

是非題 (每題 2 分)

1. 如果要比較臺灣大學 2011 年到 2015 年間，這 5 年內入學的研究生的男女分布可以利用 5 個盒鬚圖(boxplot)來比較。
2. 布阿松分配(Poisson distribution)比較適合描述次數資料(count data)的分布。
3. 若想檢定「體重過重的人是否有較高的比例罹患高血壓」，可以利用勝算比(odds ratio, OR)或卡方(Chi-square)統計量來檢定。
4. 變異數分析(Analysis of variance, ANOVA)的檢定必須有變異數相同的假設，但兩個獨立樣本的 t 檢定不一定需要變異數相同的假設。
5. 如果要比較臺灣大學 2014 年跟 2015 年的研究生入學考試之報考人數是否達統計顯著的不同，可以檢定兩個母體的期望值是否相等。
6. 為了要了解就醫民眾與醫師的互動關係，某研究進行深度訪談，訪問了 9 位就醫民眾。如果要檢定「民眾就醫前後對醫師的滿意程度(0-100%)是否相同」，需使用兩個成對樣本的 McNemar's 檢定。
7. 兩個常態分配之間可以互相轉換，但必須是不同期望值、相同變異數的情況下。
8. 卡方(Chi-square)檢定可以檢定「兩個變數之間的相關程度」是否達統計顯著。
9. 若隨機抽選 100 位臺灣大學在學學生並量測其血壓，所得出的血壓分布可以利用枝葉圖(stem-and-leaf plot)來呈現，其結果也可以利用長條圖來呈現。
10. 特異度(specificity)可以用來評估流感檢驗試劑錯誤地將感染的人篩檢出感染的機率。

單選題 (每題 3 分) 請於試卷內之「選擇題作答區」依序作答。

1. 為了要比較北部地區與南部地區民眾求醫行為的異同，某研究隨機抽選具代表性的 100 位北部地區民眾與 100 位南部地區民眾，進行問卷填寫、訪談、並連結其健保資料。可得知前一年中，這些民眾每個人的就醫次數，並繪製長條圖(histogram)。請問以下哪一個敘述**最不合適**?
 - (A)利用北部地區民眾的就醫次數樣本可以推論北部地區民眾就醫次數之母體分布的期望值與變異數。
 - (B)如果南部地區民眾樣本的就醫次數的長條圖，比北部地區民眾的長條圖還向右偏斜(skewed to the right)，表示相較於北部地區而言，有比較多的南部民眾就醫次數比較高。
 - (C)這些長條圖的偏斜(skewed)趨勢在盒鬚圖(boxplot)也可以呈現出來，但盒鬚圖中若有離群值(outlier)不一定能在長條圖中看出來。
 - (D)如果南部地區民眾樣本的就醫次數的長條圖向右偏斜，表示南部地區民眾的就醫次數的平均值，會小於母體的中位數。

見背面

2. (延續上題) 如果南部地區民眾樣本的就醫次數平均為 12.31, 樣本變異數為 12.25, 請問以下哪一個最接近全體南部地區民眾的平均就醫次數之 95% 信賴區間?
- (A)(5.45, 19.17)。 (C)(11.62, 13.00)。
(B)(12.24, 12.38)。 (D)(12.07, 12.55)。
3. (延續上題) 如果北部地區民眾樣本的就醫次數長條圖也有偏斜趨勢, 樣本平均值為 12 次, 樣本變異數為 16, 請問以下哪一個敘述最合適?
- (A) 比起北部地區民眾的就醫次數分布, 南部地區民眾的就醫次數分布比較適合使用布阿松分布(Poisson distribution)來描述。
(B) 如果檢定「兩個地區的民眾平均就醫次數是否相同」, 會得到統計顯著的結論。
(C) 不適合使用變異數分析(Analysis of variance, ANOVA)來檢定「兩個地區的民眾平均就醫次數是否相同」的原因是因為資料中只包含了兩組獨立樣本。
(D) 分別進行兩個地區民眾就醫次數的盒鬚圖, 其中每個盒鬚圖可得到第一個四分位(first quartile)及第三個四分位(third quartile)的區間, 比較這兩個區間是否重疊, 等同於檢定「兩個地區的民眾平均就醫次數是否相同」。
4. 某研究探討民眾對「健保財務的樂觀程度」與民眾「每月收入」的關係, 隨機抽選訪談了 70 位民眾得到以下數據:

	每月收入低於 3.5 萬	每月收入高於 5 萬	
高度樂觀	20	10	30
低度樂觀	25	15	40
總和	45	25	70

進行卡方檢定之後, 得到卡方檢定統計量為 0.01, p 值 = 0.9 > 0.05, 請問以下何者不正確?

- (A) 民眾對「健保財務的樂觀程度」與其「每月收入」沒有統計顯著的關係。
(B) 因為 p 值比 0.05 大很多, 可以得知這個結論非常不顯著。
(C) 這個卡方檢定不能推論每月收入介於 3.5 萬及 5 萬之間的民眾對健保財務的樂觀程度。
(D) 由資料可推論, 在顯著水準 0.05 的條件下, 全體民眾當中對健保財務的樂觀程度較低的比例在統計上沒有顯著不同於 0.5。
5. (延續上題) 利用同樣 70 位民眾的資料及其性別, 若探討樂觀程度與性別的關係, 請問以下何者不正確?
- (A) 可以利用卡方檢定來探討樂觀程度高或低與性別是否有關。
(B) 若利用卡方檢定來探討樂觀程度高或低與性別是否有關, 其結論不能進一步推論男性樂觀程度高的比例是否顯著高於女性。
(C) 若利用卡方檢定來探討樂觀程度高或低與性別是否有關, 其自由度為 1。
(D) 可以利用兩個比例的 Z 檢定(two-proportion Z test)來探討樂觀程度高或低與性別是否有關。

單選題 (每題 5 分)

6. 以下有四個敘述

(甲) $P=0.000001$ 代表對立假設成立

(乙) 虛無假設不成立代表統計機率值低於顯著水準

(丙) 若減少試驗的型一誤差，則型二誤差會跟著減少

(丁) 一試驗的檢定力為 0.8，代表在兩組具有顯著差異的資料中進行 100 次實驗理論上能獲得 80 次統計顯著結果

請回答下列何項說明最為正確？

(A) 甲乙選項正確

(B) 甲丙選項正確

(C) 乙丁選項正確

(D) 丙丁選項正確

7. 請問下列何者”不是”無母數統計的特性？

(A) 以中位數代表資料平均數

(B) 常就大小及出現先後順序排列的資料進行分析

(C) 適合用於小樣本

(D) 適合處理交互作用

8. 請問下列何者”不是” Bonferroni correction 的特性？(假設試驗進行次數為 N 次)(A) 針對某次試驗得到的統計機率值 p ，校正後的統計機率值為 Np

(B) 透過改變試驗進行次數減少偽陽性

(C) 主要目標為減少型一誤差

(D) 在多重檢定校正的方法中，屬於較為保守的方法

9. 大選在即，電視台針對兩位候選人 A 與 B 進行民調，A 的支持率為 50%，B 的支持率為 30%，請問若此次民調是針對 100 位選民，請回答下列何者”不”正確

(A) A 的支持度在 95% 的信心水準下之信賴區間為 40.2%-59.8%

(B) A 的支持度在 95% 的信心水準下之信賴區間為 21.0%-39.0%

(C) A 的支持度在 99% 的信心水準下之信賴區間為 36.1%-63.9%

(D) B 的支持度在 99% 的信心水準下之信賴區間為 18.2%-41.8%

10. 醫院開發了一新型的檢測試劑，在病患與正常對照組之測試結果如下表，請回答下列何者正確

(A) 敏感度(sensitivity)為 0.7

(B) 特異度(specificity)為 0.8

(C) 陽性預測值(positive prediction value)為 0.75

(D) 陰性預測值(negative prediction value)為 0.78

檢測試劑		病患實際情形	
		陽性	陰性
試驗預測情形	陽性	80	30
	陰性	20	70

見背面

11. 下列關於接收者操作特徵曲線(receiver operating characteristic curve)的敘述

何者”不”正確

- (A) 曲線越靠向圖形右上方越好
- (B) 曲線下面積越大越好
- (C) X軸為 False positive rate
- (D) Y軸為 True positive rate

12. 一研究指出吸菸與肺癌的關係如下表，則下列何者敘述”不”正確？

(A) 吸菸在肺癌的風險勝算比為 4.03

(B) 勝算比的 95%信賴區間為 $4.03 \times e^{\pm 1.96 \sqrt{\frac{1}{380} + \frac{1}{220} + \frac{1}{120} + \frac{1}{280}}}$

(C) 吸菸在肺癌的相對風險為 2.11

(D) 相對風險的 95%信賴區間為 $2.11 \times e^{\pm 1.96 \sqrt{\frac{1}{380} + \frac{1}{600} + \frac{1}{120} + \frac{1}{400}}}$

	肺癌患者	正常對照組	合計
吸菸	380	220	600
不吸菸	120	280	400
合計	500	500	1000

13. 下列敘述何者正確？

- (A) 皮爾森相關係數(Pearson's correlation coefficient)愈大，代表兩變項的線性相關愈強
- (B) 皮爾森相關係數(Pearson's correlation coefficient)愈大，代表兩變項的線性或曲線相關愈強
- (C) 皮爾森相關係數(Pearson's correlation coefficient)愈大，代表兩變項的因果關係愈強
- (D) 以上皆是
- (E) 以上皆非

14. 平均值加減兩倍標準差內涵蓋多少比例的資料？

- (A) 至少68%
- (B) 至少75%
- (C) 約95%
- (D) 約99%
- (E) 以上皆非

接次頁

15. 若某健康照護中心平均每天接獲 90 通民眾諮詢電話，則某天接獲 150 通以上民眾諮詢電話的機率為何？
- (A) 不超過 0.3
 - (B) 不超過 0.5
 - (C) 不超過 0.6
 - (D) 不超過 0.7
 - (E) 以上皆非
16. 以下何種統計分析方法可用來檢定兩獨立組別母體平均數是否相同？
- (A) Two-sample t test
 - (B) Analysis of Variance (ANOVA, 變異數分析)
 - (C) 迴歸分析
 - (D) 以上皆是
 - (E) 以上皆非
17. 以下何者屬於無母數分析方法？
- (A) 符號檢定 (sign test)
 - (B) Wilcoxon signed-rank test
 - (C) Kruskal-Wallis Test
 - (D) 以上皆是
 - (E) 以上皆非
18. 下列敘述何者正確？
- (A) 研究者所能容許犯型一錯誤的最大機率，稱為顯著水準
 - (B) Wilcoxon signed-rank test 是相對於paired t -test 的無母數檢定方法
 - (C) 給定對立假設下的分布才能計算出型二錯誤機率與檢定力
 - (D) 以上皆是
 - (E) 以上皆非

試題隨卷繳回