

CE5033 Statistical Methods and Data Mining

20240507 Exercise

1. 我們可以利用中央極限定理的結果，建構 μ 的95%信賴區間，稱為 (a, b) ，針對此信賴區間的描述，何者正確？

- (A) 表示 μ 有95%的機率會在 (a, b) 之間
- (B) 表示此 (a, b) 必會包含 μ
- (C) 建置 μ 的95%信賴區間，會比99%信賴區間來得長
- (D) (a, b) 區間愈長表示標準誤愈大，代表樣本數較小

(A)信賴區間不保證每次計算都會包含真正的母體平均值 μ 。95%的信賴區間意思是如果我們重複抽樣並計算無數次，大約95%的這些信賴區間會包含真實的 μ 。(B)信賴區間的寬度取決於標準誤，而標準誤反映了樣本大小的影響。標準誤越大，信賴區間越寬，通常表示樣本數較小。(C)99%信賴區間會比95%信賴區間來得長，因為要涵蓋更大的不確定性範圍，以包含母體平均值的機會更大。(D) 信賴區間的覆蓋概率是基於長期抽樣行為的概念，而不是單一樣本的機率敘述。意即不是說有95%的機率 μ 落在某個特定的區間內，而是在長期抽樣中，95%的信賴區間將包含 μ 。

2. 以下何者並非無母數檢定的優點？

- (A) 無母數檢定較不受離群點 (outlier) 影響而導致錯誤的統計推論
- (B) 無母數檢定可以分析質的資料 (qualitative data)
- (C) 對母體假設較少，並不一定要假設母體為常態分配

(D) 無母數檢定通常 power (即 $1-\beta$ ，其中 β 為型 II 誤差 (type II error) 的機率) 較大
無母數檢定通常具有比相應的母數檢定較低的檢定力，尤其是當數據真正符合母數檢定所假定的分布時。無母數檢定的主要優點在於它們對分布的要求較少，而非其檢定力較高。

3. 研究者欲了解服用避孕藥丸後對收縮壓是否會有影響，今15位婦女進入此研究，此15位婦女從未服用過避孕藥，研究者在給服藥之前，先測得每位婦女的收縮壓，給服藥兩週後再測得婦女的收縮壓值，以茲比較。試檢定服用避孕藥對收縮壓是否有影響，下列何種方法適用？

- (A) 成對樣本 t 檢定
- (B) 兩組獨立樣本 t 檢定
- (C) 兩組獨立樣本 Z 檢定
- (D) 單一樣本 Z 檢定

由於是在服用藥物前後對同一組婦女進行測量，因此適合使用成對樣本 t 檢定來評估前後差異。

4. 由樣本資料得到常態母體平均數的95%信賴區間 (confidence interval) 為 $(100, 120)$ ，若由同一筆資料得到母體平均數的90%信賴區間為 (a, b) ，則下列敘述何者正確？

- (A) $a < 100 < 120 < b$
- (B) $a < 100 < b < 120$
- (C) $100 < a < b < 120$
- (D) $100 < a < 120 < b$

90%信賴區間會比95%信賴區間窄，因此其範圍 (a, b) 必須完全包含在 $(100, 120)$ 之內。

5. 下列關於兩獨立樣本 t 檢定 (two-sample independent t test) 的敘述，何者正確？

- (A) 兩組樣本的平均值必須要相近
- (B) 兩組獨立樣本的變異數大小必須接近

(C) 兩組獨立樣本的大小必須相同 (D) 可以用來檢定左右眼近視度數的差別

(A)如果左眼和右眼的數據來自同一組人，則須使用配對樣本 t 檢定。(B)兩獨立樣本 t 檢定並無要求兩組樣本的平均值必須接近；實際上，這個檢定的目的就是要檢驗這兩個平均值是否有顯著的差異。(C)兩獨立樣本 t 檢定不要求兩組的樣本大小必須完全相同，相似的樣本大小可以增強檢定的統計能力，並使得結果更為穩健。(D)兩獨立樣本 t 檢定要求兩組數據的變異數應該是相近的，因為檢定統計量的計算基於這一假設。如果變異數顯著不同，則應該使用 Welch 的 t 檢定，此檢定不需要等變異數的假設。

6. 有關母體參數假設檢定的 p 值 (p-value)，下列敘述何者錯誤？

- (A) p 值的計算和顯著水準有關 (B) p 值的計算和虛無假設有關係
(C) p 值小，傾向於拒絕虛無假設 (D) p 值的計算和樣本的檢定統計量有關

p 值是檢定統計量給定的，和設定的顯著水準無關，顯著水準是用來與 p 值比較的閾值。

7. 某工廠生產的罐頭其 $\sigma=25$ ，今抽樣檢驗 100 個罐頭得其平均重量為 127.5 公克，當 $\Pr(Z \leq 1.28)=0.90$ ； $\Pr(Z \leq 1.645)=0.95$ 時，求該工廠罐頭重之 90% 信賴區間：

- (A) 86.4 ~ 168.6 (B) 124.3 ~ 130.7 (C) 95.5 ~ 159.5 (D) 123.4 ~ 131.6

給定平均重量為 127.5，標準差為 25，樣本大小為 100，使用 Z 分布計算信賴區間：

$$CI = \bar{x} \pm Z_{0.05} \times \frac{s}{\sqrt{n}}$$

其中， $\bar{x} = 127.5$ ， $Z_{0.05} = 1.645$ ， $s = 25$ ， $n = 100$ ，因此，該工廠罐頭重之 90%信賴區間 (123.39, 131.61)。

8. 假設母體呈常態分配，平均數 μ 未知。欲檢定 $H_0: \mu = 100$ vs. $H_1: \mu > 100$ ，顯著水準設為 0.01。若將型二錯誤 (type II error) 控制為 5%。當虛無假設 H_0 為偽，拒絕 H_0 的機率為何？

- (A) 0.99 (B) 0.95 (C) 0.05 (D) 0.01

控制型二錯誤為 5%意味著在對立假設是真的情況下，拒絕虛無假設的機率（也就是檢定力）為 95%。

9. 一般科幻小說平均 290 頁。某出版社隨機選擇他們出版的 16 部小說，其平均長度為 335 頁，標準差為 48 頁。欲檢定這出版社的小說是否明顯比一般科幻小說長，根據以上資料，得出結論為： $(t_{0.01,15} = 2.60, t_{0.05,15} = 1.75, t_{0.025,15} = 2.13)$

- (A) 若顯著水準(α)為 0.01，拒絕 H_0
(B) 若顯著水準(α)為 0.025，拒絕 H_0 ；若顯著水準(α)為 0.01，則不拒絕 H_0
(C) 若顯著水準(α)為 0.10，拒絕 H_0 ；若顯著水準(α)為 0.05，則不拒絕 H_0
(D) 若顯著水準(α)為 0.05，拒絕 H_0 ；若顯著水準(α)為 0.025，則不拒絕 H_0

H_0 : 出版社的小說比一般科幻小說長檢定統計量 $t_{df=16-1} = \frac{335-290}{48/\sqrt{16}} = t_{15} = 3.75$ ，查表知： $3.75 > 2.60, (t_{0.01,15}) \Rightarrow$ 拒絕 H_0 ；也就是說無論 $\alpha=0.1, 0.05, 0.025$ ，還是 0.01，結果都是拒絕 H_0 。

10. 下列有關型 I 錯誤機率、型 II 錯誤機率、樣本數與檢定力的敘述，何者正確？

- (A) 樣本數增加時，型 I 錯誤機率減少，檢定力減少
(B) 型 I 錯誤機率不變，則樣本數增加時，檢定力減少
(C) 樣本數固定時，型 I 錯誤機率增加，則型 II 錯誤機率也增加

(D)型 I 錯誤機率不變，則樣本數增加時，型 II 錯誤機率減少

(A)型 I 錯誤機率增加意味著拒絕虛無假設的閾值更寬鬆，因此型 II 錯誤機率會減少，因為更少的情況會錯誤地接受虛無假設。(B)增加樣本數提高了檢定的統計力，進而減少了型 II 錯誤的機率。(C)樣本數增加通常會增加檢定力，而不是減少。(D)樣本數的增加不會自動減少型 I 錯誤機率（由顯著水準 α 決定），且樣本數的增加會增加檢定力。

11. 罰球競賽中，某職籃選手罰 40 球中 26 球。假設該職籃選手罰球命中率為 p ，欲檢定 $H_0: p \leq 0.6$, $H_1: p > 0.6$ 。檢定統計量應該是多少？

- (A) 0.1021 (B) -0.1021 (C) 0.6455 (D) -0.6455

計算檢定統計量需要用到命中率的標準差和樣本大小。我們可以使用二項分布的性質來計算這個檢定統計量。使用公式

$$Z = \frac{p - p_0}{\sqrt{\frac{p_0(1 - p_0)}{n}}}$$

其中 $p = 26/40$ ， $p_0 = 0.6$ ， $n = 40$ 。

12. 變異數分析（Analysis of Variance）是用來檢定：

- (A) 數個母體的比率是否相同 (B) 數個母體的平均數是否相同
(C) 數個母體的變異數是否相同 (D) 數個母體之間是否獨立

變異數分析（ANOVA）主要用於檢定三個或更多群組的平均數是否存在顯著差異。

13. 對一母體平均數估得 95%雙尾信賴區間為（115，270），下列何者為正確？

- (A) 母體平均數在此區間內之機率為 0.95
(B) $H_0: \mu = 120$ vs. $H_1: \mu \neq 120$ ， $\alpha = 0.05$ ，不拒絕 H_0
(C) $H_0: \mu \leq 180$ vs. $H_1: \mu > 180$ ， $\alpha = 0.05$ ，拒絕 H_0
(D) $H_0: \mu \geq 100$ vs. $H_1: \mu < 100$ ， $\alpha = 0.05$ ，拒絕 H_0

因為信賴區間（115，270）包含了 120，所以沒有足夠證據在 $\alpha = 0.05$ 的水準下拒絕虛無假設 $\mu = 120$ 。

14. 一地區連續 60 個月的月平均濕度的平均值為 74%，中位數為 80%，眾數為 84%，則月平均濕度的分配形狀為：

- (A) 高峰偏右長尾向左延伸 (B) 高峰偏左長尾向右延伸 (C) 左右對稱 (D) 均勻分配

平均值小於中位數，中位數小於眾數：當眾數大於中位數，中位數又大於平均值時，這通常指示數據分佈是偏左的（左偏分佈）。這種分佈的形狀具有一個高峰（尖峰在眾數處），並向左（低值方向）延伸出一條較長的尾巴。尾部的方向：因為數據的低值端（小於平均值的部分）較多，這意味著分佈的左側尾部比右側尾部長。(A)描述了一個左偏分佈，也就是數據向左（低值）延伸的尾巴較長，平均值、中位數和眾數的排序指出數據分佈是向左偏斜的。(B)是對右偏分佈的描述，其中數據向右（高值）延伸的尾巴較長，與此問題的數據不符。(C)左右對稱意味著平均值、中位數和眾數應該接近或相等，顯然不適用於此問題。(D)均勻分配意味著所有數據點都大致相等，這在統計量明顯不同的情況下是不可能的。

15. 若 26 位參加統計學測驗之考生的成績皆不相同，在進行複閱之後，發現最高分者分數有誤，還要再加 2 分。試問以下統計量何者不受最高分者加分影響？

- (A) 算術平均數 (B) 標準差 (C) 中位數 (D) 全距

(A)算術平均數（平均值）是所有數值的總和除以數值的數量。在本案例中，只有一個數值改變（最高分加 2 分），這將改變總和，從而改變算術平均數。因此，算術平均數會受到影響。
(B)標準差是測量數據點與平均值的偏差的一種方式。當最高分者的分數增加時，平均數會變化，且該分數與新平均數的偏差也會變化。因此，這會影響到標準差的計算。所以，標準差會受到影響。
(C)位數是將數據點從小到大排列後位於中間的數值。在 26 位考生的情況下，中位數是第 13 位和第 14 位考生分數的平均值。因為修改的是最高分，除非最高分已經是中位數之一，否則這種改變不會影響排在中間的分數。在本案例中，中位數不受影響，因為只有最高分被改變。
(D)全距是最高分和最低分之間的差異。由於最高分增加了 2 分，這直接增加了全距的計算結果。因此，全距會受到影響。

16. 下列何者錯誤？

- (A) 當一組資料均為正數時，平均數、眾數、中位數及變異數均是正值
- (B) 若一組資料的變異數為零，則其平均數、眾數、中位數皆相等
- (C) 若分配是單峰對稱分配，則平均數、眾數、中位數皆相等
- (D) 若一組資料的全距愈大，則其眾數、中位數及平均數也會愈大

全距是數據中最大值和最小值之間的差距，全距的大小並不直接影響平均數、中位數或眾數的絕對大小。全距的增大可能是由於數據中新增了更大或更小的值，但這並不表示平均數、中位數或眾數必然增大。

17. 在進行 3 個處理，15 個樣本的單因子變異數分析，已知 $MSE = 5$ ， $SST(\text{total}) = 100$ ，則檢定處理平均數差異的 F 統計量值為何？

- (A) 2
- (B) 3
- (C) 4
- (D) 5

變異數分析的步驟：

- 1) 計算總變異數和（SST）：這已給出為 100。
- 2) 計算均方誤差（MSE）：這是誤差（或稱殘差）的均方，也就是不可解釋的變異的均方，已給出為 5。
- 3) 計算處理間均方（MSB）：使用 SST 和 SSE（均方誤差乘以自由度）來計算均方處理（ $MSB = SSB/df_{\text{between}}$ ）。

首先，我們需要計算處理間的平方和（SSB）： $SSB = SST - SSE$ ，其中，SSE 可以從 MSE 推導出來。MSE 的計算是： $MSE = SSE/df_{\text{error}}$ ，其中， df_{error} 是誤差的自由度，計算方式是： $df_{\text{error}} = n - k$ ，其中， n 是總樣本數， k 是處理組數。對於 15 個樣本和 3 個處理組： $df_{\text{error}} = 15 - 3 = 12$ ，所以， $SSE = MSE \times df_{\text{error}} = 5 \times 12 = 60$ ，然後， $SSB = SST - SSE = 100 - 60 = 40$ ，接下來，處理間的自由度（ df_{between} ）是： $df_{\text{between}} = k - 1 = 3 - 1 = 2$ ，因此，MSB（處理間均方）為： $MSB = SSB/df_{\text{between}} = 40/2 = 20$ ，最後，F 統計量計算為： $F = MSB/MSE = 20/5 = 4$ 。

18. 在單因子變異數分析中，Tukey 多重比較的目的為何？

- (A) 檢定所有母體是否具常態性
- (B) 檢定所有成對母體平均數是否有差異
- (C) 檢定所有母體平均數是否有差異
- (D) 檢定所有母體變異數是否有差異

(A) 檢定所有母體是否具常態性：這不是 Tukey 多重比較測試的目的。常態性檢定通常在進行 ANOVA 之前作為假設檢驗的一部分。
(B) 檢定所有成對母體平均數是否有差異：這正是 Tukey 多重比較測試的主要目的。在 ANOVA 確認至少存在一個平均值的差異之後，Tukey 測

試進行成對比較，以確定具體哪些組之間的平均數顯著不同。(C) 檢定所有母體平均數是否有差異：這是 ANOVA 的目的，而不是 Tukey 多重比較的目的。Tukey 測試是在 ANOVA 顯示存在差異後，進行更細節的成對比較。(D) 檢定所有母體變異數是否有差異：這與變異數同質性檢定相關，通常是進行 ANOVA 前的另一項假設檢驗，而非 Tukey 測試的目的。

19. 若右尾檢定的顯著水準 (α 值) 愈小，則以下何者為真？

- (A) 母體平均數愈大
- (B) 樣本平均數愈大
- (C) p 值 (p-value) 愈大
- (D) 臨界值 (Critical Value) 愈大

顯著水準 α 越小，意味著要求更強的證據來拒絕虛無假設。因此，臨界值 (例如 Z 值或 t 值) 會增大，以反映更嚴格的檢驗標準。

20. 假設隨機抽取一組樣本數等於 100 的樣本，發現樣本平均數是 10、樣本標準差是 10，那麼根據中央極限定理 (Central Limit Theorem)，上述樣本平均數的抽樣分配最接近下列那一個機率分配？

- (A) 有著平均數等於 0，變異數等於 1 的常態分配
- (B) 有著平均數等於 0，變異數等於 10 的常態分配
- (C) 有著平均數等於 10，變異數等於 1 的常態分配
- (D) 有著平均數等於 10，變異數等於 10 的常態分配

樣本平均數的抽樣分布：

平均數：抽樣分布的平均數是樣本平均數，即 10。

變異數：抽樣分布的變異數是原始數據的變異數除以樣本數量。原始數據的變異數是標準差的平方，即 $10^2 = 100$ 。因此，抽樣分布的變異數為 $100/100 = 1$ 。