

長庚大學114學年度碩士班考試招生筆試試題

系所：資訊管理學系碩士班

考試科目：統計學

注意：請詳細閱讀下列試題，並請標明題號依試題順序將答案書寫於答案卷上。 本試題共 6 頁：第 1 頁

一、單選題（每題 4 分，共 80 分）

1. 有個資料集包含了從 2013 年至 2024 年當中，天母棒球場每場球賽觀眾人數的時間序列。請問下列哪種圖表最能正確呈現人數趨勢？
 - (a) 圓餅圖
 - (b) 人數和時間點的散佈圖
 - (c) 柏拉圖
 - (d) 次數分配圖。
2. 依據資料整理結果和敘述性統計分析所得結論，某研究推斷母體不符合常態分佈。如果要用小樣本來檢定母體的中心位置(measure of center)，換句話說，要應用樣本中心值落點判定母體中心位置的可能值，以下哪個敘述較為恰當：
 - (a) 用 t-檢定做檢定
 - (b) 用中位數來做檢定，例如無母數的方法
 - (c) 把資料轉換成 Z 值，然後用 z-檢定做檢定
 - (d) 以上皆非。
3. 某教育背景之研究者執行一項實驗研究。研究者從一個學區隨機選擇 38 所學校，並對這 38 所學校的所有教師進行訪談。請問該研究者是使用何種抽樣方法？
 - (a) 群集抽樣 (cluster sampling)
 - (b) 系統隨機抽樣法
 - (c) 分層抽樣 (stratified sampling)
 - (d) 以上皆非。
4. 已知某個母體平均值之 98%信賴區間的範圍為 $22.0 < \mu < 24.0$ 。針對虛無假設 $H_0: \mu = 21.0$ ，在 1%的顯著水準下，結論是_____。以下哪個陳述最能正確地總結上句？
 - (a) 無法推翻此虛無假設，因為 21.0 是小於 24.0
 - (b) 推翻此虛無假設，因為 21.0 不等於 23.0.
 - (c) 推翻此虛無假設，因為 21.0 是小於 22.0
 - (d) 無法下結論。
5. 相對於有母數的分析方法(parametric method)來說，請問哪一項是無母數分析方法(nonparametric method)的相對優點？
 - (a) 樣本數很小和母體不能假設為常態分佈時
 - (b) 資料是次序尺度(ordinal scale)的資料類別
 - (c) (a)(b)皆正確
 - (d) (a)(b)皆非。

長庚大學114學年度碩士班考試招生筆試試題

系所：資訊管理學系碩士班

考試科目：統計學

注意：請詳細閱讀下列試題，並請標明題號依試題順序將答案書寫於答案卷上。

本試題共 6 頁：第 2 頁

6. 關於型一錯誤(type 1 error)的描述，請問哪一個選項最正確？
 - (a) 型一錯誤機率值等於本假設檢定之顯著水準(significance level)
 - (b) 此假設檢定的統計檢定力(power of the test)是等於(1 - 型一錯誤機率值)
 - (c) 型一錯誤機率值等於本假設檢定之樣本推導出的 p 值
 - (d) 以上皆非。
7. 迴歸模型中，特定自變數(given x)對應到的應變數(y')的周圍的預測區間是 _____。下列哪一個敘述最能補充上面的句子？
 - (a) 特定 x 的某段區間值用於預測 y'
 - (b) 用於計算迴歸線斜率的 y 值
 - (c) y 值和 x 值之間的差
 - (d) 與 x 相對應的 y 值的真實平均值的信賴區間。
8. 下列關於多元線性迴歸分析(Multiple regression analysis)的敘述何者正確？
 - (a) 如果增加足夠多的自變數，即使變數在模型中沒有特定意義，也有可能會讓本迴歸模型獲得接近 1 的 R^2 值。
 - (b) 如果在迴歸模型中加入越多的自變數， R^2 的值一定會逐步的提高。
 - (c) 為了要尋找最有能力的預測變數 y ，建模過程一定要選最大 R^2 值的模型，因為此迴歸模型自變數能解釋最大的應變數變異。
 - (d) 以上皆是。
9. 下列關於單因子變異數分析(1-ANOVA)的敘述何者正確？
 - (a) 在單因子變異數分析(1-ANOVA)中，若虛無假設被拒絕，則我們可以得知所有母體平均數皆不相同之結論(即沒有任何兩個母體平均數相等)。
 - (b) 對於三個或以上的組別，單因子變異數分析(1-ANOVA, F-test)可以顯示所有母體平均數之間是否存在差異，以及差異所在之處。
 - (c) 單因子變異數分析(ANOVA)檢定可以是右尾檢定、左尾檢定或雙尾檢定，這取決於對立假設。
 - (d) 在單因子變異數分析(1-ANOVA)中，誤差均方(the error mean square)用於衡量樣本內的變異性，而處理均方則衡量樣本平均數之間的變異性。
10. 如果某迴歸模型的相關係數(correlation coefficient)是等於 0.820，那此迴歸模型之應變數的變異中無法由自變數解釋部分所占的比例是等於：
 - (a) 82.0%
 - (b) 67.24%
 - (c) 32.76%
 - (d) 18.0%。

長庚大學114學年度碩士班考試招生筆試試題

系所：資訊管理學系碩士班

考試科目：統計學

注意：請詳細閱讀下列試題，並請標明題號依試題順序將答案書寫於答案卷上。

本試題共 6 頁：第 3 頁

11. 下列哪個陳述最適合描述卡方適合度檢定(Chi-square Goodness of Fit)？
 - (a) 檢驗觀察次數遠小於相應的預期次數，因此它是左尾檢定
 - (b) 檢驗觀察次數與相應的預期次數足夠相似或相等，因此它是雙尾檢定
 - (c) 檢驗觀察次數遠大於相應的預期次數，因此它是右尾檢定
 - (d) 以上皆非。
12. 某廠牌手電筒電池的平均運行時間為 9.7 小時。製造商對生產方法進行更改，並希望執行假設檢驗以確定新電池的平均運行時間是否因此增加。電池廠商針對上述的假設檢定，隨機抽生產線上的 100 個新電池來做實驗(在標準的測試機台上測試每個電池的使用時間)，以 3% 的顯著水準檢驗新電池是否具有更長的平均壽命。本樣本得到平均 9.9 小時的結果。使用 SPSS 的單一獨立樣本 T 檢定來做分析結果，算出 0.05 雙尾的顯著機率值。使用 p-值(p value)的決策方式，本假設檢定的結論會是：
 - (a) 接受 H_0
 - (b) 推翻 H_0
 - (c) 無法做決定
 - (d) 以上皆非
13. 某市場調查想了解消費者對四種品牌的速食料理包有無特定偏愛。市場調查人員隨機抽取 800 個人之後，邀請測試者個別使用及評價這些產品，並記錄他們對這四個品牌的評價排名(1 代表最不满意，4 代表最滿意)。請問下列何者為正確分析方法？
 - (a) 適合用變異係數分析來做
 - (b) 適合使用卡方檢定來做
 - (c) 適合用迴歸分析來做
 - (d) 適合用 t-檢定來做。
14. 針對兩母體平均值的檢定，下列敘述何者正確？
 - (a) 若兩個獨立樣本的個數皆大於 30，則假設變異數一樣的獨立樣本 t-檢定之結論，將與假設變異數不一樣的獨立樣本 t-檢定產生不同結論
 - (b) 若兩組樣本個數相同並樣本個數皆大於 30，則獨立樣本 t-檢定將與成對樣本 t-檢定的檢定結果相同
 - (c) 以上皆非
 - (d) 以上皆是。
15. 汽車製造商希望測試新火花塞與傳統火花塞並進行比較。製造商隨機抽取 50 輛排氣量(cc.)差不多但品牌不同的汽車進行傳統火花塞測試，另隨機抽取 50 輛車型大小差不多但是品牌不同的汽車用於新火花塞測試。每輛汽車都填滿 10 升汽油，並記載其行駛里程數。下列何者最適合用來檢驗新火花塞是否優於傳統火花塞？
 - (a) Wilcoxon 符號秩檢定(Wilcoxon signed rank test)
 - (b) 成對樣本 t-檢定
 - (c) 變異數不一樣或一樣的獨立樣本 t-檢定
 - (d) 以上皆可。

長庚大學114學年度碩士班考試招生筆試試題

系所：資訊管理學系碩士班

考試科目：統計學

注意：請詳細閱讀下列試題，並請標明題號依試題順序將答案書寫於答案卷上。

本試題共 6 頁：第 4 頁

16. 下列哪一個選項能有效提高假設檢定的統計檢定力(power of the test)？
- (a) 透過拒絕不良的樣本可以提高統計檢定力
 - (b) 透過降低樣本大小可以提高統計檢定力
 - (c) 透過添加樣本，直到達成所需的顯著水準，來提高統計檢定力
 - (d) 透過增大顯著水準或增加樣本大小可以提高統計檢定力。
17. 某城市經普查發現，當地單人單日的平均生活費是\$1200，標準差是\$135。依據柴比雪夫定理 (Chebyshev's theorem)，從多少到多少會包含 81.1%可能觀察到的單人單日的平均生活費，請問以下哪個結論是最接近正確的？
- (a) \$930.5 ~ \$1470.5
 - (b) \$889.5 ~ \$1510.5
 - (c) \$795.5 ~ \$1605.5
 - (d) 以上皆否，無法計算。

題組 18-20

某製造廠想要調查了解不同生產線組合以及製程方法會不會在生產績效上產生差異。假設製造廠有三條不同生產線組合與兩種不同製程方法要測試平均產量是否不同。實驗採用了平衡實驗設計(各組觀察值都要相等)，總共收集到 24 筆的觀察產量資料進行此分析，用雙因子變異數分析(2-ANOVA)套用 5%的顯著水準來檢驗。將雙因子變異數分析(2-ANOVA)得到的表(輸出表摘自於 SPSS 報表)如下表，

主旨間效果檢定				
因變數：生產績效				
來源 (Source)	第 III 類平方和 (SS)	自由度 (D. F.)	平均值平方 (Mean Square)	顯著性 (P-值)
截距	9168648.167	1	9168648.167	.000
生產線	519407.583	(乙)	(己)	.000
生產方法	726.000	(丙)	(庚)	.837
交互作用	125071.750	(丁)	(辛)	.043
錯誤(組內)	(甲)	(戊)	(壬)	
校正後總數(總和)	944657.833	23		

請回答下列問題：

18. 關於報表內的(丁)值，請選出正確選項：

長庚大學114學年度碩士班考試招生筆試試題

系所：資訊管理學系碩士班

考試科目：統計學

注意：請詳細閱讀下列試題，並請標明題號依試題順序將答案書寫於答案卷上。

本試題共 6 頁：第 5 頁

- (a) 2
 - (b) 3
 - (c) 6
 - (d) 以上皆非。
19. 哪個選項依序顯示出報表內(乙)和(丙)值？
- (a) 1 和 2
 - (b) 2 和 3
 - (c) 3 和 2
 - (d) 2 和 1
20. 雙因子變異數分析(2-ANOVA)需不需要各組母體之標準差需相等？本次檢定沒有交代有沒有滿足本條件，因此會不會影響到結論？下列關於雙因子變異數分析(2-ANOVA)的回答何者正確？
- (a) 雙因子變異數分析在設計上太複雜，所以不需要各組母體之標準差需相等。因此不會影響到結論。
 - (b) 雙因子變異數分析需要組母體之標準差需相等。因此會影響到結論。
 - (c) 本實驗採用了平衡實驗設計，所以可以忽略各組母體之標準差需相等。因此不會影響到結論。
 - (d) 以上皆非。

長庚大學114學年度碩士班考試招生筆試試題

系所：資訊管理學系碩士班

考試科目：統計學

注意：請詳細閱讀下列試題，並請標明題號依試題順序將答案書寫於答案卷上。

本試題共 6 頁：第 6 頁

二、計算題（共 20 分）

1 延續上述的題組 18-20

1.1 請針對報表的變數寫出並解釋本雙因子變異數分析想要驗證的完整數學模型(model)

1.2 請依據報表的 p-值(顯著差異)寫出本雙因子變異數分析想要驗證的虛無及對立假說(hypothesis)與分析後的結論

1.3 請問在報表錯誤列的(甲)、(戊)和(壬)的值分別為何？請寫出計算過程

2 某研究想要針對職業及退休年齡之間的關聯性進行研究。實驗設計針對 8 大職業類別及對應的退休年齡年齡層(退休年齡劃分為 4 個不同的層)。為了檢驗這不同的職業的平均退休年齡和職業之間是否存在關係，從最近退休員工資料集中隨機抽樣挑選 2000 個樣本收集並記錄到列聯表(contingency table)中。用 5% 的顯著水準進行假設檢驗，數據是否提供了足夠的證據來得出退休年齡取決於職業的結論。

2.1 本次統計檢定的假設為？

2.2 會用哪種檢定方法進行此研究？

2.3 本次檢定的自由度(degree of freedom)是多少？

2.4 假設上述檢定結果是顯著的，我們是否可以進一步得出結論，某項工作的退休人員比另一項工作的年齡大得多，因此人們更喜歡該工作？為什麼我們能或不能得出這個結論？解釋你的答案。