

繳交截止日期: 2020/04/17(五), 24:00

授課教師: 吳漢銘 (臺北大學統計學系副教授)

請仔細閱讀每一個注意事項 (禁止討論)

1. 寫作業要點

- (a) 可參考課本、上課講義 (包含電子檔) 及其它資料。
- (b) 儘量不要與別人 (或同學) 討論, 自己做, 不可參考同學的答案, 不可抄襲。
- (c) 程式設計題, 若程式碼直接複製 (或照抄) 講義上的以不給分為原則。
- (d) 有問題者, 請發 e-mail 或 FB 私訊問助教或老師。
- (e) 不按照規定作答者, 酌量扣分。
- (f) 請參照下列文件第 2 ~ 4 頁寫作規定, 不按照規定作答者, 會扣分。

<http://www.hmwu.idv.tw/web/teaching/doc/R-how-homework.pdf>

2. 下載題目卷, 上傳答題檔案:

- (a) 於課程網站下載題目卷。
- (b) 上傳答題檔案: 於教師網站首頁登入 [作業考試上傳區], 帳號: dm108。密碼: xxx (上課教室號碼)。
- (c) 請上傳「學號-姓名-DM-HW1.docx」。(目錄: 「20191030-exam1」。)

3. 答題檔案原則:

- (a) 請依照「R 程式作業繳交方式」, 複製 Console「程式執行及結果」至答案卷。圖形複製, 請注意大小, 內容數字文字需可辨識。
- (b) 程式設計題, 若程式碼直接複製 (或照抄) 講義上的以不給分為原則。
- (c) 若上傳檔案格式錯誤, 內容亂碼, 空檔等等問題。請自行負責。
- (d) 若要重覆上傳 (第 2 次以上), 請在檔名最後加「-2」、「-3」, 例如: 「學號-姓名-DM-HW1-2.docx」、「學號-姓名-DM-HW1-3.docx」等等。
- (e) 上傳兩次 (含) 以上、格式不合等等酌量扣分。
- (f) 如果上傳網站出現「You can modify the html file, but please keep the link 'www.wftpservers.com' at least.」, 請將滑鼠移至「網址列」後, 按「Enter」即可。若再不行, 請換 (IE/Edge/Firefox/Chrome)。

我已經仔細閱讀上述各注意事項, 若有違背, 會自行負責。

統計圖形、探索性資料分析

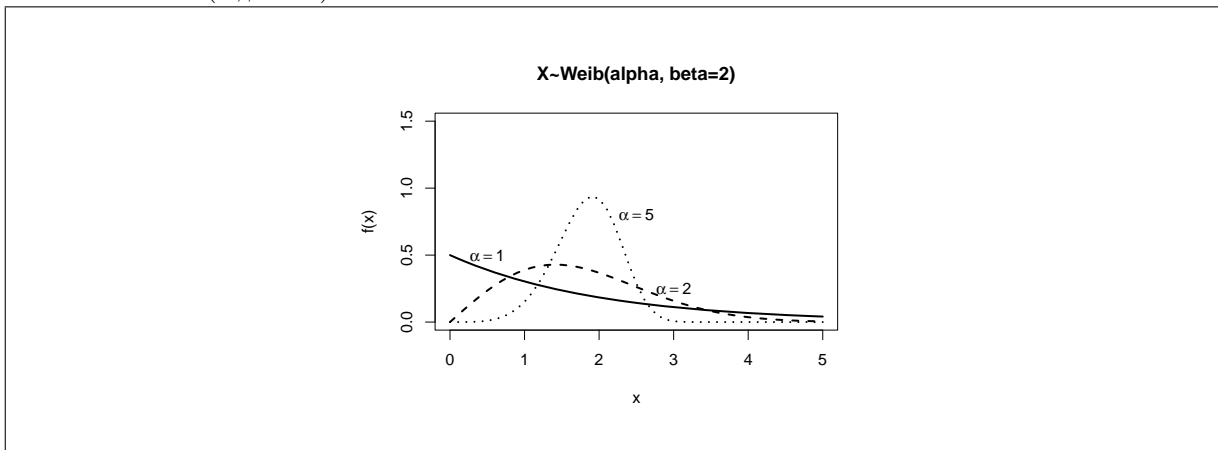
1. (統計圖形)

若隨機變數 X 服從 Weibull 分配 (簡記為 $X \sim Weib(\alpha, \beta)$)，其機率密度函數為

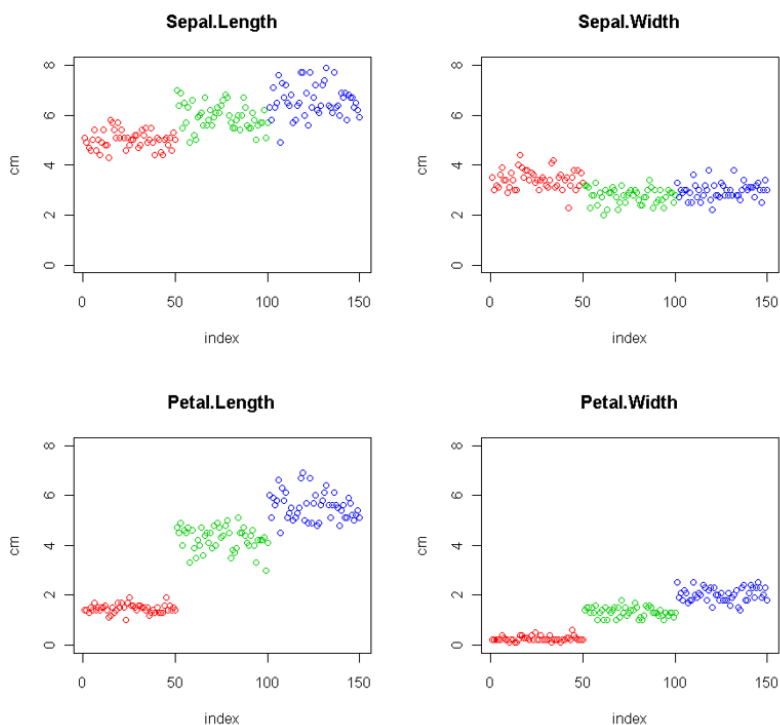
$$f(x|\alpha, \beta) = \alpha\beta^{-\alpha}x^{\alpha-1}e^{-(x/\beta)^\alpha}, \quad x > 0.$$

(a) 若 `x <- seq(0, 5, 0.1)`，寫一函式計算 $f(x|\alpha=1, \beta=2)$ 之值。

(b) 畫出 $X \sim Weib(\alpha, \beta=2)$ 之圖形如下：



2. (統計圖形) 利用 `iris` 資料畫出下圖 (4 圖一頁):



3. **(探索性資料分析)** 資料來源: 政府資料開放平台。資料檔: 不動產實價登錄資訊-買賣案件。<https://data.gov.tw/dataset/26820>
- (a) 下載資料，並讀入 RStudio。印出資料摘要 (`summary`) 及結構 (`str`)。請確認每一變數 (欄位) 皆是正確的 R 類別 (例如: 數值變數、日期變數，文字的「一、二、三...」等等轉成數值型的「1,2,3,...」)。若不是請做必要的類別轉換。
 - (b) 進行探索性資料分析。(畫出各式統計圖 (包含索引圖、直方圖、長條圖、盒形圖、餅圖、2D 散佈圖、3D 散佈圖、熱圖、地圖等等)(需對圖形結果作一些簡單解釋。)) (請自行編子題號: 2.1, 2.2, ...)
 - (c) 依此資料探索之結果，列舉一些想了解或解決的問題。這些問題需要其它的輔助資料嗎? 若有請說明。
4. **(資料摘要)** 呈上題之「不動產實價登錄資訊-買賣案件」資料。依 `district`(鄉鎮市區)，計算此資料各個連續變數之平均。例如: 依各 `district`(鄉鎮市區)，變數「`rps22`(單價每平方公尺)」之平均為何。

注意: 上傳檔案之後，請刪除作答目錄 及答案卷，清空資源回收筒，關機。交回題目卷。