

# 公共衛生核心課程基本能力測驗

## 100 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）

- 下列何者成立時，A 與 B 兩事件為獨立？
  - $P(A)+P(B)=1$
  - $P(A)\times P(B)=0$
  - $P(A \cap B) = P(A)\times P(B)$
  - $P(A \cup B) = P(A)+P(B)$
- 已知一名公衛系學生通過專業能力學科測驗的機率為 0.5，試問隨機抽取四名公衛系學生，在符合二項分配的條件下，四人皆通過的機率有多大？
  - 0.9375
  - 0.8000
  - 0.0625
  - 0.0020
- 有關資料種類（types of data）的描述以下何者正確？
  - 連續型資料（continuous data）不可轉化為等級資料（ranked data）、序位資料（ordinal data）、或類別資料（categorical data）
  - 血壓值（mmHg）是連續型資料（continuous data）
  - 喜好程度是等級資料（ranked data）
  - 對於序位資料（ordinal data）來說，各分類的大小次序有意義，且各分類間差異幅度要相同
- 關於盒鬚圖（boxplot），下列敘述何者正確？
  - 適用於連續變項
  - 是一種將偏態資料（skewed data）轉換成常態分布資料的方法
  - 能畫出一組資料的可能信賴區間的分布
  - 提供一個母體抽出之可能樣本平均數分布之視覺描述
- 對一個右偏態（right skewed）的分布，一般而言：
  - 中位數大於平均數
  - 半數的觀察值大於平均數
  - 分布的左尾較右尾短
  - 眾數大於平均數

6. 請問下列敘述何者為真？
- (A) 長條圖 (bar chart) 常用以表示連續型資料的分布
  - (B) 直方圖 (histogram) 常用以圖示類別與序位資料
  - (C) 圓餅圖 (pie chart) 為常用於呈現連續資料的圖形
  - (D) 二維散布 (two-way scatter plot) 是用來表現兩個不同的連續型資料之間的關係
7. 有一項膽固醇的研究，已知 20-40 歲的人其血液中膽固醇平均數為 180 mg/dL。想證明 41-60 歲中老年人的膽固醇高於 180 mg/dL，如果中老年人膽固醇平均數以  $\mu$  表示。虛無假設及對立假設要如何陳述？
- (A)  $H_0: \mu = 180$  VS  $H_1: \mu \neq 180$
  - (B)  $H_0: \mu \neq 180$  VS  $H_1: \mu = 180$
  - (C)  $H_0: \mu \leq 180$  VS  $H_1: \mu > 180$
  - (D)  $H_0: \mu \geq 180$  VS  $H_1: \mu < 180$
8. 關於中央極限定理 (Central Limit Theorem)，下列何者敘述錯誤？
- (A) 原本之母群體分布可以為離散型均勻分布 (discrete uniform distribution)
  - (B) 樣本平均值的抽樣分布之期望值 (mean) 等於母群體的平均值
  - (C) 中央極限定理成立的條件之一為母體分布「對稱」 (symmetry)
  - (D) 樣本數愈大，樣本平均值的抽樣分佈愈趨近於常態分佈
9. 一間大型醫學研究中心為了解成年之病人肥胖之比例，隨機抽取 100 位成年之病人發現其中 18 位屬肥胖。請問根據此樣本所得之肥胖比例估計的標準差約為何？
- (A) 0.2
  - (B) 0.16
  - (C) 0.4
  - (D) 0.038
10. 中央極限定理是指「樣本平均數的抽樣分配會趨近常態分配」，但有幾項條件要成立，不包含下列何者？
- (A) 樣本數要足夠大，比方說大於 30
  - (B) 樣本必須是隨機樣本
  - (C) 母體分配限定是常態分配
  - (D) 若樣本數夠大，母體分配對稱與否並不會影響此定理之結論

11. 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，並且求得樣本中收縮壓的平均數 (Mean: 80 mmHg) 以及標準差 (Standard Deviation: 12 mmHg)。此樣本平均數之標準誤 (Standard Error) 為何？
- (A) 0.75  
 (B) 3  
 (C) 9  
 (D) 12
12. 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		
有 心 肌 梗 塞	糖 尿 病		1.有	0.無
		1.有	9	37
		0.無	16	82

請問該用何種檢定方法進行檢定？

- (A) 卡方檢定  
 (B) McNemar's 檢定  
 (C) Wilcoxon 符號等級檢定  
 (D) ANOVA 檢定
13. 研究 50 位糖尿病患在服藥前後的血糖差異是否因性別不同，則在收集服藥前後的血糖差異與性別資料之後應採何種分析？
- (A) 相依樣本  $t$  檢定 (paired-sample  $t$  test)  
 (B) 獨立樣本  $t$  檢定 (independent-sample  $t$  test)  
 (C) 單一樣本  $t$  檢定 (one-sample  $t$  test)  
 (D) 皮爾森相關分析 (Pearson's correlation coefficient analysis)
14. 下列何者不是變異數分析(ANOVA)的基本假設？
- (A) 獨立性 (independence)，每一觀察值彼此獨立  
 (B) 常態性 (normality)，母群體資料依循常態分布  
 (C) 變異數同質性 (homoscedasticity)，每個母體變異數相同  
 (D) 線性關係 (linearity)，母體資料存在線性關係

15. 有關假設檢定的敘述，下列何者為非？
- (A) p-value 的意義為虛無假設為真時，觀察到比目前手上資料算出的檢定統計量值更為極端的機率
  - (B) 顯著性水準為犯型一誤差的最大上限
  - (C) 只有在對立假設為真時，才有可能犯型二誤差
  - (D) 犯型一誤差及型二誤差的機率加總必為 1
16. 在進行卡方檢定時，若有某一個細格之期望值小於或等於 5，則須選用何種統計方法較適當？
- (A) Sign test
  - (B) S-W test
  - (C) Fisher's exact test
  - (D) McNemar's test
17. 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
18. 簡單線性迴歸係數可以下列何種原理直接估計？
- (A) 最小平方法
  - (B) 最大平方法
  - (C) 卡方檢定方法
  - (D) 皮爾森相關係數分析法
19. 已知迴歸模式為  $\mu_{y|x}=4.0+0.8x$ ，試估計若懷胎週數 (x) 為 26 週，則出生嬰兒頭圍 (y) 的平均值 (期望值) 為何？
- (A) 24.80
  - (B) 4.69
  - (C) 20.35
  - (D) 22.63
20. 承上題，試估計懷胎週數每增加 10 週，則嬰兒平均頭圍會增加為何？
- (A) 6.35
  - (B) 0.8
  - (C) 8
  - (D) 6.5

21. 有關愛滋病盛行率研究，某地區隨機抽取一千名，結果陽性率為 2% 也就是  $P(\text{test is positive}) = 0.02$ ，此愛滋病篩檢敏感度為 0.99 也就是  $P(\text{test is positive} | \text{Disease is present}) = 0.99$ ，特異性為 0.995 也就是  $P(\text{test is negative} | \text{Disease is absent}) = 0.995$ ，使用一些機率法則計算愛滋病的盛行率也就是  $P(\text{Disease is present})$  大約為何？
- (A) 1.52%
  - (B) 2.05%
  - (C) 5.51%
  - (D) 0.91%
22. 假設成年人有規律作運動（每星期三次或三次以上）的機率為 42%，隨機抽出 15 個成年人，已知彼此之間運動與否為相互獨立，令  $X$  表示其中有規律運動的人數。試計算  $X$  分布的平均值及變異數。
- (A) 平均值=8.7，變異數= 1.912
  - (B) 平均值=6.3，變異數= 1.912
  - (C) 平均值=8.7，變異數=3.654
  - (D) 平均值=6.3，變異數=3.654
23. 假設全部成年男子的身高為一常態分佈，則從成年男子中隨機抽取 100 人，其可能的樣本組合很多，每一樣本組合都可算出樣本平均值，請問所有可能樣本平均值間的標準差與原來母群體之標準差相比：
- (A) 資料不全，不能比較
  - (B) 比母群體之標準差小
  - (C) 等於原來母群體之標準差
  - (D) 比母群體之標準差大
24. 令  $X_1, X_2, \dots, X_{36}$  為來自  $N=600, P=0.4$  的二項式分布之一組隨機樣本，則  $\Pr(\bar{X} > 243.3)$  的值最接近以下何者？
- (A) 0.95
  - (B) 0.90
  - (C) 0.05
  - (D) 0.10

25. 若一常態母體的實際平均值為 211，標準差為 46，令顯著水準為 0.01，在僅容許 0.05 的機會錯誤地接受虛無假設下，檢定  $H_0: \mu \leq 180$  至少需多少樣本？

**(P(Z > 2.32) = 0.01)**

- (A) 41
- (B) 39
- (C) 37
- (D) 35

26. 某研究者從甲社區隨機抽出 100 名 40-60 歲居民，其中的 10 名被診斷為罹患冠狀動脈心臟疾病 (CHD)，請估計甲社區居民患有 CHD 盛行率 (prevalence) 之 95% 信賴區間為何？

- (A) 0.04-0.16
- (B) 0.06-0.14
- (C) 0.08-0.12
- (D) 0.09-0.11

以下 27~28 題為題組：

新生兒的種族與心跳頻率 (beats per minute) 關係如下表：

	Mean	SD	樣本數
種族 A	125	11	218
種族 B	133	12	156

27. 若已知新生兒的心跳頻率為常態分布，且種族間的變異相同，下列何者可用於檢定新生兒的種族與心跳頻率的關係？

- (A) Two Sample T Test
- (B) Paired T Test
- (C) Mann-Whitney U Test
- (D) Signed Rank Test

28. 若新生兒的心跳頻率為常態分布，假設不同種族的心跳頻率變異數相同，試估計此共同變異數 (Common Variance) ？

- (A) 11.4
- (B) 130.6
- (C) 162.6
- (D) 3.5

29. 下列敘述何者不適當？

- (A) 某族群之年齡和收入之間的皮爾森相關係數是0.27
- (B) 身高和體重相關係數為0則表示身高和體重兩變數互相獨立
- (C) 學生對教師的教學評量，和他對自己的學習評量之皮爾森相關係數為  $\pi/4$
- (D) 18 歲以前的小孩其身高和體重有很高的線性相關

30. 以下何者為非？

- (A) 迴歸線的截距項不具意義
- (B) 迴歸線代表的是自變數與對應之依變數平均值的線性關係
- (C) 迴歸線可以用來預測在不同自變數數值下對應之依變數的平均數值
- (D) 迴歸線的斜率代表當自變數改變一單位時，依變數平均值變動的量

#### 附錄

若  $Z$  為標準常態分布，則  $P(|Z| \geq 1.96) = 0.05$ ， $P(|Z| \geq 1.64) = 0.10$

若  $\chi^2$  為自由度為 1 的卡方分布，則  $P(|\chi^2| \geq 3.84) = 0.05$