

108 學年度第一學期
電腦概論與程式設計: 作業 (5) 第 1 頁/共 3 頁

繳交日期: 2019/11/25(一), 24:00 前
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) 問問題，請多利用課程助教。
- (c) 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。
- (d) 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。
- (e) 有問題者，請發 e-mail 或 FB 私訊問助教或老師。
- (f) 不按照規定作答者，酌量扣分。
- (g) 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. 上傳答題檔案:

- (a) 於課程網站上登入 [作業考試上傳區]，帳號: r108。密碼: xxxxx。
- (b) 上傳答題檔案時，請注意「正確目錄」。
- (c) 若傳錯，請最終要上傳一份正確的答題檔案。被移至 {fail} 資料夾的檔案一律不批改。
- (d) 請上傳「學號-姓名-R-HW5.txt」。(學號及姓名，改成自己的)
- (e) 若上傳檔案格式錯誤，內容亂碼，空檔等等問題。請自行負責。
- (f) 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-HW5-3.txt」等等。
- (g) 上傳兩次 (含) 以上、格式不合 (答題檔案內容格式不對、學號多或少一碼、檔名多一空格等等) 酌量扣分。

我已經仔細閱讀上述各注意事項，若有違背，會自行負責。

R: 自定函式

1. 某地購買物品需加收增值稅 (VAT)，而增值稅會根據所物品類別不同而有不同之稅率，今稅率表如下：

類別	產品	VAT
印刷品類	書, 雜誌, 報紙等等	8%
食物類	蔬菜, 肉品, 飲料等等	10%
衣服類	T 恤, 牛仔褲, 上衣等等	20%

某人買了下列 5 樣物品 (括弧內數字為不含稅之花費金額): {書 (50 元)、肉品 (200 元)、上衣 (299 元)、牛仔褲 (1200 元)、飲料 (20 元)}。(提示: %in%, ifelse)

- (a) 造出以下資料框 (data.frame)。

	品項	價格	類別	VAT
1	書	50	印刷品類	0.08
2	肉品	200	食物類	0.10
3	上衣	299	衣服類	0.20
4	牛仔褲	1200	衣服類	0.20
5	飲料	20	食物類	0.10

- (b) 請計算此次消費所付出之總稅額。

2. 卡方獨立性檢定適用於分析兩組類別變數 (X, Y) 是否為相互獨立。例如: 薪資的高低和學歷的程度是否獨立? 婚姻狀況與宗教信仰是否有關? 通常資料會表示成 $r \times c$ 的列聯表 (Contingency Table)(格式如下) , 其中 $\{A_1, A_2, \dots, A_r\}$, $\{B_1, B_2, \dots, B_c\}$ 分別為兩組類別變數 (X, Y) 之類別, 其中 O_{ij} 為 ($X = A_i, Y = B_j$) 觀察次數, n 為樣本數:

(X, Y)	B_1	B_2	\dots	B_j	\dots	B_c	列總和
A_1	O_{11}	O_{12}	\dots	O_{1j}	\dots	O_{1c}	$O_{1\cdot}$
A_2	O_{21}	O_{22}	\dots	O_{2j}	\dots	O_{2c}	$O_{2\cdot}$
\vdots	\vdots			\vdots		\vdots	\vdots
A_i	O_{i1}	O_{i2}	\dots	O_{ij}	\dots	O_{ic}	$O_{i\cdot}$
\vdots	\vdots			\vdots		\vdots	\vdots
A_r	O_{r1}	O_{r2}	\dots	O_{rj}	\dots	O_{rc}	$O_{r\cdot}$
行總和	$O_{\cdot 1}$	$O_{\cdot 2}$	\dots	$O_{\cdot j}$	\dots	$O_{\cdot c}$	n

卡方獨立性檢定的統計值公式為

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}},$$

其中 E_{ij} 為期望次數, 計算方式是以第 i 列總和與第 j 行總和之乘積除以總樣本數 n 而得:

$$E_{ij} = \frac{O_{i\cdot} \times O_{\cdot j}}{n}.$$

現有一資料如下, 想了解性別與政黨傾向是否獨立。

Gender	Party Identification		
	Democrat	Independent	Republican
Females	762	327	468
Males	484	239	477

- (a) 將上述資料表格，輸入 R 成為一矩陣 (matrix) 類別的物件後，將之轉成表格 table 類別物件 (命名為 `GenderParty.observed.ct`)，並加上維度名稱後印出如下。(提示: `as.table`, `dimnames`)

```
> GenderParty.observed.ct
      party
gender Democrat Independent Republican
F         762           327           468
M         484           239           477
```

- (b) 計算並印出列總和、行列總和及樣本總個數。
- (c) 計算並印出期望次數表格 (命名為 `GenderParty.expected.ct`) (注意，此物件之 R 類別同 `GenderParty.observed.ct`)。
- (d) 計算並印出此資料的卡方獨立性檢定統計值 χ^2 。