

公共衛生核心課程基本能力測驗

103 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. 一般而言「樣本資料分佈之圖形為偏右分佈」時，下列何者正確？

- (A) 表示圖形之主峰偏右
- (B) 表示資料中有若干個極小數值之資料點
- (C) 樣本之眾數(mode) > 樣本之平均數(mean)
- (D) 樣本之平均數(mean) > 樣本之中位數(median)

2. 下列對卜瓦松隨機變數 X 的機率函數之表示，何者最正確？

- (A) $f(x) = \frac{e^{-\mu}\mu^x}{x!}, x = 0,1,2, \dots, n$
- (B) $f(x) = \frac{e^{-\mu}\mu^x}{x!}, x = 0,1,2, \dots, \infty$
- (C) $f(x) = \binom{n}{x}p^x(1-p)^{n-x}, x = 0,1,2, \dots, n$
- (D) $f(x) = \binom{n}{x}p^x(1-p)^{n-x}, x = 0,1,2, \dots, \infty$

3. 以下為一組年齡資料所繪製之莖葉圖(stem-and-leaf display)，則眾數(mode)是？

4	57
5	0133559
6	0237778
7	59
8	0015
9	0123

- (A) 53
- (B) 55
- (C) 67
- (D) 以上皆非

4. 下列何者不適合用來描述 6 年 8 班的血型分佈？
- (A) 長條圖(bar chart)
 (B) 盒鬚圖(box plot)
 (C) 圓餅圖(pie chart)
 (D) 次數多邊圖(frequency polygons)
5. 下表為 76 名高血壓住院病人之收縮壓(SBP)測量值分佈情形，以下哪一個統計圖最適合用於描述此高血壓病人樣本 SBP 之分佈？

SBP 區間	頻率
220–229	5
210–219	7
200–209	20
190–199	16
180–189	15
170–179	8
160–169	5

- (A) 長條圖(bar chart)
 (B) 圓餅圖(pie chart)
 (C) 盒鬚圖(box chart)
 (D) 直方圖(histogram)
6. 下列何者為常態分佈(normal distribution)之特性？
- ①對稱 ②連續性 ③橫軸上只有正值 ④兩端均伸展至無限
 ⑤兩屈折點(Inflection point)間的面積約等於 0.68
- (A) ①②③⑤
 (B) ①②③④
 (C) ①②④⑤
 (D) ②③④⑤
7. 利用一個 95% 的信賴區間估計「執行一個新的保險政策所需要的平均時間 μ 」，得到信賴區間為 $12 \leq \mu \leq 14$ 天，其意為？
- (A) 所有的政策中只有 5% 之執行會小於 12 天或大於 14 天
 (B) 所有的政策中只有 5% 之執行會介於 12 天與 14 天之間
 (C) 在樣本大小相同的樣本中，每 100 次中約有 95 次所得的信賴區間會包含母群體的平均值 μ
 (D) 所有的政策執行約有 95% 的機會介於 12 天與 14 天之間

8. 今有一研究，16 位病人使用一種新藥後，收縮壓平均下降 5.2mmHg，這 16 位病人血壓差(difference)的標準差為 4.2mmHg。下列那一項陳述是正確的？
- (A) 參與者平均血壓差的標準誤為 1.05mmHg
 (B) 以 t 檢定探討此新藥是否有效，其 t 檢定值=1.24
 (C) 此研究虛無假設為血壓差平均值不為零
 (D) 以 t 檢定探討此新藥是否有效，其自由度為 16
9. 有項膽固醇相關研究，已知 20-40 歲者其血液中膽固醇平均數為 180 mg/dL。若想證明 41-60 歲中老年人的膽固醇高於 180 mg/dL，且中老年人膽固醇平均數以 μ 表示，則虛無假設及對立假設要如何陳述？
- (A) $H_0: \mu = 180$ VS $H_1: \mu > 180$
 (B) $H_0: \mu > 180$ VS $H_1: \mu \leq 180$
 (C) $H_0: \mu \geq 180$ VS $H_1: \mu < 180$
 (D) $H_0: \mu \leq 180$ VS $H_1: \mu > 180$

[以下 10-12 題為題組]

10. 想要檢定大學男生的體重是否為 70 公斤，若自大學男生中抽出一個 49 人的樣本，此樣本平均數為 73.3 公斤，設母體標準差為 10 公斤，顯著水準=0.05。則下列敘述何者為非？
- (A) 體重平均數之標準誤約為 1.43
 (B) 不需假設母體的分配，中央極限定理亦可能成立
 (C) 標準誤即樣本資料之標準差
 (D) 標準誤也是體重平均數分配的標準差
11. 檢定大學男生的體重平均數是否為 70 公斤，其檢定統計量為？
- (A) 0.231
 (B) 16.17
 (C) 1.617
 (D) 2.31
12. 若要求檢定力為 0.9(在對立假設下，大學男生的體重為 74 公斤，根據文獻提供)，最少需多少樣本數才能達到此檢定力的水準？
- (A) 40 人
 (B) 30 人
 (C) 54 人
 (D) 66 人

13. 甲乙兩種診斷工具評估 100 個病人的血液樣本，其中 25 個病人甲乙診斷工具都出現陽性反應，35 個病人甲乙診斷工具都出現陰性反應，30 個病人甲診斷為陽性而乙診斷為陰性，10 個病人甲診斷為陰性而乙診斷為陽性。如要檢定甲乙兩種診斷工具是否有差異，則檢定量為？
- (A) 5.872
 - (B) 10
 - (C) 0.015
 - (D) 4.89
14. 變異數分析(ANOVA)檢定多組平均數是否相同，其檢定原理主要是利用哪兩種變異的比值進行檢定？
- (A) 組間變異/組內變異
 - (B) 組間變異/總變異
 - (C) 組內變異/總變異
 - (D) 迴歸變異/總變異
15. 若欲分析男生與女生的 BMI(kg/m²)有無差異(假設不確定男女生母群體的 BMI 資料是否成常態分配)，今收集 10 位男生與 8 位女生之 BMI (kg/m²)資料(已知 BMI 範圍介於 18.9~36.8)，請問下列何種分析方法最恰當？
- (A) 獨立樣本 t 檢定(independent-sample t-test)
 - (B) 成對樣本 t 檢定(paired-t test)
 - (C) Wilcoxon 等級和檢定(Wilcoxon rank sum test)
 - (D) 卡方檢定(chi-square test)
16. 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(\bar{X})為 488 公克。下列何者為假設檢定之拒絕域？(Z 為標準常態分配，顯著水準為 0.05)
- (A) $\bar{X} \leq 496.08$
 - (B) $Z \geq 1.95$
 - (C) $\bar{X} \leq 503.92$
 - (D) $Z \leq 1.645$

17. 某流病專家想探討喝茶與否和罹患代謝症候群可能的相關性。他在某工廠內問了每個工人是否有喝茶的習慣，同時作體檢，確定是否有代謝症候群，請問該用哪一種統計方法？
- (A) F 檢定
 - (B) Chi-square 檢定
 - (C) Paired-t 檢定
 - (D) McNemar's 檢定
18. 下列有關相關係數的敘述中，何者正確？
- (A) 係數數值從-1 到 1
 - (B) 相關係數用作測量兩組類別數據的相關性
 - (C) 係數是否有統計意義，由其數值決定
 - (D) 一組連續性數據中每個數值減半，另一組數值不變，則係數數值會因而改變
19. 要評估兩個連續型變項間的相關性，若資料含有極端值(extreme value)，則應採用哪種統計相關係數較佳？
- (A) Kappa statistic
 - (B) Pearson correlation coefficient
 - (C) Kendall's tau coefficient
 - (D) Spearman correlation coefficient
20. 關於相關係數(r)與迴歸係數(b)兩者間的關係，下列敘述何者正確？
- (A) r 是指 X 解釋變數(explanatory variable)每增減一個測量單位，Y 反應變數(response variable)增減若干單位
 - (B) b 是指 X 解釋變數預測 Y 反應變數的準確度有多少
 - (C) r 是有單位的，b 是沒有單位的
 - (D) r 為正值，b 一定也為正值，反之亦然
21. 任兩個隨機變數(X,Y)，下列何者正確？
- (A) $\text{Var}(X+Y) = \text{Var}(X)+\text{Var}(Y)$
 - (B) $\text{Var}(X+Y) \geq \text{Var}(X)+\text{Var}(Y)$
 - (C) 若 $E[X]>E[Y]$ 則 $\text{Var}(X)>\text{Var}(Y)$
 - (D) 假設 X 和 Y 服從 Poisson 分佈，若 $E[X]>E[Y]$ 則 $\text{Var}(X)>\text{Var}(Y)$

22. 關於 Bernoulli(p)分配下列何者錯誤？
- (A) 只有兩種結果，成功或失敗
 - (B) 期望值為 p
 - (C) 當 $p = 0.5$ 時，實驗結果的變異是最小的
 - (D) 獨立的 Bernoulli 分配相加會形成 Binomial 分配
23. 對於母體的述敘，下列述敘何者為非？
- (A) 母體的平均數，通常以 μ 表示；母體的變異數，通常以 σ^2 表示
 - (B) 母體的參數，如： μ 、 σ^2 ，固定不變
 - (C) 當母體個數很龐大時，母體的參數常常是已知的
 - (D) 母體的定義，會隨著不同的研究主題而有所改變
24. 一個假說檢定的 p -value 為 0.999，下列何者正確？
- (A) 虛無假說為真
 - (B) 對立假說為真
 - (C) 虛無假說為真的機率是 0.999
 - (D) 虛無假說對於資料提供了一個相當可信的解釋
25. 以下何者為型一誤差(type-I error)？
- (A) 在虛無假設成立之下接受虛無假設
 - (B) 在虛無假設成立之下拒絕虛無假設
 - (C) 在對立假設成立之下接受虛無假設
 - (D) 在對立假設成立之下拒絕虛無假設
26. 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) 成對樣本 t 檢定
 - (D) ANOVA
27. 有關 Wilcoxon 等級和檢定(Wilcoxon rank sum test)方法之敘述，下列何者正確？
- (A) 無需假設測量值是常態分佈
 - (B) 樣本數小於 30 時則必須使用無母數方法
 - (C) 適用在成對樣本的檢定
 - (D) 適合用在兩組以上的樣本檢定

28. 今有一健康檢查資料，針對 10 名國小學童分別測量在一年級及三年級時的身高，資料如下(單位：cm)：

	1	2	3	4	5	6	7	8	9	10	平均數和標準差
一年級	102	113	112	116	115	114	97	119	114	111	$\bar{X}_1 = 111.3, S_1 = 6.7$
三年級	128	126	141	133	132	155	128	130	131	137	$\bar{X}_2 = 134.1, S_2 = 8.6$
	-26	-13	-29	-17	-17	-41	-31	-11	-17	-26	$\bar{X}_d = -22.8, S_d = 9.39$

研究目的為學童在兩年的時間中，身高是否有增加的現象。則檢定之虛無假設為兩年前後平均體重($\mu_1 - \mu_2$)的差異？(顯著水準 $\alpha = 0.05$)

- (A) 大於等於零
- (B) 小於等於零
- (C) 等於零
- (D) 以上皆非

29. 某研究隨機調查 100 位 12 歲女童的身高(X，公分)與體重(Y，公斤)，得到以下之迴歸方程式 $\hat{y} = -25 + 0.5x$ ，且此迴歸線的判定係數(coefficient of determination)為 0.64。下列敘述何者有誤？

- (A) 此組資料身高與體重的相關係數為 0.8
- (B) 12 歲女童的身高與體重有高度的線性正相關
- (C) 12 歲女童的身高可以解釋其體重變異量的 64%
- (D) 12 歲女童的身高每增加 10 公分，其平均體重增加 0.5 公斤

30. 某公司進行成本分析，於是做了一個工作時數(X)與產量(Y)的迴歸分析 $Y = \hat{\beta}_0 + \hat{\beta}_1 X$ ，及得到的結果如下表：

	估計值	標準誤
β_0	10	3
β_1	4	1

請問下列哪一個統計值可以拿來檢定工作時數是否為產量之重要解釋變數？

- (A) 10
- (B) 4
- (C) 3
- (D) 1