

日期: 2020/11/10(二) 上機考: Open Book
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. 考試期間

- (a) 請按照平時上課之座位入座。若同學右邊有人，則該同學應使用隔板 (或拿包包擋著)。
- (b) 可參考課本、上課講義 (包含電子檔) 及其它資料，但不能與別人討論。
- (c) 可使用計算機、自己的筆記型電腦及平板電腦，不可使用手機。
- (d) **全程可上網查詢，但不能用通訊軟體 FB/LINE 等討論。**
- (e) 有問題者，請舉手發問。勿與同學交談。
- (f) 不按照規定作答者，酌量扣分。
- (g) 不可使用它人之隨身碟。「作弊」或「疑似作弊」，往後各項考試不予評分。

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: 於課程網站上登入 [作業考試上傳區]，帳號: r1091。密碼: xxxxx。
- (c) 上傳答題檔案時，請注意「正確目錄」。
- (d) 答題是從「Console」複製程式執行過程及結果，然後貼到答題案。
- (e) 若傳錯，請最終要上傳一份正確的答題檔案。
- (f) 請上傳「學號-姓名-R-exam-extra.txt」。(學號及姓名，改成自己)
- (g) 若上傳檔案格式錯誤，內容亂碼，空檔等等問題。請自行負責。
- (h) 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-R-exam-extra-2.txt」等等。
- (i) 上傳兩次 (含) 以上、格式不合等等酌量扣分。

3. 完成考試

- (a) 上傳完畢，確認無誤，請刪除作答目錄 及答案卷，清空資源回收筒，並關機。即可離席。

我已經仔細閱讀上述各注意事項，若有違背，會自行負責。

R: 基礎物件、輸入/輸出

1. 產生數列:

(a) 1 4 7 10 13 16 19

(b) $1, -\frac{1}{2}, \frac{1}{3}, -\frac{1}{4}, \dots, \frac{1}{19}, -\frac{1}{20}$

2. 讀取下列檔案，列印出資料前 5 筆，及後 5 筆紀錄; 同時檢查 (印出) 資料每一變數 (欄位) 是否有符合 R 的類別物件，若沒有，請更改。(http://www.hmwu.idv.tw/web/R/data/)

(a) R-score.xlsx

(b) 20140714-weather.txt

(c) weather_delays14.csv

3. 「statlog_vehicle_846x18.txt」是以 tab 為分隔的資料，具有 18 個變數，請讀入 R 之後，列出資料框維度、前後各 5 筆紀錄及儲存此資料框物件所佔用的記憶體。(原始資料說明: [https://archive.ics.uci.edu/ml/datasets/Statlog+\(Vehicle+Silhouettes\)](https://archive.ics.uci.edu/ml/datasets/Statlog+(Vehicle+Silhouettes))。)

4. 某銷售人員在 2018 年的網路銷售紀錄從公司資料庫隨機抽樣 10 筆如下:

```
Dates: 0924, 1112, 1231, 1105, 0604, 0219, 0416, 0611, 0813, 1029
Time: 01:00, 04:00, 16:00, 23:00, 08:00, 09:00, 07:00, 17:00, 03:00, 14:00
Items: shirt, shirt, pants, jacket, jacket, shirt, jacket, jacket, shoes, shirt
Volume: 7951, 159,1958, 6848, 3762, 3678, 8696, 9045, 6208, 1425
```

(a) 請將上述資料儲存成一資料框 (data.frame) 類別之物件，命名 mySale，使得第一個欄位為銷售日期時間 (DateTime)，類別為 POSIXct，時區為世界協調時間 (UTC); 第二個欄位為銷售品項 (Items)，類別為 factor; 第三個欄位為銷售量 (Volume)，類別為 numeric。印出 mySale。

(b) 本資料中，七月 (含) 之後的銷售品項為何？其總銷售量為多少？

5. 有一班學生之座號 (ID) 及性別 (student.gender) 的資訊如下。某日小考兩科: 微積分 (score.calculus) 及英文 (score.english)，成績如下，其中有三位同學缺考。

```
set.seed(12345)
ID <- paste("No.", 1:50, sep="")
score.calculus <- sample(0:100, 50, replace=T)
score.english <- sample(0:100, 50, replace=T)
student.gender <- as.factor(sample(c("f", "m"), 50, replace=T))
absence.id <- sample(1:50, 3)
score.calculus[absence.id] <- score.english[absence.id] <- NA
```

(a) 算出微積分平均分數及標準差。(提示: (1) 缺考不計入; (2) ?mean)

(b) 男生英文成績平均多少分? (提示: 缺考不計入)

(c) 將缺考成績記為 0 分後，請問有哪些同學兩科成績同時及格? (列出座號)

(d) (承上小題) 兩變數 $(x, y)_{i=1}^n$ 的相關係數之公式如下:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}$$

試計算微積分及英文兩成績之相關係數，並與 `cor` 之結果相比較。(提示: `sqrt`, `sum`, `mean`)

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。