

※ 注意：請於試卷內之「選擇題作答區」依序作答。

單選題（每題 5 分）

- 下列哪一項資料散佈(dispersion)指標較不受極端值影響？
 - 全距(range)。
 - 標準差(standard deviation)。
 - 平均絕對離差(mean absolute deviation)。
 - 四分位差(interquartile range)。
 - 變異係數(coefficient of variation)。
- 有關盒鬚圖(box-and-whisker plot)，下述何者錯誤？
 - 中位數位於盒子的中間位置。
 - 盒子的寬度(或高度)代表數據的四分位差。
 - 鬚的長度代表數據中的極端值範圍。
 - 若資料呈右偏分佈，則盒子的右側(或上側)通常較長。
 - 盒鬚圖可用於顯示資料的集中趨勢與散佈情況。
- 有關隨機變數(random variable)及機率分佈(probability distribution)，下述何者錯誤？
 - 兩個獨立的(independent)連續型(continuous)隨機變數相加，仍為連續型隨機變數。
 - 兩個獨立的離散型(discrete)隨機變數相減，仍為離散型隨機變數。
 - 離散型隨機變數，其可能結果僅可為整數。
 - 離散型機率分佈函數，其函數值即為機率。
 - 連續型機率分佈須透過機率密度函數(probability density function)的積分，才可求得機率。
- 有關常態分佈(normal distribution)，下述何者錯誤？
 - 常態分佈乘以正數，仍為常態分佈。
 - 常態分佈乘以負數，仍為常態分佈。
 - 兩個獨立的常態分佈相加，仍為常態分佈。
 - 兩個獨立且相同(identical)的常態分佈相減，仍為常態分佈。
 - 兩個獨立的標準常態分佈(standard normal distribution)相乘，為常態分佈。
- 有關二項式分佈(binomial distribution)及布阿松分佈(Poisson distribution)，下述何者錯誤？
 - 兩個獨立且相同的二項式分佈相加，仍為二項式分佈。
 - 兩個獨立但不相同的布阿松分佈相加，仍為布阿松分佈。
 - 二項式分佈與布阿松分佈若期望值(expected value)相同時，布阿松分佈有較大的變異數(variance)。
 - 二項式分佈與布阿松分佈皆為離散型分佈。
 - 二項式分佈與布阿松分佈皆為右偏(positively skewed)分佈。
- 有關卡方分佈(chi-square distribution)，下述何者錯誤？
 - 自由度為 1 的卡方分佈的平方根，即為標準常態分佈。
 - 兩個獨立的卡方分佈相加，仍為卡方分佈。

見背面

- (C) 卡方分佈為連續型分佈。
- (D) 卡方分佈為右偏分佈。
- (E) 卡方分佈的變異數為期望值的兩倍。

是非改錯題 (每題 5 分)

7. 當信賴區間的長度變短時，可能是由於顯著水準下降或標準誤減少造成的。
答案：
解釋：
8. 在雙尾檢定中，當實際效應接近零時，使用更大的樣本數會增加拒絕虛無假說的可能性。
答案：
解釋：

計算問答題

9. 一項研究欲比較三種不同飲食方案對減重的效果。研究中，40 名受試者被隨機分派到三組，分別採用 (A 組) 低碳水化合物飲食、(B 組) 低脂肪飲食、與 (C 組) 高蛋白飲食。研究結果顯示，三組的平均體重減輕如下：
- A 組：平均減重 5.2 ± 1.5 kg, $n_A = 13$
 - B 組：平均減重 4.8 ± 1.2 kg, $n_B = 14$
 - C 組：平均減重 6.0 ± 1.4 kg, $n_C = 13$
- 研究者使用單因子變異數分析(One-way ANOVA)檢驗這三種飲食方案在減重的效果上是否有顯著差異。
- (1) 請問進行 ANOVA 的前提假設有哪些？(5 分)
 - (2) 請問 ANOVA 的虛無假說 (H_0) 和對立假說 (H_1) 分別是什麼？(3 分)
 - (3) 為什麼 ANOVA 檢定適用於比較三組平均值，而不是進行多次 t 檢定？(2 分)
 - (4) 請簡述 ANOVA 檢定的流程。(8 分)
 - (5) 假設 ANOVA 檢定的結果為 $F = 4.72$, $p\text{-value} = 0.015$ 。請解釋 F 值和 p-value 值的意義，而在顯著水準為 $\alpha = 0.05$ 下，你會如何下結論？(5 分)
 - (6) 請依據事後(post-hoc)檢定顯示，A 組和 C 組之間有顯著差異，但 B 組與其他組無顯著差異，請問應如何解釋這些結果？(2 分)

接次頁

某傳染病研究學者想探討氣候變遷對登革熱病媒蚊分布的影響。在研究中，他利用迴歸分析探討 15 年來高雄與台南地區登革熱流行期間，降雨情況與布氏指數的關係，結果如下表。請回答 10-14 題。

迴歸統計	
R ²	0.8803
adjusted R ²	0.8792
標準誤	8.2669
觀察值個數	420

ANOVA

	自由度	SS	MS	F 統計	顯著值
迴歸	4	208635.834	52158.958	763.208	<0.0001
殘差	415	28361.829	68.342		
總和	419	236997.663			

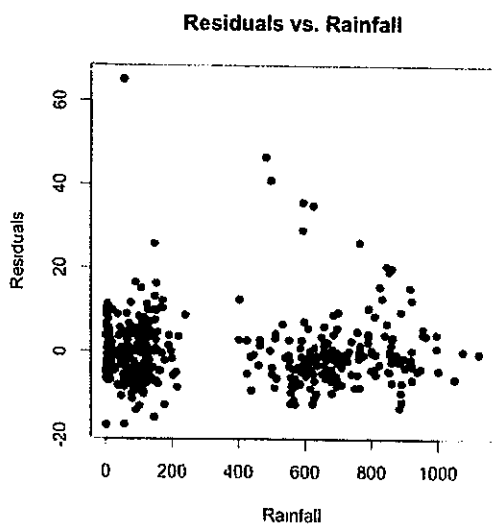
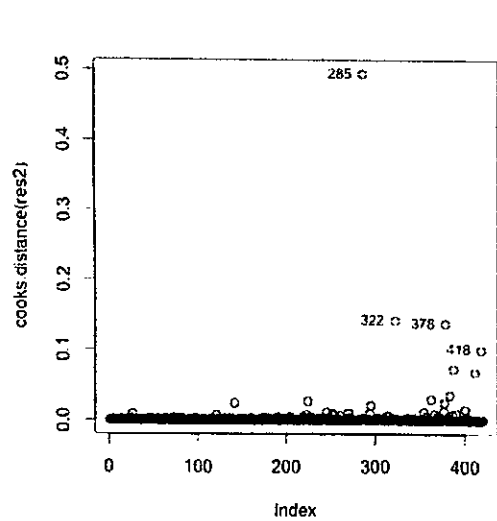
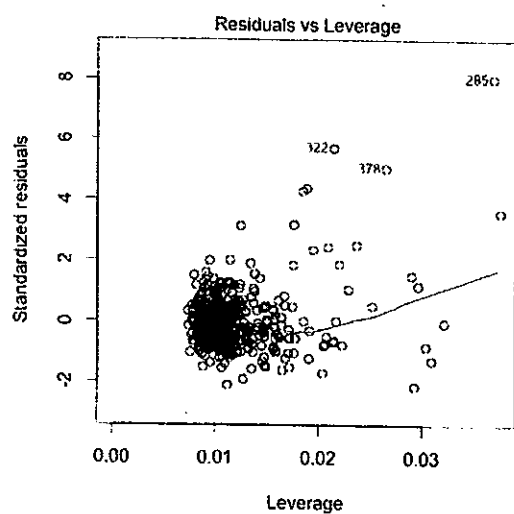
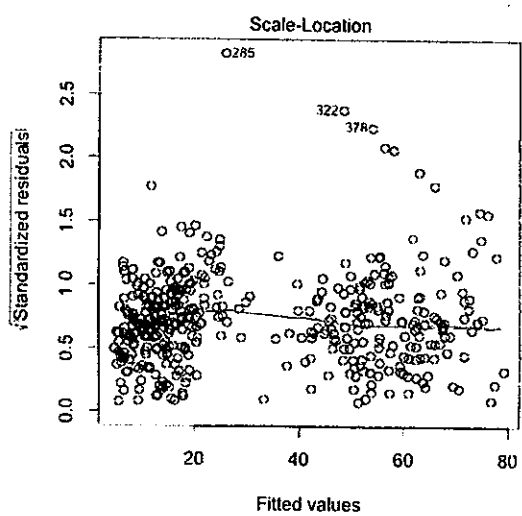
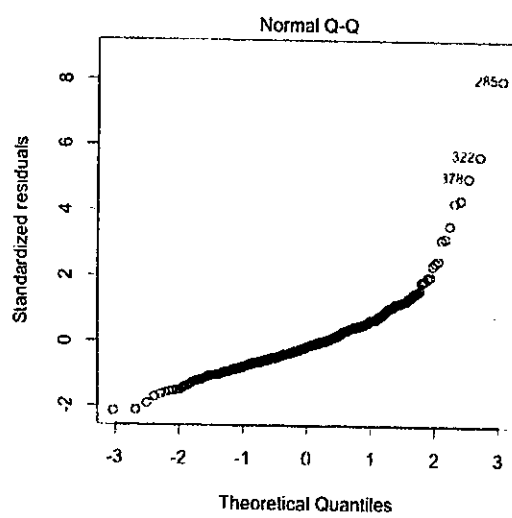
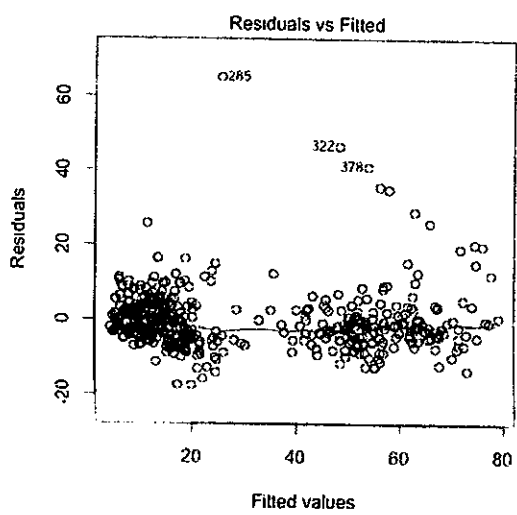
	係數	標準誤	t 統計	P-值	95% 信賴區間
截距	11.264	1.899	5.932	<0.0001	(7.532, 14.997)
臺南 vs. 高雄	-2.879	0.807	-3.567	0.0004	(-4.465, -1.292)
山地 vs. 平地	-1.524	0.807	-1.888	0.0597	(-3.110, 0.063)
降雨量(mm)/月	0.056	0.002	35.596	<0.0001	(0.053, 0.059)
降雨天數/月	0.895	0.093	9.604	<0.0001	(0.712, 1.078)

10. (4%) 請問說明 ANOVA table 中的 F 統計量的虛無假說(null hypothesis)及對立假說(alternative hypothesis)為和，並解釋檢定結果。
11. (4%) 請解釋「降雨天數/月」的迴歸係數的 95%信賴區間(confidence interval)所代表的意義。
12. (4%) 請說明模型中 R²所代表的意義。
13. (15%) 請判斷下列關於線性迴歸模型(linear regression model)的說明，正確與否，有錯誤者請更正之。
 - (A) 線性迴歸模型中「線性」的意思是指自變數(independent variable) X 與依變數(dependent variable) Y 之間的關係是線性。
 - (B) 線性迴歸模型中假設自變數(independent variable) X 要服從常態分配(normal distribution)。
 - (C) 根據分析結果，山地鄉的布氏指數會比平地地區少 1.524。
 - (D) 此分析結果的迴歸模型為：布氏指數 = 11.264 - 2.879 × 臺南 - 1.524 × 山地 + 0.056 × 降雨量 + 0.895 × 降雨天數
 - (E) 因為增加放入模型中自變數(independent variable) 的數量，可以提高模型的 R²，所以建立迴歸模型時，為了讓模型有更高的解釋力或更好的配適能力(goodness-of-fit)，有多少自變

見背面

數就放入多少，不會有任何問題。

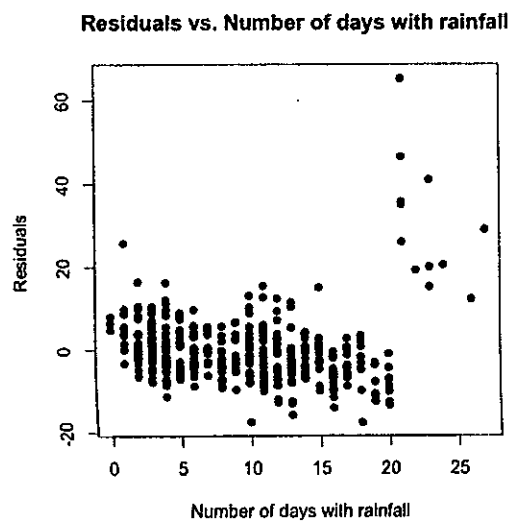
14. (8%) 請利用下圖對此迴歸模型進行診斷，並說明此模型是否有任何問題或是違反任何線性迴歸分析的假設。



接次頁

題號： 283 國立臺灣大學114學年度碩士班招生考試試題
科目：生物醫學統計學
節次： 7

題號：283
共 5 頁之第 5 頁



試題隨卷繳回