(本頁中文版如下一頁)

National Chengchi University, 113-1 Academic Year
Midterm Exam of Statistics (I), Bonus Test, R Programming

Department/Grade:_____ ID:_____ Name:_____

Subject: Statistics (I)

Date: 2024/11/14

This test consists of 5 major questions. (20% each, total score: 100%)

Time period: 15:00~16:00 (total 60 minutes)

Notes:

1. Download the R exam sheet (`113-1-Stat-R-Midterm.zip`) from the course website and unzip in your laptop. The zip file contain the question sheet, the answer sheet, and the datasets.
2. Answers for this exam should be provided using the R programming language (either Rgui or RStudio). Other programming languages are not permitted.
3. During the exam, you may refer to textbooks, lecture notes (including videos, Please bring your own headphones), or browse the internet. However, the use of communication software/APP such as Messenger, IG, Line, etc., is strictly prohibited.
4. Any form of cheating or suspicious behavior is not allowed.
5. On this answer sheet, please ensure you copy the "executed code and its results (including graphics)" from the **R Console** and paste it here (in `Courier New font`, size 10, black text on a white background). This should include both the code and the output, not just one or the other. Finally, the answers for each sub-question should be highlight by yellow color (not just printing the report; the TA shouldn't have to search for the answers)
6. Please label your answers in sequence, e.g., (1)a, (1)b, (2)a, etc.
7. After completing your answers, save this Word document with the filename "`StudentID-FamilyName-Midterm.docx`" (replace with your actual "`Student ID` and `FamilyName`") and upload it to http://hmwu.nccu.edu.tw/login.html .
8. Username: `stat113`, Password: (classroom number) `26xxxx`, Folder: "`20241114-MidtermExam`".
9. If the upload site displays a "blank page", move your cursor to the "address bar" and press "Enter". If that doesn't work, try using a different browser (IE/Edge/Firefox/Chrome).
10. Uploaded files cannot be deleted. If you need to upload a revised file, please add "-2" to the main filename, e.g., " `StudentID-FamilyName-Midterm-2.docx` ".

_____Wishing you a successful exam_____

(English version on the previous page)

國立政治大學 113 學年度第一學期

統計學(二) 期中 R 程式加分考

系級:_____ 學號:_____ 姓名:_____

考試科目: 統計學(一)

考試日期:2024/11/14

本試題共 5 大題 (各 20%)

考試時間:15:00~16:00 (共 60 分鐘)

注意事項:

1. 從教學網站下載電子考卷 (`113-1-Stat-R-Midterm.zip`)，並於自己的筆電解壓縮。壓縮檔包含題目卷、答案卷和資料集。
2. 本次考題以 R 程式(Rgui 或 RStudio)方式作答，其他程式不允許。
3. 考試過程中可查詢書本、教學講義或上網(含上課影片，請自備耳機)，禁止利用 messenger, IG, Line 等等通訊軟體。
4. 禁止疑似作弊行為。
5. 本答案卷上請務必於 R Console 內**複制「執行後的程式碼及結果(含圖形)」**，於本答案卷貼上(`Courier New, 10 點字，白底黑字`)，不是只有程式碼，不是只有報表。最後，將每小題之答案以黃色底高亮起來(不能只印出報表，要助教去找答案)。
6. 請依序註明題號: (1)a, (1)b, (2)a 等等。
7. 作答完請將此 word 檔存檔，檔名為「`StudentID-FamilyName-Midterm.docx`」(更改成自己「學號」、「姓」)並上傳至 http://hmwu.nccu.edu.tw/login.html
8. 帳號: `stat113`，密碼: (上課教室號碼) `26xxxx`，資料夾:「`20231111-MidtermExam`」
9. 如果上傳網站出現「空白頁」，請將滑鼠移至「網址列」後，按「Enter」即可。若再不行，請換其它瀏覽器(IE/Edge/Firefox/Chrome)
10. 上傳檔案無法刪除，若要上傳更新檔，請於主檔名後加「-2」，例如:「`StudentID-FamilyName-Midterm-2.docx`」。

_____祝考試順利_____

How to Upload the Answer Sheet?

Step 1: Check the file name is correct: StudentID-FamilyName-Midterm.docx.
(Example: **`110101069-Tsai-Midterm.docx`**)

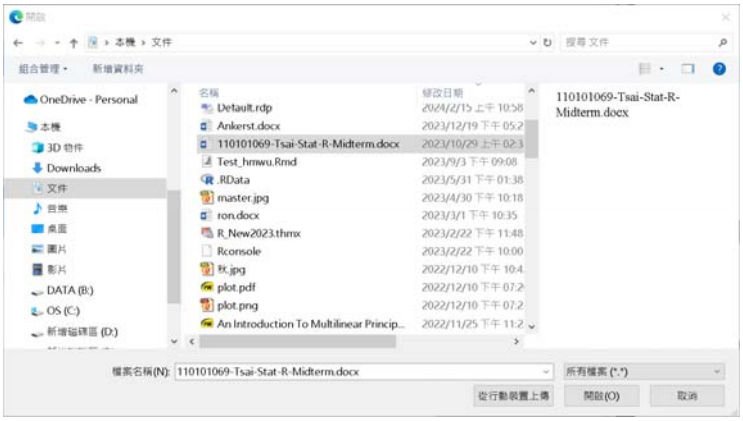
Step 2: Go to the upload site: http://hmwu.nccu.edu.tw/login.html

Step 3: Type in the login name: stat113 and password: 26xxxx

Step 4: Click to enter the correct folder before uploading.

Step 5: Click the upload file button.

Step 6: Select your answer sheet file.



Step 7: Close the select file window.

Step 8: Ensure the uploaded file is correct and then log out.

How to Write the Answer Sheet?

(1)

Data file: NCAA

NCAA Basketball Game Scores. A sample of 10 NCAA college basketball game scores provided the following data.

Winning Team	Points	Losing Team	Points	Winning Margin
Arizona	90	Oregon	66	24
Duke	85	Georgetown	66	19
Florida State	75	Wake Forest	70	5
Kansas	78	Colorado	57	21
Kentucky	71	Notre Dame	63	8
Louisville	65	Tennessee	62	3
Oklahoma State	72	Texas	66	6
Purdue	76	Michigan State	70	6
Stanford	77	Southern Cal	67	10
Wisconsin	76	Illinois	56	20

a. Compute the mean and standard deviation for the points scored by the winning teams.

b. Assume that the points scored by the winning teams for all NCAA games follow a bell-shaped distribution. Using the mean and standard deviation found in part (a), estimate the percentage of all NCAA games in which the winning team scores 84 or more points. Estimate the percentage of NCAA games in which the winning team scores more than 90 points.

c. Compute the mean and standard deviation for the winning margin. Do the data contain outliers? Explain.

```
> # 1(a)
> library(readxl)
> NCAA <- read_excel("data/NCAA.xlsx")
> head(NCAA)
# A tibble: 6 × 5
  WinningTeam   PointsW LosingTeam   PointsL WinningMargin
  <chr>           <dbl> <chr>          <dbl>         <dbl>
1 Arizona            90 Oregon            66            24
2 Duke               85 Georgetown        66            19
3 Florida State      75 Wake Forrest      70             5
4 Kansas             78 Colorado          57            21
5 Kentucky           71 Notre Dame        63             8
6 Louisville         65 Tennessee         62             3
> mean(NCAA$PointsW)
[1] 76.5
> sd(NCAA$PointsW)
[1] 7.011895
>
> # 1(b)
> xbar <- mean(NCAA$PointsW)
> s <- sd(NCAA$PointsW)
> z <- (84 - xbar) / s
> z # about one-sd
[1] 1.069611
> cat("The percentage of all NCAA games in which the winning team scores 84 or more
points is: ", (100 - 68)/2, "%.")
The percentage of all NCAA games in which the winning team scores 84 or more points
is:  16 %.
> z <- (90 - xbar) / s
> z # about two-sd
[1] 1.9253
> cat("The percentage of NCAA games in which the winning team
+ scores more than 90 points: ", (100 - 95)/2, "%.")
The percentage of NCAA games in which the winning team
scores more than 90 points:  2.5 %.
>
> # 1(c)
> xbar_wm <- mean(NCAA$WinningMargin )
> xbar_wm
[1] 12.2
> s_wm <- sd(NCAA$WinningMargin)
> s_wm
[1] 7.885289
>
> Detect_Outlier <- function(x){
+   Q1 <- quantile(x, type = 6, probs = 0.25)
+   Q3 <- quantile(x, type = 6, probs = 0.75)
+   IQR.x <- IQR(x, type = 6)
+   Lower.Limit <- Q1 - 1.5 * IQR.x
+   Upper.Limit <- Q3 + 1.5 * IQR.x
+   ifelse(x < Lower.Limit | x > Upper.Limit, "Outlier", "Not")
+ }
> Detect_Outlier(NCAA$WinningMargin)
 [1] "Not" "Not" "Not" "Not" "Not" "Not" "Not" "Not" "Not" "Not"
> cat("The data did not contain outliers.")
The data did not contain outliers.
```