

公共衛生核心課程基本
能力測驗

生物統計
考試試卷

中華民國 108 年 9 月 7 日

公共衛生核心課程基本能力測驗

108 年生物統計考試試卷

一、選擇題（答案 4 選 1，1-20 題每題 3 分，21-30 題每題 4 分，共 30 題）

- 下列哪一個圖可以幫助我們瞭解身高和體重的相關？
 - 直方圖 (histogram)
 - 散佈圖 (scatter plot)
 - 莖葉圖 (stem-and-leaf diagrams)
 - 長條圖 (bar chart)
- 有項大台北地區吸菸狀況及血壓的研究，隨機抽樣二千名成年市民，吸菸狀況分成吸菸及不吸菸二種類別，血壓用量性資料表示。吸菸狀況這隨機變數為何種變數？何種分佈？
 - 連續變數，常態分佈 (Normal distribution)
 - 連續變數，指數分佈 (Exponential distribution)
 - 離散變數，二項分佈 (Binomial distribution)
 - 離散變數，卜松分佈 (Poisson distribution)
- 有關一組數據平均值 (mean)、中位數 (median) 與眾數 (mode) 的敘述下列何者正確？
 - 眾數 (mode) 只能有一個
 - 平均值 (mean) 永遠會大於或等於中位數 (median)
 - 中位數不容易受到極端值 (extreme observation) 的影響
 - 三者之中眾數 (mode) 數值的大小最容易受到極端值 (extreme observation) 的影響
- 假設有一族群的自殺率為 3%，貧窮率為 20%，若這兩個事件是獨立的，請問隨機選取一個人為貧窮且自殺的機率是多少？
 - 0.6%
 - 0.15%
 - 0.3%
 - 0.5%
- 已知某考試分數為常態分佈，隨機抽了 100 名，樣本平均數為 480 樣本標準差為 100，則母群體平均數 95% 信賴區間大約為？
 - (380, 580)
 - (460, 500)

(C)(280, 680)

(D)(470, 490)

6. 假設檢定中，以 β 為發生型二錯誤的機率， $1-\beta$ 為統計檢定力。何謂統計檢定力？

(A) 當 H_0 是真的，推翻 H_0 的機率

(B) 當 H_0 是真的，沒有推翻 H_0 的機率

(C) 當 H_0 是真的，推翻 H_0 的機率

(D) 當 H_0 是真的，沒有推翻 H_0 的機率

7. 以下 7~8 題為題組：

已知收縮血壓為常態分布。某研究者從某一醫院，隨機收集 16 位 35 至 44 歲成年男子，並且求得樣本收縮壓的平均數(Mean: 80 mm Hg)以及標準差(Standard Deviation: 12 mm Hg)。

此樣本平均數之標準誤(Standard Error)為何？

(A) 0.75

(B) 3

(C) 9

(D) 12

8. 比較此母體平均數之 90%信賴區間與 95%信賴區間得知：

(A) 兩信賴區間一樣長

(B) 90%信賴區間比 95%信賴區間較長

(C) 90%信賴區間比 95%信賴區間較短

(D) 無法比較

9. 有一個病例對照研究調查，探討糖尿病（“1”代表有、“0”代表無）與心肌梗塞（“病例”代表有心肌梗塞、“對照”代表沒有心肌梗塞）是否相關，每一心肌梗塞(病例)配對一同齡健康者作為對照，獲得下列結果：

		無心肌梗塞		
		糖尿病		
		1. 有	0. 無	
有心 肌 梗 塞	糖 尿 病	1. 有	9	37
		0. 無	16	82

糖尿病與心肌梗塞是否相關？設 $\alpha=0.05$ 。

(A) 接受虛無假設，糖尿病與心肌梗塞無關

- (B) 接受虛無假設，糖尿病與心肌梗塞有關
- (C) 拒絕虛無假設，糖尿病與心肌梗塞無關
- (D) 拒絕虛無假設，糖尿病與心肌梗塞有關

10. 當進行單尾單一樣本檢定時，實際母群體平均值為 45 時，我們下的結論為平均值大於 45，此結果是什麼情形？

- (A) 我們犯第一型誤差(type I error)
- (B) 我們犯第二型誤差(type II error)
- (C) 我們做出了正確的判定(made a correct decision)
- (D) 我們有足夠的檢力(power)

11. 某研究想知道三種不同飲食習慣(A, B, C)的男性其平均體重是否不同，已知各種飲食習慣男性的體重呈常態分布且變異數相等，經隨機抽取部份男性，得到下列的變異數分析表：

變異來源	自由度	平方和	平均平方和	F 值
組間變異	****	1294	①	④
組內變異	****	②	③	
總變異	131	3116		

請寫出上表中的①的值分別為何？

- (A) 647
- (B) 315
- (C) 200
- (D) 35

12. 某研究欲分析男女性血膽固醇有無差異，收集資料如下

	人數	平均	標準差
男生組	500	200	12
女生組	500	205	10

在顯著性水準為 0.05 下，採用雙樣本獨立 t-test 的敘述何者正確？

- (A) 男生組與女生組可以有血緣關係
- (B) 男生組與女生組人數一定要相同
- (C) 檢定之臨界值約為1.645
- (D) 若計算共同樣本標準差(pooled standard deviation)約為11

13. 為分析肥胖與憂鬱症的關係，以年齡進行肥胖與體位正常者之配對，觀察有無憂鬱症發生狀況。請問哪種分析方法較合適？

- (A) McNemar' s test
- (B) Fisher' s exact test

- (C) Chi-square test
(D) Paired t-test
14. 下列何者不是無母數(nonparametric)檢定方法?
(A) Wilcoxon rank sum test
(B) Chi-square test
(C) ANOVA
(D) Sign test
15. 下列哪個狀況不適合使用雙樣本(two sample) t-test?
(A) 變異數同質性
(B) 資料呈現常態分佈
(C) 兩組樣本可包含雙胞胎的樣本
(D) 兩組樣本總人數可以不同
16. 某一衛生教育介入的研究,分別測量實驗組及對照組在衛教介入前後知識的分數,若要探討實驗組的衛教介入前後知識分數是否有改變,請問要用下列哪種統計方法?
(A) 單一樣本 t 檢定
(B) 兩組獨立樣本 t 檢定
(C) 成對樣本 t 檢定
(D) 多組獨立樣本的 ANOVA 檢定
17. 以下是一個簡單線性迴歸模式: $\hat{y} = 0.56x + 502$, 以下何者為是?
(A) 相對於每個單位的 X 變化量, Y 值的變化量為 502 個單位
(B) 相對於每個單位的 Y 變化量, X 的變化量為 0.56 個單位
(C) 相對於每個單位的 Y 變化量, X 的變化量為 0.56+502
(D) 當 X 為 0 時之 Y 值的平均為 502 個單位
18. 如果所有的觀察值 Y_i 均落在迴歸線上, 則
(A) 依變數的總變異量 = 0
(B) 決定係數(R square) = 0
(C) 迴歸可解釋的變異量 = 迴歸無法解釋的變異量
(D) 迴歸無法解釋的變異量 = 0

19. 有關皮爾森相關 (Pearson' s correlation coefficient)的統計方法之敘述，何者錯誤？
- (A) 是由皮爾森(K. Pearson)所發展出來的統計方法
- (B)用 r 表示； r 係數的範圍是介於-1 至 1 之間
- (C) 相關係數 $r \neq 0$ ，表示 X 與 Y 兩連續變數之間有因果關係存在
- (D) 相關係數 $r=0$ ，表示 X 與 Y 兩連續變數之間，沒有直線關係存在
20. 某研究探討年齡與血中膽固醇濃度之線性相關，共招募 13 位志願者參加，資料經收集與分析後，獲得下列報表，請回答下列各題：

簡單直線迴歸分析摘要表

模式	迴歸係數估計值	標準誤	t 值	p 值
(常數)	-54.877	40.973	-1.339	0.207
年齡	5.590	0.744	7.517	<0.001

依變數：膽固醇

本研究年齡迴歸係數及截距項，經統計檢定(訂 $\alpha=0.001$)，下列敘述何者正確：

- (A) 年齡迴歸係數顯著，截距不顯著
- (B) 年齡迴歸係數不顯著，截距顯著
- (C) 年齡迴歸係數與截距皆顯著
- (D) 年齡迴歸係數與截距皆不顯著
21. 給一出現正面機率為 $1/3$ 的硬幣，在已出現 1 次正面的情況下，投出反面的機率為何？
- (A) $1/3$
- (B) $2/3$
- (C) $1/9$
- (D) $2/9$
22. 假設 X 服從標準常態分布 $N(\mu = 0, \sigma^2 = 1)$ 。請問以下敘述何者正確？
- (A) X 數值落在 $[-1,1]$ 之間的機率約為 0.95
- (B) X 的期望值會比中位數來的大
- (C) $X + 5$ 會服從常態分布 $N(\mu = 5, \sigma^2 = 1)$
- (D) $5X$ 會服從常態分布 $N(\mu = 0, \sigma^2 = 5)$

23. 從一隨機變數 $X \sim N(\mu, \sigma^2)$ ，抽 16 人 ($n_1=16$)，得 \bar{X}_1 ，又從相同母體抽 64 人 ($n_2=64$)，得 \bar{X}_2 。問 \bar{X}_1 與 \bar{X}_2 所分別得出之 μ 的 95% 信賴區間寬度之關係？
- (A) \bar{X}_1 得之 CI = 0.5 倍 \bar{X}_2 之 CI
- (B) \bar{X}_1 得之 CI = 1 倍 \bar{X}_2 之 CI
- (C) \bar{X}_1 得之 CI = 2 倍 \bar{X}_2 之 CI
- (D) \bar{X}_1 得之 CI = 4 倍 \bar{X}_2 之 CI
24. 今有一研究，有 100 位參與者，年齡為 45 ± 1 歲 (mean \pm standard error)。下列那一項陳述是錯誤的：
- (A) 參與者年齡的標準差為 10 歲
- (B) 我們有 95% 的信心，相信母體平均年齡介於 43~47 歲之間
- (C) 此研究可以應用中央極限定理，推論母體平均年齡
- (D) 這 100 位參與者年齡的分佈必須服從常態分佈，才能對母體平均年齡進行推論
25. 一般認為 6 歲以下小孩暴露在高濃度鉛的環境下，其血紅素濃度普遍較低。已知血紅素平均值 $\mu = 12.29 \text{ g}/100 \text{ ml}$ ，標準差 $\sigma = 0.85 \text{ g}/100 \text{ ml}$ ，現隨機抽取 74 位暴露在高濃度鉛的小孩，得平均血色素 $\bar{x} = 10.6 \text{ g}/100 \text{ ml}$ ，哪種檢定較宜？(假設 $\alpha = 0.05$)
- (A) 雙尾檢定 $H_0: \mu = 12.29$, $H_1: \mu \neq 12.29$
- (B) 雙尾檢定 $H_0: \mu \geq 12.29$, $H_1: \mu > 12.29$
- (C) 單尾檢定 $H_0: \mu \geq 12.29$, $H_1: \mu < 12.29$
- (D) 單尾檢定 $H_0: \mu \leq 12.29$, $H_1: \mu > 12.29$
26. 有一研究收集總計 210 位病人，比較三種降血壓藥物的效果，進行 one-way ANOVA 檢定結果 $p\text{-value} = 0.04$ ，在顯著性水準為 0.05 下，下列敘述何者有誤？
- (A) $p\text{-value}$ 可自 F 分配自由度為 (2, 207) 查出
- (B) 此分析結果為三種降血壓藥物的效果全不相同
- (C) 此分析結果須再進行事後分析 (post-hoc analysis)
- (D) 三種降血壓藥物的病人彼此是不相關的

27. 在一個比較四組平均數的one-way ANOVA摘要表，以下何者為非？
- (A) 組間之自由度=3
- (B) 組平方和 (group sum of squares) + 誤差平方和 (error sum of squares) = 總平方和 (total sum of squares)
- (C) 組均方 (group mean square) + 誤差均方 (error mean square) = 總平均均方 (total mean square)
- (D) 各組之個案數不必要相等

28. 以下是一個隨機化設計(Randomized Design)試驗的變異數分析(ANOVA)摘要表，部分數據並未顯示。試問總平方和是多少？

變異來源	平方和	自由度	均方	F
處理方式		3	21	
誤差	48	16		
總和				

- (A) 111
- (B) 146
- (C) 258
- (D) 359
29. 下列敘述何者不正確？
- (A) Spearman' s correlation 比起 Pearson' s correlation 較不受到特異值(outlying values)的影響
- (B) 相關係數(Pearson' s correlation coefficient)僅可表達線性(linear)相關
- (C) Spearman' s correlation coefficient 的數值是-1 到 1
- (D) 當相關係數 $r=0.12$ 且檢定 " $H_0: \rho=0$ " 的 p 值是 0.001，則可以解釋為兩個數值變項有很強的線性正相關之關係
30. 簡單迴歸方程式 $E(Y) = \alpha + \beta X$ ， α 估計值為 = 15.6， β 估計值為 4.24，斜率之標準誤為 = 2.8，以下何者為非？
- (A) 該模式之決定係數是兩變項之皮爾森相關係數的平方
- (B) 以 t 檢定 (顯著水準=0.05) 推論 $\beta=0$ ，若無法拒絕虛無假設，則指出 X 與 E(Y) 沒有線性相關
- (C) X 每增加 1 個單位，Y 平均估計增加 4.24 單位
- (D) 若樣本數為 100，此迴歸模式中斜率達統計上顯著不同於 0