

114-1-Programming and Statistical Software Exam Paper

Date: 2025/10/09 (Thu) 10:30~11:50

Page 1 of 4

Notes

1. **Time** (Please sit according to your usual seat):
 - (a) During the exam period (10:30~11:50), the internet will be disconnected. You may only use the public computer and your USB drive for answering questions. Laptops, tablets, and mobile phones are not allowed (even with network functions disabled).
 - (b) Exam paper download: Download the exam paper from the course website within 5 minutes before the exam (10:25~10:30) onto the public computer.
 - i. Students who arrive late or need to (re)download the exam after disconnection must bring their USB drive to the instructor's laptop to download (a small point deduction will apply).
 - ii. The download file name is: 1141-StatSoft-exam1.zip. Don't tell the instructor you don't know how to unzip it or where the file was downloaded!
 - iii. The zip file contains the question paper, answer sheet, dataset, and some R packages (zip files). Don't tell the instructor you don't know how to install R packages locally!
 - (c) Exam submission: Upload the answer sheet within 5 minutes after the exam (11:50~11:55). [Students submitting early must bring their USB drive to the instructor's laptop to upload.]
2. The exam must be answered using R (Rgui or RStudio). Other programs are not permitted.
3. You may refer to any materials during the exam, including textbooks, lecture notes, and electronic files (documents or videos).
4. The use of communication software such as FB Messenger, IG, Line, etc., for discussion is strictly prohibited. The use of ChatGPT or similar AI-assisted tools is

also prohibited. Any cheating or suspected cheating behavior is forbidden. [Please don't be the one who ruins it for everyone!]

5. Answer Sheet:

- (a) Please use the provided answer sheet template: MS WORD file named "Name-StatSoft-exam1.docx" (replace "Name" with your own).
- (b) In the answer sheet, be sure to copy both your executed R code and results (including figures) from the R Console, and paste them into the answer sheet using **Courier New**, 10pt font, black text on white background. Not just code, not just output tables. (The instructor has demonstrated this in class—see the recorded lecture.)
- (c) Indicate question numbers in order, e.g., # ex1(a), # ex1(b), # ex2, etc.
- (d) Even if your answer is incomplete or the code produces an error, you must still paste the executed code and results (including figures). Blank answers cannot receive partial credit.

6. Uploading the Answer Sheet:

- (a) Log in to the [Assignment Upload Area] on the course website or visit <http://hmwu.nccu.edu.tw/login.html?lang=tchinese>.
- (b) Account: statsoft, Password: xxxx (announced on the course FB group), Folder: "20251009-exam1" .
- (c) If a "blank page" appears when uploading, move your mouse to the URL bar and press "Enter." If that doesn't work, try another browser (IE/Edge/Firefox/Chrome).
- (d) After uploading, verify the file size. Once uploaded, submission is complete. Files cannot be deleted or re-uploaded. If you encounter problems, contact the instructor.
- (e) Do not tell the instructor that your answer file disappeared or became blank—the instructor cannot fix it.
- (f) Students who finish early during the disconnection period should bring their USB drive to the instructor's laptop for upload. Once uploaded, you may leave.
- (g) One day before the exam, you may practice uploading any file following the same steps into the folder: "upload_testing" .

7. If your public computer encounters an issue that cannot be resolved, please move to the "Free Zone" (the back two rows without fixed seats).
8. If you have any problems, try to resolve them yourself first, or raise your hand to ask the instructor. Do not communicate with other students.
9. If the public computer cannot read your USB drive, informing the instructor won't help either (test this during class beforehand!).
10. During the exam, there is only one instructor and one TA. Please stay calm and wait patiently if you encounter issues.
11. Please use the restroom before the exam. Leaving the classroom during the exam is not allowed except under special circumstances.
12. Please read all instructions carefully to ensure a smooth and successful exam.

R Problems (Total: 120 points)

1. (10 points) Use `readline` to display prompts and read your department, student ID, and name, then print the following declaration using the entered data:

"I, (department, student ID, name), solemnly abide by all exam regulations. If I violate them, I am willing to accept the university's most severe punishment. I hereby declare."

2. (15 points) A student analyzed the relationship between wind speed (`Wind`) and temperature (`Temp`) in the `airquality` dataset using regression and covariance analysis, as follows:

```
lm.obj <- lm(airquality$Wind ~ airquality$Temp)
lm.anova <- anova(lm.obj)
lm.summary <- summary(lm.obj)
```

- (a) (5 points) Print only the parameter estimation table as shown:

| Estimate | Std. Error | t value | Pr(> t) |
|------------------|------------|------------|------------------------|
| (Intercept) | 23.2336881 | 2.11239468 | 10.998744 4.901351e-21 |
| airquality\$Temp | -0.1704644 | 0.02692606 | -6.330835 2.641597e-09 |

- (b) (5 points) Print only the value `"2.641597e-09"` or `"2.642e-09,"` which is the p-value for the regression coefficient of `airquality$Temp`.
 - (c) (5 points) Print only the Adjusted R-squared value: 0.2045. (You must extract Adjusted R-squared from the report. Knowing what it means is not required.)
3. (30 points) The file `"Husehold_funds_data.txt"` records household financial data collected by an interviewer, including columns `"Survey Date (YYYY/MM/DD), Time (HH:MM), City, Household Funds (10,000 NT), and Rating."`
 - (a) (5 points) Read in the data and print the first five records and their structure.
 - (b) (5 points) Convert the rating variable into an ordered factor with levels `A<B<C` and print it.
 - (c) (10 points) Print all household data with rating B or higher and calculate the mean household funds (10,000 NT). (Use relational operators `<`, `<=`, `>`, `>=`.)

- (d) (10 points) Print all household data surveyed in 2022 and calculate the mean household funds (10,000 NT). (Use relational operators $<$, $<=$, $>$, $>=$.)
4. (25 points) There is a score data file "106-1-DA-Score2.xlsx."
- (a) (5 points) Read in the data and print the first and last three students' records.
- (b) (10 points) Based on the weights for each exam (attendance not counted; absences counted as zero), calculate the total score for each student. Then create a data frame with columns "Student ID, Name, Total Score" and print it.
- (c) (10 points) Print the student ID, name, and total score of all male students whose total score is between 50 (inclusive) and 70 (exclusive).
5. (20 points) The probability density function of the normal distribution is given by:

$$f(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

Compute the probability density values **fx** for **x** values $(-3, -2, -1, 0, 1, 2, 3)$, mean $\mu = 3$, and standard deviation $\sigma = 2$. Also, use **R**'s built-in **dnorm** function to compute the same values, and print the table below. (**fx** should not be written as an R function.)

| | x | fx | dnorm |
|---|----------|-------------|--------------|
| 1 | -3 | 0.004431848 | 0.004431848 |
| 2 | -2 | 0.053990967 | 0.053990967 |
| 3 | -1 | 0.241970725 | 0.241970725 |
| 4 | 0 | 0.398942280 | 0.398942280 |
| 5 | 1 | 0.241970725 | 0.241970725 |
| 6 | 2 | 0.053990967 | 0.053990967 |
| 7 | 3 | 0.004431848 | 0.004431848 |

6. (20 points, 5 points each) Given the series S_n :

$$S_n = \sum_{i=1}^n \frac{(-1)^{i+1}}{2i-1} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots + \frac{(-1)^{n+1}}{2n-1}.$$

It is known that $\lim_{n \rightarrow \infty} S_n = \pi/4$.

- (a) Generate $(-1)^{i+1}, i = 1, \dots, n$ (where $n = 10$), and print: 1 -1 1 -1 1 \dots -1.
- (b) Generate $2i - 1, i = 1, \dots, n$ (where $n = 10$), and print: 1 3 5 7 \dots 19.
- (c) Generate $(-1)^{i+1}/(2i - 1), i = 1, \dots, n$ (where $n = 10$), and print: 1 -1/3 1/5 -1/7 \dots 1/19.
- (d) Compute $4S_{10}$, $4S_{100}$, and $4S_{1000}$.