

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 106 年 9 月 2 日

公共衛生核心課程基本能力測驗

106 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)

1. 以下為一復健中心病人的資料，何者為質性變項？
(A) 病人的年齡
(B) 病人住院的星期數
(C) 自從住院後接受治療的總時數
(D) 住院的原因
2. 一組五名病人所構成的樣本中，其體重分別為 65，68，62，64，66 公斤，請問這一組樣本之標準差為幾公斤？
(A) $\sqrt{2}$
(B) $\sqrt{5}$
(C) $\sqrt{16}$
(D) $\sqrt{20}$
3. 描述資料時，為何用標準差(standard deviation)而非變異數(variance)？
(A) 標準差比較好計算
(B) 標準差和平均值有相同的單位
(C) 計算變異數一定要知道標準差
(D) 用哪一個都一樣，因為兩者相等
4. 某班公衛系學生 50 名參加衛生行政與環境衛生兩科測驗，其中衛生行政的全班平均分數為 78 分、標準差為 8 分；環境衛生的全班平均分數為 73 分、標準差為 7.6 分。該兩科目測驗分數分布之相對離散情形為何？
(A) 衛生行政=環境衛生
(B) 衛生行政>環境衛生
(C) 衛生行政<環境衛生
(D) 資訊不足無法判定

5. 以下哪一種量度法可以指出體重樣本間的變異程度大於身高樣本間的變異程度？
- (A) 標準差(standard deviation)
 - (B) 標準誤(standard error)
 - (C) 變異係數(coefficient of variation)
 - (D) 四分差(interquartile range)
6. 下列關於 Poisson(λ) 分布的敘述，何者錯誤？
- (A) 是一對稱分布
 - (B) $E(X) = \lambda$
 - (C) $\text{Var}(X) = \lambda$
 - (D) 常用來描述某事件發生次數的分布
7. 檢定一組樣本之母體期望值 μ ，且虛無假設 H_0 為 $\mu = \mu_0$ ，又對立假設 H_1 為 $\mu = \mu_1$ ，在其他條件相同下，統計檢力(power)會隨著下列何者增加而下降？
- (A) μ_0 與 μ_1 的差異
 - (B) 型一誤差的機率(Type I error)
 - (C) 樣本數
 - (D) 變異數
8. 隨機抽取 100 人，其中有 12 個人為左撇子。請問這個族群左撇子比率的 95% 信賴區間為：
- (A) 0.056 to 0.184
 - (B) 0.037 to 0.203
 - (C) 0.056 to 0.284
 - (D) 1.521 to 2.384
9. 有項膽固醇的研究，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。我們調查 100 名中老年人膽固醇，假設母群體標準差為 100 mg/dL，在型一錯誤機率 $\alpha=0.05$ 設定下，樣本平均數要大於多少才會拒絕虛無假設，也就是臨界值(critical value)為何？
- (A) 220.8
 - (B) 196.5
 - (C) 192.6
 - (D) 199.6

10. 下列何者敘述錯誤？
- (A) 在顯著水準 α 相同時，若要推翻虛無假設，則雙尾檢定所需之樣本數小於單尾檢定；
 - (B) 通常 α -error 與 β -error 是對立的，若要使二者同時變小，則必須增加抽樣的樣本數；
 - (C) 進行統計檢定時，若 p -value 小於顯著水準 α ，則應推翻虛無假設；
 - (D) 進行區間估計時，若希望信賴區間短，同時又希望信賴水準高，則必須要增加樣本數。
11. 針對勝算比(Odds ratio, OR)取自然對數可得 $\ln(\text{OR})$ ，因此 OR 或 $\ln(\text{OR})$ 之 95% 信賴區間，下列何者正確？
- (A) $[\ln\widehat{OR} - 1.645 \times SE(\ln\widehat{OR}), \ln\widehat{OR} + 1.645 \times SE(\ln\widehat{OR})]$
 - (B) $[\widehat{OR} - 1.645 \times SE(\widehat{OR}), \widehat{OR} + 1.645 \times SE(\widehat{OR})]$
 - (C) $[\ln\widehat{OR} - 1.96 \times SE(\ln\widehat{OR}), \ln\widehat{OR} + 1.96 \times SE(\ln\widehat{OR})]$
 - (D) $[\widehat{OR} - 1.96 \times SE(\widehat{OR}), \widehat{OR} + 1.96 \times SE(\widehat{OR})]$
12. 在台灣，母親懷孕 40 週出生的嬰兒，體重分佈近似常態，其平均數為 2500 公克，標準差為 500 公克。當樣本數為 25，則樣本平均數大於或等於多少時，會屬於樣本平均數抽樣分佈的 95% 以上？
- (A) 2696
 - (B) 2665
 - (C) 2336
 - (D) 2304
13. 下列何者不為描述資料離散程度的統計量？
- (A) 標準差(standard deviation)
 - (B) 四分位差(interquartile range)
 - (C) 中位數(median)
 - (D) 全距(range)
14. 欲探討學歷(國小、國中、高中、大學)與「平均每週看電視時數」的關係，應採下列何種統計分析方法？
- (A) 卡方檢定
 - (B) 單一樣本 t 檢定
 - (C) Paired t 檢定
 - (D) ANOVA

15. 藥廠宣稱，其生產之藥品每罐平均重量至少為 500 公克，標準差為 14 公克。今自其藥品中隨機抽取 49 罐測定其重量，得其平均值為 488 公克。試檢定此藥廠之宣稱是否成立？顯著水準為 0.05。其 p 值(p value)為何？
- (A) 約為 1
(B) 約為 0
(C) 約為 0.7
(D) 約為 0.3
16. 下列何者敘述為真？
- (A) 型 II 誤差 = $\Pr(\text{reject } H_0 | H_0 \text{ is true})$
(B) 型 I 誤差=0.05
(C) 若型 I 誤差(Type I error)愈大，則型 II 誤差(Type II error)愈大
(D) 型 I 誤差 = $\Pr(\text{Do not reject } H_0 | H_a \text{ is true})$
17. 下列有關迴歸係數(β)和相關係數(r)的敘述何者正確？
- (A) r永遠都大於零
(B) r愈大， β 也就愈大
(C) β 指自變項每增加一個單位，依變項平均值增加 β 單位
(D) r 有自變項及依變項之分別
18. 如果針對簡單線性迴歸的斜率項 β 之檢定結果為拒絕 $H_0: \beta=0$ ($\alpha=0.05$)。則以下何者為非？
- (A) 足夠證據顯示自變數和依變數間存在線性關係
(B) β 的信賴區間不可能為(-1.28, 0.5)
(C) β 的信賴區間不可能為(-0.04, 0.02)
(D) 檢定 β 的 p 值大於 0.05
19. 有關簡單線性迴歸的敘述何者有誤？
- (A) 殘差(residual)為隨機誤差項的估計值
(B) 殘差通常被用來做為迴歸模式假設之檢查
(C) 最小平方方法是在考量使殘差平方和最大的要求下進行迴歸係數的估計
(D) 迴歸係數與截距項並無同號關係

20. 假設行政院農委會官員想要了解農民氮肥用量與稻穀收量之關係，今以 x 代表氮肥用量(公斤)， y 代表稻穀收量(公斤)，每一氮肥用量下可獲得一稻穀收量，其結果如下表：

用量 (x)	0	0.5	1.0	1.5	2.0	2.5	$S_{xx}=4.375,$ $\sum x = 7.5, \bar{x} = 1.25$
稻穀 收量 (y)	10	18	32	48	55	62	

今若以一直線模式 $y=b_0+bx$ 來表示 x 與 y 的關係，利用最小平方法估算 b_0 得 9.86 及 b 得 22.11。

請問以下敘述何者正確？

- (A) 氮肥用量與稻穀收量有正相關
 (B) 氮肥用量與稻穀收量有負相關
 (C) 如果氮肥用量為 100 公斤，則稻穀收量超過 2000 公斤
 (D) 如果氮肥用量增加 1 公斤，則稻穀收量增加 9.86 公斤
21. 某醫院急診室平均每天(24 小時)進來的個案數為 18 人，急診每天收到個案數呈卜瓦松分布(Poisson distribution)。若目前沒有個案進來，醫護人員也沒有正在處理收案，請問此醫護人員，從此刻算起可以休息 8 小時，而不被收案干擾的機率為何？
- (A) 0.0498
 (B) 0.0025
 (C) 0.0183
 (D) <0.0001
22. 進行變異數分析(ANOVA)，如欲檢定 4 個群體的收縮壓平均值是否相等，則下列敘述何者為誤？
- (A) 4 組組內(Within-group)均方可以用來估算這 4 個母群體收縮壓的變異數
 (B) 如推翻虛無假設，可再進行事後比較
 (C) 如接受虛無假設，不必再進行事後比較
 (D) F 檢定值=組內(Within-group)均方/組間(Between-group)均方

23. 若已知男生身高呈常態分布(Normal distribution)，期望值(μ)為 162 公分，標準差(σ)為 3 公分，則屬於身高最矮的 2.5%，其身高是幾公分以下？
- (A) 167.9 公分。
 - (B) 158.1 公分。
 - (C) 157.1 公分。
 - (D) 156.1 公分。
24. 設一二元資料，母體成功比例 $\pi=0.1$ 。今隨機抽出 100 位，問樣本成功比例 P 在 0.05 或以下之機率最接近以下哪一個答案？
- (A) 0.02
 - (B) 0.05
 - (C) 0.29
 - (D) 0.58
25. 如果我們進行母群體平均數 95%信賴區間估計，假設要使區間範圍縮減為原來的一半。假設母群體標準差為 46，樣本數至少要多少個，才能達到在 95%信賴度下，樣本平均數估計母群體平均數的誤差範圍不超過 5？
- (A) 115
 - (B) 180
 - (C) 36
 - (D) 326
26. 估計出的信賴區間是否包含全班統計學的平均成績值。
- (A) 一定包含
 - (B) 有 90%的機率會包含
 - (C) 一定不包含
 - (D) 無法確定是否一定包含

27. 有 20 位學生參加微積分考試，若學生來自三個科系，欲比較三個科系的學生成績是否有差異，下表為分析之部分結果，下列答案何者正確？

ANOVA

Grade					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	3308.811		(C)	(D)	.003
Within Groups	(A)		201.082		
Total	(B)				

- (A) 3619.476
- (B) 6928.287
- (C) 1102.937
- (D) 8.228**

28. 一個肺部疾病研究，將吸菸狀態分為數組，記錄其肺功能(FEV, L/s)之後，利用單因子變異數分析(One Way ANOVA)，將部分結果整理如下：

變異(Source of Variation)	Sums of Squares	Degrees of Freedom
組間 (Between)	184.38	5
組內 (Within)	663.87	1044

此變異數分析模型，吸菸狀態對於肺功能的變異，共解釋了多少？

- (A) 18.4%
- (B) 21.7%**
- (C) 36.2%
- (D) 78.3%

29. 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

利用簡單線性迴歸，分析 30 歲以上成人的年齡(自變數為 X: 年)與膽固醇(應變數為 Y: mg/ml)，得到迴歸模型

$$\hat{Y} = -60 + 6x, \text{ model p-value(模型的 p 值)}=0.02,$$

intercept p-value(截距的 p 值)=0.001，以下何者解釋有誤？

- (A) P-value for the slope (斜率的 p 值)為 0.02
 (B) 截距項達統計顯著水準
 (C) 每增加一歲，膽固醇增加 6 mg/ml
 (D) 18 歲的成年人，預估膽固醇為 48 mg/ml
30. 同上題，本研究關於年齡之迴歸係數的描述，下列何者正確？
 (A) 年齡迴歸係數= -54.877，且達統計上顯著的意義(p=0.207)
 (B) 年齡每增加一個單位，血中膽固醇的變化量，經 t 檢定，是未達統計顯著意義(p<0.001)
 (C) 年齡每增加一個單位，血中膽固醇的變化量，經 t 檢定，是有達統計顯著意義(p=0.02)
 (D) 年齡與血中膽固醇的變化量成負相關