2025FALL Probability & Statistics for EECS 2025/12/08

Final Project Option 3

Due:2026/01/20 10:59pm

This is the Final Project Option 3 entitled “Reverse Engineering the Mechanism of WeChat Red Envelope”. This project can be done by a team with no more than three students. Your team is required to use Python for the programming part. Your team needs to submit all things in Jupyter Notebook format, including Python codes, simulation results, analysis, discussions, tables, figures, *etc.* Always keep the **academic integrity** in mind and remember to give credit to any source that inspires you.

- **Goal:** In this project, let's subjectively feel and objectively **reverse-engineer** the random allocation mechanism of WeChat Red Envelope, short as WRE mechanism. It is an open exploration based on the sampled data from actual WeChat application. There is no “standard/correct” answer and it is all about **logically** combine observation, modeling, reproduce and justification to clearly draw a **reasonable** conclusion.

1. **Observation.** You should collect sufficient & representative data using the WeChat application to have a direct observation of how the money in Red Envelope is distributed. Remember to clearly mention your experiment settings. For example,
   - How many scales have you tried while conducting the experiment? 5 RMB in the a group of 3 users or 50 RMB in a group of 10 users.
   - How many trials have you repeated to get the results?
2. **Visualization.** You are required to use Histogram, Boxplots and Scatterplot to plot/demonstrate your data such that you can have intuition/insights from it.
3. **Distribution Modeling.** Guessing the parametric form of hidden distribution from your data. Use Maximum Likelihood Estimation method to obtain the parameters. For example, if you think the hidden distribution is Normal with two parameters μ and σ^2. You need to use the sample data to estimate the value of μ and σ^2. The hidden distribution can be joint distributions, for example, the multinomial distribution or multi-variate normal distribution.
4. **Model Testing.** To check whether your data fits your distribution well, you are required to do the goodness of fit test with p-value. The Tests are Chi-square test and Kolmogorov-Smirnov test. If the guessed distribution is normal, you are required to do one more test: Shapiro-Wilk test. Discuss the pros and cons of your tests.

5. **Generative Modeling (Optional)** Based on the data, use the latest diffusion modeling to train a deep probability model. Such model is a proxy of the hidden distribution. Use such model to generate new samples. Then do the hypothesis test for check whether such new samples and the old samples (Collected from Wechat) are from the same distribution.

6. **From Modeling to Decision-Making Policy:** After several rounds of revision (i.e., loops of steps 1 to 4), you should be happy about your current model which can better mirror the behavior of WeChat than the initial model. Now, you should think further about further decision-making question based on your previous analysis:

   When should you start to snatch/grab Red Envelopes? At the very beginning of their appearance or waiting for a right moment. The decision-making policy should depend on the number of people in the Wechat group, and the number of people who have already obtained the red envelopes.

7. **Further Explorations:**

   (a) Does the WRE mechanism behave the same inside and outside China?

   (b) Does the WRE mechanism behave the same for a work day and a holiday? Maybe you want to have some tests on the coming Christmas.

   (c) QQ also has a builtin Red Envelope function. Are they exactly the same?

   (d) Can you propose a "user-specific" WRE mechanism such that it is optimized for groups with known backgrounds, e.g., a group of children, a company working group with the boss and employees, and a family group with close or distant relatives? For example, in a group of young children, you may want their received money to be similar in case some unlucky one would get upset upon the new year.

   (e) Can you propose an "fairness-ware" WRE mechanism such that it is optimized for successive Red Envelopes? Consider a group of five users, each of them sends a Red Envelope in their group to celebrate the new year. However, upon the revealing of the third Red Envelope, there is one unlucky user who has got significantly less money than others. How can you dynamically just the WRE mechanism to compensate unlucky users in a group for fairness?