 12 在簡單迴歸分析中，解釋變數的變異數越大，則迴歸係數估計式的變異數越大。
- 13 在迴歸分析中，解釋變數越多，則判定係數(R^2)越大。
- 14 在迴歸分析中，判定係數(R^2)越大，則校正的判定係數(adjusted- R^2)越大。
- 15 在迴歸分析中，欲檢定某解釋變數是否有顯著影響。若樣本數夠大時，則可以用標準常態分配來計算 p-value。
- 16 變異數分析(ANOVA)可以用來檢定兩個獨立母體的平均數是否相等。

選擇題(每題 3 分)

1. 以下敘述何者錯誤?
 - (A) 已知肺癌不是傳染性疾病，則每個人會不會得肺癌是獨立事件。
 - (B) 假如每一胎生女嬰的機率都是 p，生男嬰的機率都是 1-p，則每一胎生出女嬰與生出男嬰互為獨立事件。
 - (C) 已知就讀於 T 小學三年甲班的小明得了傳染性強的流行性感冒，則坐在小明旁邊的同班同學阿強罹患同一個流感的機率比一般人高。
 - (D) 假如陳先生每次買大樂透時都選擇同一組號碼，那麼他每次中獎的機率都一樣。
2. 以下有關標準差(standard deviation)與標準誤(standard error)的敘述何者錯誤?
 - (A) 標準差常用來描述一個分布的散布程度。
 - (B) 標準誤可用來描述樣本平均值的抽樣誤差。
 - (C) 標準誤的平方可用來估計樣本來源之原母體的變異數。
 - (D) 標準誤與樣本數有關。
3. 以下有關常態分佈(normal distribution)的敘述何者錯誤?
 - (A) 常態分佈是最常見的分佈。
 - (B) 常態分佈是對稱的分佈。
 - (C) 各種常態分佈彼此可以互相轉換(transformation)。
 - (D) 常態分佈可以用來估計二項式分布(Binomial distribution)。

4. 以下有關中央極限定理(central limit theorem)的敘述何者錯誤？
- (A)中央極限定理是用來描述樣本平均值在大樣本情況下的抽樣分布。
 - (B)中央極限定理說明了樣本平均值在大樣本情況下的分布會逼近常態分布。
 - (C)使用中央極限定理時不需要假設樣本來源之原母體的分布為常態。
 - (D)中央極限定理證明了樣本平均值在大樣本情況下的期望值與變異數跟樣本來源之原母體的期望值與變異數一樣。
5. 以下有關ROC曲線(receiver operating characteristic curve)的敘述何者正確？
- (A)ROC曲線離對角線($X=Y$)越遠越好。
 - (B)ROC曲線上點離座標原點(0, 0)越遠越好。
 - (C)ROC曲線上的每一點的X與Y座標值，合為1。
 - (D)ROC曲線下的面積越大越好。

計算問答題

- 1 某研究希望探討某新藥 X_1 不同劑量($X_1 = 1:0\text{ml}\sim 10\text{ml}$, $X_1 = 2:11\text{ml}\sim 20\text{ml}$, $X_1 = 3:21\text{ml}\sim 30\text{ml}$, $X_1 = 4:31\text{ml}\sim 40\text{ml}$)對病毒含量 Y 是否有顯著影響。研究者另外蒐集可能的兩個影響因子，包括性別 X_2 ($X_2 = 1$:女性)以及白血球含量 X_3 。過去的研究發現白血球含量和劑量之間存在著交互作用。請利用迴歸分析(假設誤差項服從常態分配 $N(0, \sigma^2)$)回答以下問題。
- (1) 根據以上描述，請寫出 Y 對 X_1, X_2, X_3 以及 (X_1, X_3) 交互作用之迴歸模型。(6分)
 - (2) 根據(1)中的模型參數，請寫出檢定所有的解釋變數對 Y 都沒有影響的虛無假說。若利用F-test，請寫出檢定統計量的虛無分配。(6分)
 - (3) 根據(1)中的模型參數，請寫出檢定劑量多寡對病毒含量無影響的虛無假說。若利用partial F-test，請寫出檢定統計量的虛無分配。(6分)
- 2 某研究者於某社區隨機抽樣收案 30 位居民，估計出該社區之吸菸盛行率。試問如何量化這個估計值的精確度 (precision)? 並請說明你將如何計算? 另外，若該研究者不滿意這樣的精確度，你會如何建議? 你的這個建議的依據是什麼? (11分)
- 3 試舉一個生物醫學實例，說明二組獨立樣本平均值如何檢定其平均值是否有差異，包括如何訂定虛無假說(null hypothesis)及對立假說(alternative hypothesis)，如何計算統計檢定值，如何決定自由度，以及如何求得p值等等。這個檢定依賴什麼假設(assumption)? (12分)
- 4 某研究者收案 1000 位男性及 1000 位女性，欲檢定血型 (分成 A, B, AB, O 等四型) 的分佈是否有性別差異。你將採用什麼統計檢定方法檢定之? 並請說明你將如何進行之? 包括如何訂定虛無假說及對立假說，如何計算統計檢定值，如何決定自由度，以及如何求得p值等等。這個檢定方法是否依賴大樣本理論? 如何判斷該筆資料是否適用? (12分)

試題隨卷繳回