

一、選擇題 (每題 3 分，共 10 題，合計 30 分) ※ 注意：請於試卷內之「選擇題作答區」依序作答。

1. 以下敘述何者最不恰當?
 - (A) 常態分配跟 t 分配一樣，都是對稱的分配。
 - (B) 不同的常態分配彼此可以互相變換；不同的 t 分配彼此也可以互相變換。
 - (C) 自由度為 15 的 t 分配的變異數比標準常態分配還大。
 - (D) 自由度為 103 的 t 分配可以被標準常態分配逼近(approximate)。
2. 以下敘述何者最不恰當?
 - (A) 假設 X 服從二項式(Binomial)分配， $\text{Binomial}(n, p)$ ，其中 n 是總共的試驗數，p 是每次試驗的成功機率。如果 n 很大，p 很小，np 趨向定值，則 X 可以利用一個布阿松分配(Poisson distribution)來逼近。
 - (B) 布阿松分配(Poisson distribution)可以用來描述台灣南部地區罹患兒童白血病的人數。
 - (C) 假設 X 服從布阿松分配(Poisson distribution)，參數為 λ ，則 X 的變異數為 2λ 。
 - (D) 假設 X 服從布阿松分配(Poisson distribution)，參數 $\lambda=1$ ，則 $X=0$ 的機率比 0.3 大。
3. 假設 X 服從常態分配，期望值為 μ ，變異數為 1，而 $\{x_1, x_2, \dots, x_{100}\}$ 為從母體隨機抽取之樣本。請問以下敘述何者最不恰當?
 - (A) $\{x_1, x_2, \dots, x_{100}\}$ 的樣本平均值是 μ 的最大概似估計值(maximum likelihood estimator)。
 - (B) $\{x_1, x_2, \dots, x_{100}\}$ 的樣本平均值是 μ 的不偏(unbiased)估計值。
 - (C) $\{x_1, x_2, \dots, x_{100}\}$ 中的第一個觀察值是 μ 的不偏(unbiased)估計值。
 - (D) $X=3$ 的機率會大於 0.1。
4. 以下有關 p 值的敘述何者最不恰當?
 - (A) p 值越小表示檢定越顯著。
 - (B) p 值是一種機率。
 - (C) 計算 p 值時需要假設虛無假說為真。
 - (D) p 值不是顯著水準。
5. 以下敘述何者最恰當?
 - (A) 卡方分配是對稱的分配。
 - (B) 卡方檢定可以用來進行適合度(goodness-of-fit)檢定。
 - (C) 卡方檢定可以用來進行兩個連續變數之間的相關性(association)檢定。
 - (D) 卡方分配的隨機變數的值的範圍與常態分配一樣。
6. 假設全台大學生的血壓分布為常態分配，某研究想探討全台大學生的血壓平均值 μ 是否為 104 毫米汞柱，現隨機抽取 100 位台大學生並測量其血壓，得平均值 96 毫米汞柱，標準差 16 毫米汞柱，請問以下何者最接近 μ 的 95% 信賴區間?
 - (A) (64, 128)
 - (B) (80, 112)
 - (C) (92.8, 99.2)
 - (D) (94.4, 97.6)
7. (延續上題)有關以上的信賴區間，以下敘述何者最恰當?
 - (A) 這個信賴區間有 0.95 的機率會包含 μ 。
 - (B) 這個信賴區間有至少 0.95 的機率會包含 μ 。
 - (C) 這個 95% 的信賴區間的寬度會比 99% 的信賴區間窄。
 - (D) 這個 95% 的信賴區間的寬度與信心水準有關，每次抽樣所得的寬度都會一樣。

見背面

8. (延續上題)根據以上數據，以下敘述何者最恰當?
- (A) 全台大學生的血壓平均值 μ 有可能是 104 毫米汞柱。
 - (B) 全台大學生的血壓平均值 μ 不太可能是 90 毫米汞柱。
 - (C) 該研究對 μ 的估計誤差為 16 毫米汞柱。
 - (D) 這 100 個樣本的分布圖會是對稱的分布。
9. (延續上題)若使用統計檢定方法，在顯著水準 0.05 時回答該研究問題，以下敘述何者最不恰當?
- (A) 虛無假說為全台大學生的血壓平均值 μ 是 104 毫米汞柱。
 - (B) 這個檢定會是雙尾檢定。
 - (C) 檢定的結論會拒絕虛無假說。
 - (D) 檢定的結論會是，有不到 0.05 的機率虛無假說為真。
10. 某研究想探討規律運動一年之後對體重的影響，假設體重服從常態分配，現隨機抽取十位過重民眾，測量其規律運動一年之前與之後的體重，請問以下敘述何者最不恰當?
- (A) 應該使用成對樣本的 t (paired-t) 檢定，不能使用兩組獨立樣本的 t 檢定。
 - (B) 這個檢定的自由度為 $10+10-2=18$ 。
 - (C) 若是畫出運動之前的體重的盒鬚圖(boxplot)，與運動之後的體重的盒鬚圖，沒辦法看出每個人的體重變化。
 - (D) 這 20 筆資料間的相關存在於每個人的兩筆體重測量值之間，不存在於人與人之間。

二、計算簡答題 (共 3 題，合計 50 分)

1. (20 分) 陳先生想競選美麗市市長，某民調公司隨機抽樣 900 位民眾並詢問是否願意投票給陳先生，其中 180 位回答願意。請問：
- (1) 根據上述資料，該民調公司估計陳先生的得票率與估計誤差應為多少? (3 分)
 - (2) 請根據上述資料，檢定在顯著水準 0.05 時，陳先生的得票率能否超過 25%?(請寫出虛無假說、對立假說、統計量、及結論) (4 分)
 - (3) 根據上述資料，該民調公司對陳先生的得票率的 95%信賴區間為何? (3 分)
 - (4) 請問上述信賴區間的計算是否使用了中央極限定理?請說明為什麼? (5 分)
 - (5) 王小姐也想競選美麗市市長，該民調公司也為王小姐進行了同樣目的的抽樣調查，隨機抽樣 400 位民眾並詢問是否願意投票給陳先生，其中 100 位回答願意。請問這兩位候選人的得票率的差異的 95%信賴區間為何? (5 分)
2. (15 分) 菁英電子公司新任的 CEO 為了增進公司的營運績效，要求人力資源部門馬上規劃人事調整方案。有經驗的人資部經理請你先做一份問卷調查，瞭解不同位階員工對人事調整的看法。下表為不同位階員工對於人事調整看法的次數分配資料，請你選用合適的統計方法進行分析，並協助人資部經理判斷是否「可以立即規劃人事調整方案」，或是「應當先請 CEO 深入瞭解員工對人事調整的看法，再做後續規劃」。請使用 5%顯著水準，進行統計分析並說明理由。

[註： $Z_{0.05}=1.645$; $t_{0.05, 4}=2.132$; $\chi^2_{0.05, 4}=9.488$; $F_{0.05, 2, 2}=19$]

員工位階	對公司人事調整的看法			合計
	贊成	反對	無意見	
生產線工人	180	60	60	300
工廠領班	20	20	20	60
管理階層	5	25	10	40
合計	205	105	90	400

3. (15 分) Toyota 是台灣汽車銷售量第一品牌，為提升女性的購買率，該公司擬透過抽樣調查了解台灣女性對汽車顏色與車型大小的偏好。因此，銷售部門先選出紅、灰、白、黑四種顏色(Color)與大、中、小三種車型(Size)，在 Toyota 的旗艦店展出結合上述特性的 12 輛新車，而後在參觀者中依系統隨機抽樣方法選出 24 位女性，請她們依其個人喜好為每一輛新車評分。下表是根據調查數據執行變異數分析(ANOVA)的結果，請回答以下問題：
- (1) 變異數分析表格的內容仍有缺漏，請寫出下表(A)~(E)的數值。(10 分)
 - (2) 請引用表中合適的數據，並自行決定顯著水準，說明台灣女性對於 Toyota 汽車在顏色、車型的偏好特性。(5 分)

Analysis of Variance					
Source	df	SS	MS	F	p-value
Color	3	2711.17	(A)	(B)	0.000
Size	2	579.00	289.50	23.16	0.000
Color*Size	(C)	150.33	25.06	(D)	0.144
Error	12	150.00	12.50		
Total	(E)	3590.50			

三、統計報表解讀 (20 分)

1. 佳佳健身中心開授的減重課程極為熱門，參與的學員都需配合進行體重管理，藉以瞭解學員在參加課程之前的 BMI 值 (Kg/m^2) 和課程結束後減去的體重 (BW, 單位：公克) 之間是否有關係。該中心管理人員隨機選取 70 名學員的 BMI 與 BW 資料進行簡單迴歸分析，統計分析的結果如下，請據此回答以下兩個問題：
 - (1) 請引用合適的數據，列出迴歸方程式，並說明參與減重課程學員的 BMI 值與其減去的體重有何實質關係。(10 分)
 - (2) 請估計參與減重課程學員 BMI 值之迴歸係數的 95% 信賴區間，並解釋其意義。(10 分)

The REG Procedure					
Model: MODEL1					
Dependent Variable: BW					
Number of Observations Read				70	
Number of Observations Used				70	
Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	4628670	4628670	30.62	<.0001
Error	68	10280044	151177		
Corrected Total	69	14908714			
Root MSE		388.81501	R-Square	0.3105	
Dependent Mean		2924.28571	Adj R-Sq	0.3003	
Coeff Var		13.29607			
Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > t
Intercept	1	1873.62795	195.48294	9.58	<.0001
BMI	1	27.32022	4.93741	5.53	<.0001