

CE5033 Statistical Methods and Data Mining

20240423 Exercise

1. 下列關於單因子變異數分析 (one-way analysis of variance) 的敘述，何者錯誤？
- (A) 變異數分析的假設，其中之一為母群體變異數相等
 - (B) 單因子變異數分析用來比較三組的母群體平均數 μ_1 , μ_2 和 μ_3 是否相等時，虛無假設是 $\mu_1 = \mu_2 = \mu_3$
 - (C) 如果用變異數分析比較兩組平均數相等時，其結果會等於兩獨立樣本 t 檢定的結果
 - (D) 對立假設為母群體平均數不相等，亦即 $\mu_1 \neq \mu_2$ 且 $\mu_1 \neq \mu_3$ 且 $\mu_2 \neq \mu_3$

對立假設應該是至少有兩組的平均數不相等，而不是所有組間平均數都不相等。

2. 若隨機分配 X 其母體為平均數 μ 及標準差 σ ，今抽取出足夠大樣本數為 n 之樣本，則下列敘述何者錯誤？
- (A) 樣本平均數的標準誤與樣本數為 n 成反比
 - (B) 樣本平均數的標準誤與標準差 σ 成正比
 - (C) 樣本平均數會近似常態分配
 - (D) 樣本平均數的標準誤與平均數 μ 成正比

標準誤是獨立於平均數 μ 的，它依賴於樣本標準差 σ 和樣本大小 n。

3. 為教導幼兒認識性別，某位幼稚園的老師提出一份教材，該師於實驗前先擬好一份認識性別問卷給受試幼兒填寫，再於教授課程後給這群幼兒再做一次相同的問卷，問卷中有問及「性別人我關係」指標，令 μ_1 及 μ_2 分別表示性別人我指標的前測與後測平均值，以下對於推論敘述何者正確？
- (A) $H_0: \mu_1 = \mu_2$ 與 $H_1: \mu_1 < \mu_2$ ，對應檢定 p 值 < 0.0001 ，拒絕 H_0 表示性別人我指標的後測平均值沒有比較大
 - (B) $H_0: \mu_1 = \mu_2$ 與 $H_1: \mu_1 < \mu_2$ ，對應檢定 p 值 < 0.0001 ，拒絕 H_0 表示性別人我指標的後測平均值明顯較大
 - (C) $H_0: \mu_1 = \mu_2$ 與 $H_1: \mu_1 \neq \mu_2$ ，對應檢定 p 值 < 0.0001 ，拒絕 H_0 表示性別人我指標的前測與後測平均值相似
 - (D) $H_0: \mu_1 = \mu_2$ 與 $H_1: \mu_1 \neq \mu_2$ ，對應檢定 p 值 < 0.0001 ，拒絕 H_0 表示性別人我指標的後測平均值明顯不同

若拒絕 $H_0: \mu_1 = \mu_2$ 且對立假設 H_1 為 $\mu_1 < \mu_2$ ，表示後測平均值大於前測平均值；若拒絕 $H_0: \mu_1 = \mu_2$ 且對立假設 H_1 為 $\mu_1 \neq \mu_2$ 表示前後測試結果不相同。

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

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. 當估計母體平均數時，若減少抽樣個數為原本的一半，則平均數的標準誤 (Standard Error of the Mean) 有何改變？
- (A) 約降低至原平均數的標準誤數值之 50%

- (B) 約降低至原平均數的標準誤數值之 70%
- (C) 約增加至原平均數的標準誤數值之 1.4 倍
- (D) 約增加至原平均數的標準誤數值之 2 倍

標準誤與樣本大小的平方根成反比，減少到原本的一半，標準誤增加到原本的 $\sqrt{2}$ (約 1.4 倍)。

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

(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 。

7. 一群過重的成年人被隨機分派成兩組，一組進入減重計畫，一組作為對照組，半年之後用 t 檢定評估兩組受試者體重改變的差異，參加減重計畫的受試者體重平均減少 5 公斤，對照組平均減少 1 公斤，二組體重改變平均值差異的 95%信賴區間是 1 公斤到 7 公斤。下列敘述何者正確？
- (A)從二組體重改變平均值差異的 95%信賴區間，可以拒絕虛無假設
 - (B) t 檢定的 P 值大於 0.05，所以不能拒絕虛無假設
 - (C)雖然兩組體重減少的差異未達到統計顯著，但仍應拒絕虛無假設
 - (D)因為是前後測量，所以應該用成對樣本的 t 檢定

信賴區間不包含 0，表明平均體重改變存在統計顯著差異。

8. 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) 可以用來檢定左右眼近視度數的差別
 - (B) 兩組樣本的平均值必須要相近
 - (C) 兩組獨立樣本的大小必須相同
 - (D) 兩組獨立樣本的變異數大小必須接近

(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。

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

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。

10. 下列有關敘述性統計（Descriptive Statistics）的敘述，何者正確？

- (A) 左偏（left skewed）的分布，中位數（median）通常會比平均數（mean）來得小
- (B) 長條圖（Histogram）可以用來描述資料的分布情況
- (C) 中位數比平均數較容易受到極端值（outlier）影響
- (D) 臺灣受僱勞工的平均月收入是一個左偏（left skewed）的分布，因為有一些高收入所得者

(A)左偏（left skewed）分布意味著數據的尾部向左延伸，即較小的值比較稀疏，但極端小的值拉低了平均數的位置。在左偏分布中，平均數通常會比中位數小，不是中位數比平均數小。

(B)長條圖或直方圖是描述數據分布的一個基本工具，特別適用於顯示數據的頻率分布，可以幫助我們直觀地理解數據的集中趨勢、擴散程度以及偏斜方向。(C)中位數是一個比平均數更不容易受到極端值影響的統計量。由於中位數是將數據分為兩等分的值，它不直接計算數據的具體數值，因此對異常值的抵抗力較強。(D)當一個分布中存在一些遠高於大多數數據的極端值時，這將導致分布是右偏（right skewed）而不是左偏。右偏分布是由於少數極高的值將平均數拉向更高的範圍，而中位數和大多數數據則仍聚集在相對較低的範圍。

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

(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。

12. 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) -0.6455 (B) -0.1021 (C) 0.1021 (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 $p = 26/40$, $p_0 = 0.6$, $n = 40$ 。

13. 在某寒冷的冬天，連續 10 天的溫度都低於 0 度。則此 10 天溫度的標準差：

- (A) 因為每天的溫度都是負的，所以標準差是負的
- (B) 標準差大於或等於 0
- (C) 因為每天的溫度都是負的，所以標準差不能算
- (D) 標準差是可正可負的

標準差是衡量數據分散程度的統計量，其值始終非負。

14. 設一項有關平均數的區間估計，在信賴水準 95%之下為(2.5,3.1)。若換成假設檢定 $H_0: \mu = 3.5$ v.s. $H_1: \mu \neq 3.5$ ，則在顯著水準 5%之下，其結論應為：
- (A) 否決 H_0 (B) 不否決 H_0
(C) 無法做結論 (D) 以上(A)(B)(C)選項皆有可能

信賴區間 (2.5,3.1) 不包含假設檢定中的虛無假設值 3.5，因此在顯著水準 5% 下，我們拒絕虛無假設 H_0 。

15. 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。

16. 製作盒形圖 (box plot) 時，必須使用下列那一個統計量數？

- (A) 平均數 (mean) (B) 眾數 (mode)
(C) 中位數 (median) (D) 第 90 百分位數 (90-th percentile)

盒形圖中展示從第一四分位數到第三四分位數的範圍，中位數是這個盒子中的一條線，顯示資料的中心趨勢。

17. 紐約市的旅館每晚平均住宿價格為\$273。假定此估計值來自 45 家旅館所構成的樣本，同時知道樣本標準差為\$65。已知前年紐約市的旅館每晚平均住宿價格為\$229，請以 45 家旅館所構成的樣本來建構每晚平均住宿價格的 95%信賴區間，並用來判定這兩年的住宿價格的改變情況。下列何者正確？ ($t_{0.025,44} = 2.015$, $z_{0.025} = 1.96$)

- (A) 每晚平均住宿價格的 95%信賴區間為 (253,293)
(B) 每晚平均住宿價格的 95%信賴區間為 (228,280)
(C) 每晚平均住宿價格的 95%信賴區間為 (215,305)；這兩年的住宿價格起伏很大
(D) 每晚平均住宿價格的 95%信賴區間為 (195,226)；這兩年的住宿價格減少

給定平均值為\$273，標準差為\$65，樣本大小為 45，使用 t 分布計算信賴區間（因為樣本大小較小）：

$$CI = \bar{x} \pm t_{0.025,44} \times \frac{s}{\sqrt{n}}$$

其中， $\bar{x} = 273$ ， $t_{0.025,44} = 2.015$ ， $s = 65$ ， $n = 45$ ，因此，每晚平均住宿價格的 95%信賴區間為 (253,293)。

18. 某工業零件廠欲檢定其所生產零件規格是否符合客戶要求。假定其所生產零件規格服從常態分配，且利用 t 分配所得的信賴區間及檢定統計量來做關於零件長度規格平均值 μ 公分的統計推論。隨機抽檢 4 個零件，其所得的標準差為 2 公分，而 μ 的 95%信賴區間為[6.818,13.182]，即在 6.818 公分到 13.182 公分之間，下列敘述何者正確？

- (A) 如果假設為 $H_0: \mu = 6$ 對 $H_1: \mu \neq 6$ ，則在 5%的顯著水準下，結論是不拒絕虛無假設 H_0
(B) 如果假設為 $H_0: \mu = 7$ 對 $H_1: \mu \neq 7$ ，則 t 統計量值為 10
(C) 如果假設為 $H_0: \mu = 10$ 對 $H_1: \mu \neq 10$ ，則 p 值 (p-value) 為 1
(D) 如果樣本數增加至 16，且這 16 個零件長度的標準差亦為 2 公分，則樣本數 16 所得 μ 之 95%信賴區間寬度為原來樣本數 4 所得 μ 之 95%信賴區間寬度的一半

從信賴區間 (6.818, 13.182) 可知，樣本平均值是 10.0 公分，即為信賴區間的中點。(C)由於樣本平均值為 10，理論上 t 統計量值為 0。此時的 p 值為 1。(D)當樣本數增加時，信賴區間的寬度會受到兩個因素的影響：(1)標準誤：由於標準誤與樣本數的平方根成反比，因此標準誤會隨著樣本數增加而減少。(2)臨界值 (t 分數)：隨著樣本數的增加，自由度增加， t 分數會稍微減少。這兩個因素共同作用於信賴區間的寬度。假定原始樣本數為 4，標準差為 2 公分。使用 t 分數為 3.182 (自由度為 3, 95%信賴區間) 進行計算。當樣本數增加至 16，使用的 t 分數會變為 2.131 (自由度為 15, 95%信賴區間)。根據計算結果，當樣本數從 4 增加到 16 時，信賴區間的寬度確實變小，但變化比例並不完全是一半。具體來說：原始樣本數 4 的信賴區間寬度約為 6.364 公分；增加到樣本數 16 後的信賴區間寬度約為 2.131 公分。雖然信賴區間的寬度有所減少，但其減少的幅度並不完全等於原來的一半，而是大約為原來的三分之一。

19. 在統計學中針對各種不同種類的數據所劃分的尺度裡，下列那一個是屬於比例 (Ratio) 測量尺度？

- (A) 性別 (B) 身分證字號 (C) 出生年月日 (D) 薪資收入

薪資收入有絕對零點 (沒有收入) 且間隔均等 (薪資差異有意義，例如 \$50,000 與 \$100,000 間的差異等於 \$100,000 與 \$150,000 間的差異)。

20. 從一母體隨機抽出資料 x_1, x_2, \dots, x_n ，第一次抽出 100 個資料，也就是 $n=100$ ，而第二次抽出 10000 個資料也就是 $n=10000$ 。假設這兩組樣本其變異數的值非常接近 (其比值接近 1)，在相同信心水準 (Confidence Level) 下，利用這兩組樣本可做出兩個估計母體平均的信賴區間。則關於第一次抽樣做出信賴區間寬度對第二次抽樣做出信賴區間寬度的比值，下列敘述何者正確？

- (A) 兩信賴區間寬度的比值與信心水準的值有關
(B) 當信心水準是 0.95 時，兩信賴區間寬度的比值約是 100
(C) 假如原先變異數值有誤，修正後的變異數值比較大且兩組樣本之修正後的變異數值依然非常接近 (比值接近 1)，則兩個修正後的信賴區間其寬度的比值變大
(D) 當信心水準是 0.7 時，則兩信賴區間寬度的比值約是 10

(A)信賴區間寬度的比值主要取決於樣本大小的比值和臨界值的比值。因為臨界值 (對於常用的信心水準如 0.95) 基本上是固定的，所以主要是樣本大小的變化影響比值。(B)信賴區間寬度的比值應該是樣本大小的平方根的比值。因此，比值應為 10。(C)如果變異數修正後依然非常接近，則修正後的信賴區間寬度的比值不會變大；它仍然取決於樣本大小的比值。(D)無論信心水準如何變化，只要標準差保持一致，信賴區間寬度的比值主要受樣本大小影響，比值為樣本大小平方根的比值，即 10。