

Statistics I

Homework 2

Due time: 2021/10/25 Monday 10:00 a.m Please submit the paperwork and **write down your student ID and your name in the upper right corner**. Make sure your handwriting is readable and highlight your answer part would be nice for scoring. You can hand in your homework at the class or hand it to the TA directly at 森林館403 any time. If you have any questions, send TA the email (d08625001@ntu.edu.tw).

1. (2 points) Random walk [1] is a mathematical object, known as a stochastic or random process, it is easy to be understood by imaging that a drunk person walking. Random walk is used for simulating the price change in the finance application. Draw a 2-dimension **random walk path** plot using H2RandomWalk.csv data.
2. (5 points) Draw a diagram, a box plot and a pie plot that describe the frequency of the data in H2Q2.csv. (You may classify the data using the interval with range 5 if needed.)
3. (3 points) There is a 4-dimension dataset (y, x_1, x_2, x_3) where x_1 is the categorical data (e.g species, sex...) and y, x_2, x_3 are continuous data (e.g temperature, age...). How would you visualize the data?

Only accept paper, thus the graphs should be printed. You may need assistance from some software. R is recommended in this course. However, if you master other computer languages, it's also okay for the homework, but I may not be able to answer the programming problem. Your TA is good at golang, Python and VB, if your problem belongs to the above fields, it is welcome to discuss with me.

References

- [1] Karl Pearson. The problem of the random walk. *Nature*, pages 294–294, 1905.