

# CE5033 Statistical Methods and Data Mining

## 20240430 Exercise

1. 針對統計檢定力 (statistical power,  $1-\beta$ ) 的敘述，下列何者較正確？

- (A) 當虛無假設是真的，卻沒有推翻虛無假設的機率
- (B) 當對立假設是真的，卻推翻虛無假設的機率
- (C) 當虛無假設為真的，卻推翻虛無假設的機率
- (D) 當對立假設是真的，卻沒有推翻虛無假設的機率

統計檢定力 ( $1-\beta$ ) 是指當對立假設實際上為真時，正確地拒絕虛無假設的機率。

2. 若右尾檢定的顯著水準 ( $\alpha$  值) 愈小，則以下何者為真？

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

顯著水準  $\alpha$  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。

3. 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  $\mu$  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  $\mu$  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  $\mu$  之 95%信賴區間寬度為原來樣本數 4 所得  $\mu$  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。

4. 若某組資料之標準差為 0，則下列敘述何者為真？

- (A) 資料分配成右偏分配
- (B) 中位數大於平均數
- (C) 資料中所有觀察值都相同
- (D) 資料中觀察值的數值，正負各佔一半

標準差為 0 表示所有的觀察值與平均值之間沒有任何差異，因此所有的觀察值必須是相同的。

5. 所謂“顯著水準”（level of significance）是指：

- (A) 最大可容忍型 II 錯誤（type II error）之機率 (B) 檢定力（power of test）  
(C) 信賴係數（confidence coefficient） (D) 最大可容忍型 I 錯誤（type I error）之機率

顯著水準  $\alpha$  是用來設定拒絕虛無假設的臨界值的標準，它代表的是在虛無假設實際為真時，錯誤地拒絕它的最大可容忍機率。

6. 在“騎機車不戴安全帽”實行處罰以前，我們調查發現不戴安全帽的比率是 80%。現在實施之後，我們想知道不戴安全帽的比率  $p$  是否變小，那麼檢定假設應該如何設？

- (A)  $H_0: p = 0.5$  vs.  $H_1: p \neq 0.5$  (B)  $H_0: p = 0.8$  vs.  $H_1: p \neq 0.8$   
(C)  $H_0: p = 0.8$  vs.  $H_1: p > 0.8$  (D)  $H_0: p = 0.8$  vs.  $H_1: p < 0.8$

設置虛無假設為沒有變化（ $p=0.8$ ），對立假設為比率下降（ $p<0.8$ ），這是最直接的方式來檢驗比率是否因政策實施而下降。

7. 續上題。若現在實際隨機調查 3000 人，發現不戴安全帽的比率是 20%，在顯著水準 1% 之下：（ $Z_{0.005} = 2.58$ ， $Z_{0.01} = 2.33$ ）

- (A) 拒絕  $H_0$  (B) 接受  $H_0$   
(C) 吾人不應該隨便做結論，應該繼續研究 (D) 以上皆可以

給定的數據顯示，在政策實施後，不戴安全帽的比率降至 20%。要決定是否拒絕虛無假設，可以利用比例檢定來評估這個變化是否在統計上顯著。由於這是一個左尾檢定（檢測比率是否下降），所以會查找  $-Z_{0.01} = -2.33$ 。由於這是一個左尾檢定（因為我們檢測比率是否下降），檢定統計量為

$$\frac{0.2 - 0.8}{\sqrt{(0.8 \times 0.2)/3000}} = -82.16 < -2.33$$

進行比例檢定證實此差異在統計上是顯著的，因此，拒絕  $H_0$  是正確的。

8. 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) 124.3 ~ 130.7 (B) 95.5 ~ 159.5 (C) 123.4 ~ 131.6 (D) 86.4 ~ 168.6

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）。

9. 在假設檢定中，下列何者為真？

- (A)  $\beta$  值與  $\alpha$  值互不影響 (B)  $\alpha$  值愈大，愈不容易拒絕原始假設  
(C)  $p$ -value 愈大，愈不容易拒絕虛無假設 (D)  $p$ -value 隨著  $\alpha$  值增加而減少

$p$  值是在虛無假設為真的前提下，觀察到當前或更極端結果的概率。 $p$  值越大，我們越缺乏證據來拒絕虛無假設。

10. 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)  $H_0: \mu \geq 100$  vs.  $H_1: \mu < 100$ ， $\alpha = 0.05$ ，拒絕  $H_0$

(C)  $H_0: \mu \leq 180$  vs.  $H_1: \mu > 180$ ,  $\alpha = 0.05$ , 拒絕  $H_0$

(D)  $H_0: \mu = 120$  vs.  $H_1: \mu \neq 120$ ,  $\alpha = 0.05$ , 不拒絕  $H_0$

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  $\mu = 120$ 。

11. 下面那一個不是用以測量離散度的測度量？

- (A) 標準差      (B) 眾數      (C) 全距      (D) 四分位距

眾數是數據集中出現次數最多的值，並不是用來測量數據的離散度。標準差、全距和四分位距都是描述數據分佈的離散性的常用統計量。

12. 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) 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

(B) 兩信賴區間寬度的比值與信心水準的值有關

(C) 當信心水準是 0.7 時, 則兩信賴區間寬度的比值約是 10

(D) 當信心水準是 0.95 時, 兩信賴區間寬度的比值約是 100

信賴區間的寬度主要由標準誤決定, 而標準誤與樣本大小的平方根成反比。因此, 當從  $n = 100$  到  $n = 10000$ , 信賴區間寬度的比值應等於這兩個樣本大小平方根的比值, 即  $\sqrt{10000/100} = 10$ 。

13. 檢定母體平均數是否為 10 時, 何種情況下不會拒絕虛無假設？

(A) 母體平均數的信賴區間包含 10

(B) 檢定統計量值落在拒絕域

(C) 顯著水準大於  $p$  值

(D) 樣本平均數等於 15

如果信賴區間包含了虛無假設指定的母體平均數值 (本例為 10), 則沒有足夠證據拒絕虛無假設。

14. 在統計檢定問題中, 有關「型 I 誤差」與「型 II 誤差」的敘述何者為正確？

(A) 型 I 誤差與型 II 誤差和為 1.00

(B) 型 I 誤差永遠大於型 II 誤差

(C) 檢定力 (power) 與型 I 誤差和為 1.00

(D) 增加樣本數可同時降低型 I 誤差與型 II 誤差

增加樣本數可以提高檢定的精確度, 進而降低錯誤接受虛無假設的風險 (型 II 誤差) 以及錯誤拒絕虛無假設的風險 (型 I 誤差)。

15. 以下何者並非無母數檢定的優點？

(A) 對母體假設較少, 並不一定要假設母體為常態分配

(B) 無母數檢定較不受離群點 (outlier) 影響而導致錯誤的統計推論

(C) 無母數檢定通常 power (即  $1 - \beta$ , 其中  $\beta$  為型 II 誤差 (type II error) 的機率) 較大

(D) 無母數檢定可以分析質的資料 (qualitative data)

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。

16. 請判斷下列的研究資料分析方法適用於配對資料或是雙樣本資料：

- (i) 品名藥 (generic drug) 與原廠藥 (reference drug) 的藥效研究。受測者隨機分為兩組，第一組服用品名藥 24 小時後測量血液中藥含量。第二組則服用原廠藥 24 小時後測量血液中藥含量。比較服用品名藥與原廠藥後血液中藥含量的差異。
- (ii) 鈣對血壓影響的研究。受測者隨機分為實驗組 (服用鈣片) 及對照組 (服用安慰劑)，一週後比較兩組血壓改變量。
- (iii) 身體的電磁場對皮膚癒合影響的研究。將隨機選出的水蜥兩隻以刀割相同長度的傷口，隨機選擇右或左足以電極改變傷口的磁場，另一隻則以半數電極量改變傷口的磁場，兩小時後比較傷口癒合率。

(A) (iii) 適用於配對資料分析，(i) (ii) 適用於雙樣本資料分析

(B) (i) (iii) 適用於配對資料分析，(ii) 適用於雙樣本資料分析

(C) (ii) 適用於配對資料分析，(i) (iii) 適用於雙樣本資料分析

(D) (i) (ii) (iii) 都適用於雙樣本資料分析

(i) 品名藥與原廠藥的研究：受試者被分為兩組進行不同治療，這是典型的雙樣本設計。

(ii) 鈣對血壓影響的研究：受試者同樣被分為實驗組和對照組，這也是雙樣本設計。

(iii) 電磁場對皮膚癒合的研究：在同一個生物體上的兩個不同部位進行不同處理，這是典型的配對資料設計。

17. 若信賴水準為不變，樣本大小增加 3 倍，則常態信賴區間的長度是原來信賴區間長度的多少倍？

- (A) 3      (B) 1/3      (C) 1.732      (D) 0.577

信賴區間的寬度與樣本大小的平方根成反比。增加樣本大小至原來的三倍，信賴區間的寬度將變為原來的  $1/\sqrt{3} \approx 0.577$  倍。

18. 颱風來襲時，桃園市政府依據颱風是否經過桃園市來決定要不要放假，若認為會經過則會放假。若虛無假設與對立假設如下，下列何者是型 I 錯誤？

$H_0$ ：颱風會經過桃園市

$H_1$ ：颱風不會經過桃園市

- (A) 颱風不會經過桃園市，也沒放假      (B) 颱風不會經過桃園市，卻放了假  
(C) 颱風會經過桃園市，卻沒放假      (D) 颱風會經過桃園市，也放了假

型 I 錯誤是虛無假設 ( $H_0$ ) 實際為真時，錯誤地拒絕它。這裡虛無假設是颱風會經過桃園市，型 I 錯誤發生在預測颱風會經過但實際上沒有。

19. 為了解市面上販賣的飲料中某物質 X 的平均值，隨機抽取 36 杯飲料的樣本，這 36 杯飲料 X 的平均值為 10。假設 X 在母群體及此樣本的標準差剛好皆為 4。下列敘述何者錯誤？

- (A) 如果只抽取 20 杯，X 的平均值分布未知  
(B) X 的平均值近似常態分布  
(C) 標準誤 (Standard Error) 為 0.667

(D) 標準誤 (Standard Error) 為量化 X 分散程度的敘述性統計指標

標準誤衡量樣本平均值估計總體平均值的準確性，而不是資料本身的離散度。

20. 在 95% 的信心水準下，某航空公司每天取消航班的比率的信賴區間是(2.5%, 3.9%)，試問每天取消航班比率的點估計值為何（假設為大樣本調查）？

(A) 0.03      (B) 0.032      (C) 0.039      (D) 0.025

信賴區間(2.5%, 3.9%)的中點提供了點估計的合理估計，計算方法為  $(2.5\% + 3.9\%) / 2 = 3.2\%$  或 0.032。