

日期: 2020/10/27(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-exam1.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-exam1-2.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

1. 造出以下數列 (提示: `month.abb`):

(a) 1 2 3 4 5 6 2 3 4 5 6 3 4 5 6 4 5 6 5 6 6

(b) "Jan" "Mar" "May" "Jul" "Sep" "Nov" "Feb" "Apr" "Jun" "Aug" "Oct" "Dec"

2. 兩向量 $\vec{u} = (u_1, u_2, \dots, u_p)$ 及 $\vec{v} = (v_1, v_2, \dots, v_p)$ 的 cosine 夾角 θ ，可用下列式子表示:

$$\cos \theta = \frac{\vec{u} \cdot \vec{v}}{\|\vec{u}\| \|\vec{v}\|},$$

其中 $\vec{u} \cdot \vec{v} = \sum_{i=1}^p u_i v_i$, $\|\vec{u}\| = \sqrt{\sum_{i=1}^p u_i^2}$ 和 $\|\vec{v}\| = \sqrt{\sum_{i=1}^p v_i^2}$ 。現若有兩向量 $\vec{u} = (3, 4, 0)$ 和 $\vec{v} = (4, 4, 2)$ ，試計算此兩向量之的 cosine 夾角 ($\cos \theta$)。

3. 空間上有兩平面 $I_1: a_1x + b_1y + c_1z + d_1 = 0$ 及 $I_2: a_2x + b_2y + c_2z + d_2 = 0$ ，其法線 (normal) 分別為 $\vec{n}_1 = (a_1, b_1, c_1)$ 及 $\vec{n}_2 = (a_2, b_2, c_2)$ 。若 $\vec{n}_1 = k\vec{n}_2$ ，亦即 $\frac{a_1}{a_2} = \frac{b_1}{b_2} = \frac{c_1}{c_2} = k$ ，則我們稱此兩平面平行。而此兩平行的平面，其距離公式為:

$$D = \frac{|a_2x_0 + b_2y_0 + c_2z_0 + d_2|}{\sqrt{a_2^2 + b_2^2 + c_2^2}},$$

其中 $P_0 = (x_0, y_0, z_0)$ 為平面 I_1 上之一點。若兩平面分別為 $I_1: 2x + 3y + 4z - 3 = 0$ 及 $I_2: -4x - 6y - 8z + 8 = 0$ ， $P_0 = (0, 0, 3/4)$ 為平面 I_1 上之一點，試回答下列各問題:

(a) 將 \vec{n}_1 、 \vec{n}_2 及 P_0 表示為 R 向量類別的物件。

(b) 利用 \vec{n}_1 、 \vec{n}_2 ，說明平面 I_1 和 I_2 平行。

(c) 利用 \vec{n}_1 、 \vec{n}_2 及 P_0 計算此兩平行的距離。

4. 李克特量表 (Likert Scale) 的五等測量法是根據陳述語句的傾向給予各等級不同分數。對正向陳述而言，答案越正向分數越高例如：「非常同意」為 5 分，「同意」為 4 分，「普通」為 3 分，「不同意」為 2 分，「非常不同意」為 1 分。今有一問卷資料集，為一群學生 (男生、女生) 對某門課教師評分是否公平的認同程度 (consent)，如下:

```
set.seed(1234567)
n <- 60
ID <- sample(1:n) #座號
gender <- sample(c("男", "女"), n, replace=T) # 性別
consent <- sample(c("非常不同意", "不同意", "普通", "同意", "非常同意"),
                  n, replace=T, prob=c(0.1, 0.1, 0.2, 0.4, 0.3)) # 認同程度
```

試回答下列各問題:

(a) 將性別轉為 R 因子 (factor) 類別。男女生各多少人?

(b) 將認同程度轉為 R 有順序的因子 (factor) 類別。印出

- (c) 將上述資料存成一 R 資料框 (`data.frame`)，命名為 `survey.df`，欄位名稱依序為座號、性別及認同程度。印出此資料的前 5 筆紀錄。印出此資料框的結構。
- (d) 列出填寫「不同意」(含) 以下的學生座號，共有幾人。(註: 使用「`<=` 或 `>=`」)
- (e) 此次調查結果，認同程度平均為多少分?

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。