

一、是非題 20 分 (每題 2 分) ※ 注意：請於試卷內之「非選擇題作答區」標明題號依序作答。

1. 男女兩個不同性別的血型分布(A、B、O、AB)不能以兩個個別的 boxplot 來比較分布狀況。
2. 今年台灣地區家戶所得分布圖呈 M 行分布，這不是單峰(single mode)的分布。
3. 如果要比較臺灣大學甲、乙兩學院男學生所佔的比例是否相同，也可以使用卡方(chi-square)檢定。
4. 變異數分析(Analysis of variance, ANOVA)雖然檢定的是不同母體的母體期望值是否相等，但必須先有各個母體的變異數皆相同的假設才行。
5. 兩個成對母體的 t 檢定可以轉化為檢定單一母體期望值是否為 0 的檢定。
6. 如果要檢定兩組獨立樣本的母體期望值是否相等，也可以利用變異數分析(Analysis of variance, ANOVA)來進行。
7. 伯努利分布(Bernoulli distribution)也是一種二項式分配(Binomial distribution)。
8. 二項式分配(Binomial distribution)一定是對稱的分布。
9. 勝算比(odds ratio)可以拿來評估兩個連續變數之間的關係。
10. 簡單線性迴歸(simple linear regression)探討的是單一解釋變數(covariate)和反應變數(response variable)之間的所有關係。

二、選擇題 80 分 (單選，1-10 每題 3 分，11-20 每題 5 分) ※ 請於試卷內之「選擇題作答區」依序作答。

1. 某研究比較一種降血壓藥物對男性和對女性的效果是否有不同，已知下降程度的變異數在男性和女性中都相等。現有九位男性服藥後平均下降 18.9 毫米汞柱，標準差 SD 為 5.9 毫米汞柱；另外 16 位女性平均下降 11.9 毫米汞柱。請問在進行兩個獨立樣本 T 檢定時，T 檢定統計量的值為

(A)
$$\frac{18.9-11.9}{\sqrt{\frac{8 \times 5.9^2 + 15 \times 6.3^2}{23}}}$$

(B)
$$\frac{18.9-11.9}{\sqrt{\frac{9 \times 5.9^2 + 16 \times 6.3^2}{23}}}$$

(C)
$$\frac{18.9-11.9}{\sqrt{\frac{8 \times 5.9^2 + 15 \times 6.3^2}{23} \left(\frac{1}{9} + \frac{1}{16} \right)}}$$

(D)
$$\frac{18.9-11.9}{\sqrt{\frac{9 \times 5.9^2 + 16 \times 6.3^2}{23} \left(\frac{1}{9} + \frac{1}{16} \right)}}$$

2. (延續上題)請問 T 檢定的自由度為

(A) 256; (B) 24; (C) 23; (D) 22。

3. (延續上題)以下何者正確?

- (A) 上述檢定也可以利用變異數分析來完成
- (B) 上述檢定檢定的是兩個樣本平均值是否相同
- (C) 上述檢定中男性樣本的來源可以是同一個家族成員
- (D) 這是一個配對的世代研究

4. (延續上題)若另一研究人員想估計該藥物對女性的影響，則女性平均降血壓量之 95% 信賴區間為

見背面

(A) $(18.9 - 1.96 \times 5.9, 18.9 + 1.96 \times 5.9)$

(B) $(18.9 - 1.96 \times 5.9 / \sqrt{9}, 18.9 + 1.96 \times 5.9 / \sqrt{9})$

(C) $(18.9 - 1.96 \times 6.3, 18.9 + 1.96 \times 6.3)$

(D) $(18.9 - 1.96 \times 6.3 / \sqrt{16}, 18.9 + 1.96 \times 6.3 / \sqrt{16})$

5. (延續上題)由上一小題可知以下何者正確

(A) 該藥物的降血壓效果對女性有統計顯著的影響, $P < 0.05$

(B) 該藥物的降血壓效果對女性沒有統計顯著的影響, $P < 0.05$

(C) 由上一小題之 95% 信賴區間無法判斷該藥物是否對女性有效

(D) 由上一小題之 95% 信賴區間與 T 檢定之結果不一定一致

6. 某研究探討尼古丁替代療法找來 100 位癮君子參與試驗, 其中有 50 位男士、50 位女士, 分別有 10 位男士、25 位女士經由此法成功戒菸, 請推估一般女士能利用這個方法成功戒菸的比例之 95% 信賴區間為

(A) $(0.5 - 1.96 \times \sqrt{\frac{0.5^2}{50}}, 0.5 + 1.96 \times \sqrt{\frac{0.5^2}{50}})$

(B) $(0.5 - 1.96 \times \sqrt{0.5^2}, 0.5 + 1.96 \times \sqrt{0.5^2})$

(C) $(0.5 - 1.96 \times \sqrt{0.5}, 0.5 + 1.96 \times \sqrt{0.5})$

(D) $(0.5 - 1.96 \times \sqrt{0.5/50}, 0.5 + 1.96 \times \sqrt{0.5/50})$

7. (延續上題)男女性戒菸成功的比例差異之 95% 信賴區間為

(A) $(\frac{25}{50} - \frac{10}{50}) \pm 1.96 \times \frac{0.5 + 0.4}{\sqrt{50}}$

(B) $(\frac{25}{50} - \frac{10}{50}) \pm 1.96 \times \sqrt{\frac{0.5^2}{50} + \frac{0.4^2}{50}}$

(C) $(\frac{25}{50} - \frac{10}{50}) \pm 1.96 \times \sqrt{0.5^2 + 0.4^2}$

(D) $(\frac{25}{50} - \frac{10}{50}) \pm 1.96 \times (0.5 + 0.4)$

8. 某研究探討「成年時期過重」的人與他們「幼年時期過重」是否有關係得以下數據

	成年時期過重	成年時期沒有過重	
幼年時期過重	34	116	150
幼年時期沒有過重	4	109	113
	38	225	263

請問以下何者正確

- (A) 這是配對研究
(B) 這是追蹤研究
(C) 這是回溯性研究
(D) 這是世代研究
9. (延續上題)若使用卡方檢定可得卡方統計檢定量值為 17.49，請問以下何者正確
(A) 這是利用 McNemar 檢定的卡方檢定
(B) 這個檢定也可以利用 T 檢定來完成
(C) 這個卡方檢定的自由度為 2
(D) 檢定結果達統計顯著， $P < 0.05$
10. (延續上題)以下何者不正確?
(A) 成年後過重的人有高比例在幼年時也是過重的
(B) 這是一種無母數假設的檢定方法
(C) 進行卡方檢定時若加總四個格子的期望值，會是 263。
(D) 這個檢定也可以利用簡單線性迴歸來進行
11. 某班學生共 100 人，其中 90 人為近視。您遇到該班兩名學生，兩名皆未近視的機率最接近下列何值?
(A) 0.81
(B) 0.18
(C) 0.10
(D) 0.09
(E) 0.01
12. 樣本數增加，下列何者錯誤?
(A) 成本(cost)增加。
(B) 檢力(power)增加。
(C) 標準誤(standard error)變小。
(D) 偏差(bias)變小。
(E) 信賴區間(confidence interval)變窄。
13. 有關常態分佈，下列何者錯誤?
(A) 為連續性機率分佈。
(B) 具有兩個參數(parameter)。
(C) 為對稱分佈。
(D) 其機率密度函數(density function)的函數值即為機率。
(E) 常態分佈取平方，可得卡方分佈。
14. 有關二項式分佈(binomial distribution)，下列何者錯誤?
(A) 一定是離散分佈(discrete distribution)。
(B) 一定是對稱分佈。
(C) 其可能出現的數值，必為正整數(含零)。
(D) 其各種可能出現數值的機率總和必為 1。
(E) 將數個獨立的伯努利分佈(Bernoulli distribution)相加，可得到一個二項式分佈。

15. 下列何者是無母數方法(non-parametric method)?
- (A) Kruskal-Wallis test
 - (B) ANOVA (analysis of variance)
 - (C) binomial test
 - (D) Student t test
 - (E) paired t test
16. 某研究者抽樣 1000 名社區居民，調查抽菸(有/無)與喝酒(經常/偶爾/沒有)之盛行率。現欲檢定抽菸與喝酒是否有相關，可採用下述何者?
- (A) Student t test;
 - (B) McNemar test
 - (C) chi-square test
 - (D) paired t test
 - (E) F test
17. 有關線性迴歸分析(linear regression analysis)，下列何者錯誤?
- (A) 截距項(intercept)亦是其迴歸係數之一，其值可能為正，亦可能為負。
 - (B) 樣本數超過十萬個案時，受限於電腦計算能力，不建議使用線性迴歸分析。
 - (C) 自變項(independent variable)可以為任何測度尺度的變項。
 - (D) 應變項(dependent variable)的測度尺度為連續變項。
 - (E) 實際觀察值與迴歸預測值間的差(difference)，稱為殘差(residual)。
18. 給定反應變數 Y 及解釋變數 X 。下列關於簡單迴歸分析的敘述何者錯誤?
- (A) R^2 (coefficient of determination)越大，則 X 與 Y 越有線性相關
 - (B) R^2 (coefficient of determination)越小，代表 X 與 Y 無任何關係
 - (C) 若 XY 之間的相關係數絕對值越大，則越適合使用線性迴歸分析
 - (D) 解釋變數 X 可以是連續型的也可以是離散型的
 - (E) 對預測的數值建立信賴區間時，離 X 的平均數越遠的點信賴區間會越寬
19. 關於 p-value 的敘述，何者正確?
- (A) 必須由資料自行決定其數值，非人為指定
 - (B) p-value=0.000001 代表虛無假設為真
 - (C) p-value=0.000001 代表虛無假設為假
 - (D) 信心水準給定後，p-value 就確定了
 - (E) 以上皆錯誤
20. 關於假設檢定的敘述，何者正確?
- (A) 型一誤差與型二誤差必須在執行檢定之前先決定
 - (B) 型一誤差與型二誤差的和為一固定的數
 - (C) 型一誤差決定後，型二誤差也就跟著決定了，且無法改變
 - (D) 若想要有較小的型二誤差，可以採用較小的型一誤差來執行檢定
 - (E) 以上皆錯誤