Stats210 Final project : An analysis of WeChat red envelop

Jiayang Hong, Xi Chen, Tianxuan Sun, Tianji Sun

May 2022

1 Abstract

WeChat Red Envelope is a functionality developed by Tencent in 2014 to add users for WeChat Pay. This function soon became one of the most popular online recreational interactions during the holidays. Later version of it introduced fancy covers and built-in-ads added commercial values (Glossy 2021). In this project, our group experimented to reverse engineer the mechanism of Wechat red envelope distribution. We sent out 105 digital red envelopes containing the same amount of money, then we collected data according to the grabbing order to build our distribution program. Using this program we did some simulation and came to the following conclusion. WeChat distributes the Red Envelope when the open is clicked, and the amount given is uniformly distributed from 0.01 to two times the average of the remaining amount divided by the remaining portions. Later our simulations proved this hypothesis.

2 Introduction

Our group consisted of four members. To form a group with more participants, we worked together with another group of three members. The goal of this experiment is to investigate the distribution mechanism. As the main purpose of this experiment is to evaluate how the grabbing order affects the expected value of money, we designed our experiment accordingly. After analyzing the data we collected from the experiment and the simulation, the hypothesis about the red envelope mechanism is testified. We also derived some interesting findings on the techniques of grabbing red envelopes.The variance of expected money increase with order, so be the early one to grab indicates stability and late ones means higher risk but possibly gain more.

3 Procedure

In this experiment, we first formed a WeChat group to distribute the Red envelopes. We worked with another group. We have seven participants in the experiment, each of us sent out 15 5-Yuan red envelopes, then we recorded the amount according to the grabbing order. Data was recorded for initial observation. After that, we formed a hypothesis of WeChat‘s distribution mechanism. One of our group members used python to design our Red Envelope program according to the hypothesis.

We create a class called red in Python. In this class, we set the parameters of object to "people" and "money". There are also two parameters in this class: "remainingpeople" and "remainingmoney", and make the initial values of these two parameters equal to "people" and "money" respectively.

Based on what we have found before, we let the money on each poll obey the uniform distribution with the parameter 0.01 and two times the remaining mean. For the last person who take the red envelop, he takes away all the money remaining.

The remaining mean is calculated by the “remainingmoney” divided “remainingpeople”by in this object.

In process part, we first generate a random number R between 0 and 1 with uniform probability. And assign the twice of the remaining mean to the variable MAX. So MAX time R is the maximum value of each poll and minimum is 0.01. Also, we need to round since this number will have a lot of digits and we only need two digits in this experiment. After this process, “remainingpeople” will decrease one and “remainingmoney” will decrease the amount of money take away in this process. Then repeat the procedure for several times until “remainingpeople” reached 0. We used the program to simulate the Red Envelope. The first time we repeated the progress of our experiment, then we simulated 10000 red envelopes to get a general conclusion. After the simulation, we calculated the mean and variance of each grab for further analysis.

4 Data

Based on the data we collected from the red envelope experiment, we examined and analyzed the possible hypothesis. Our main purpose in this research is to test how the order of grabbing the red envelope affects the expected value of money. Figure 1 shows bar chart of the mean money of the experiment data for different orders. The horizontal axis is the order of grabbing, and the vertical axis is the mean money value. We can see that the second money grabber has the highest expected value of money of about 0.802 yuan, and the fifth money grabber has the lowest of about 0.644 yuan. Although the mean value varies only from 0.6 yuan to 0.9 yuan with no evident pattern. The difference shows that the sample size of 105 is still relatively small.

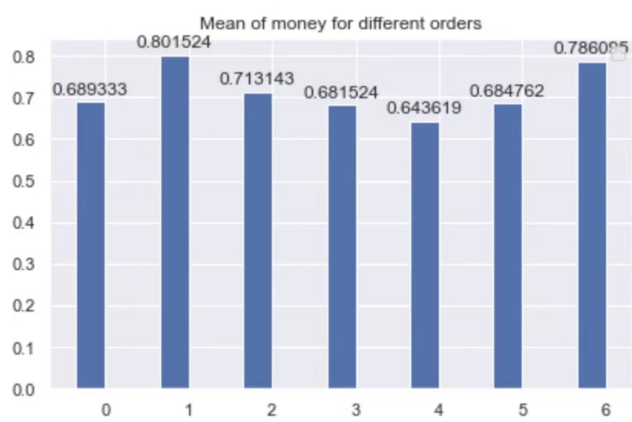

Figure 1: Mean of money for different orders

Figure 2 shows bar chart of the variance of money value for different orders based on the experiment results. The magnitude of money value is relatively large as the sample size is small. The variation of the variance at different orders still does not show a trend. Interestingly, the pattern of the variance corresponds to the pattern of mean value in this experiment. The second and last money grabber still have the larger value. But as magnitude of mean value does not affects the magnitude of variance of data, this similarity is merely a coincidence and therefore does not possess practical significance.

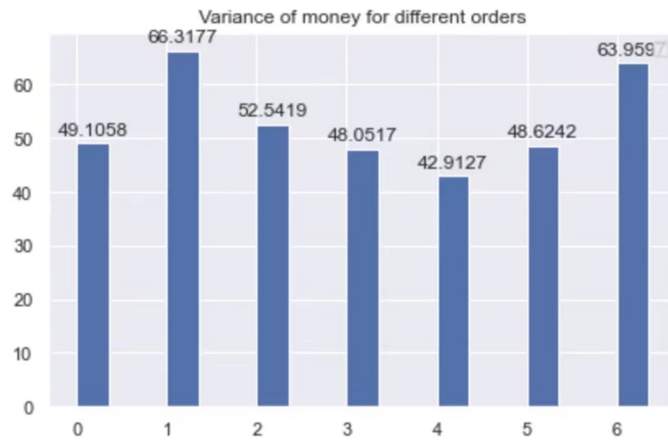

Figure 2: Variance of money for different orders

After deriving the bar chart of the mean and variance of the experiment data. We are interested in the distribution of money value of different order. When generating the distribution plots using python, we considered the bin that we set for the plot. As we have seven participants in this experiment and red envelope with money value of 5 yuan each. The initial choice was the bin of 0.5 yuan, which is what Figure 3 shows. We find that the 0.5-yuan bin was too large for red envelope of 5 yuan each, making the distribution pattern incapable of showing its distribution. Figure 4 is the distribution pattern using the 0.2-yuan bin, comparing to the 0.5-yuan bin, we can see that the distribution pattern is more evident.

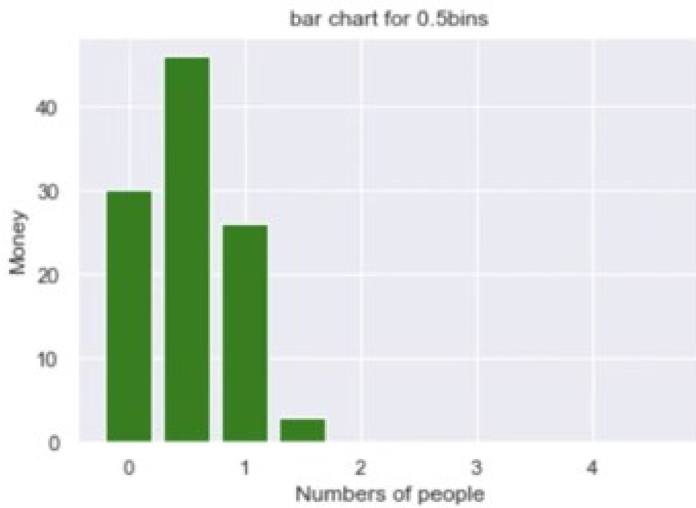

Figure 3: Distribution using 0.5 bin

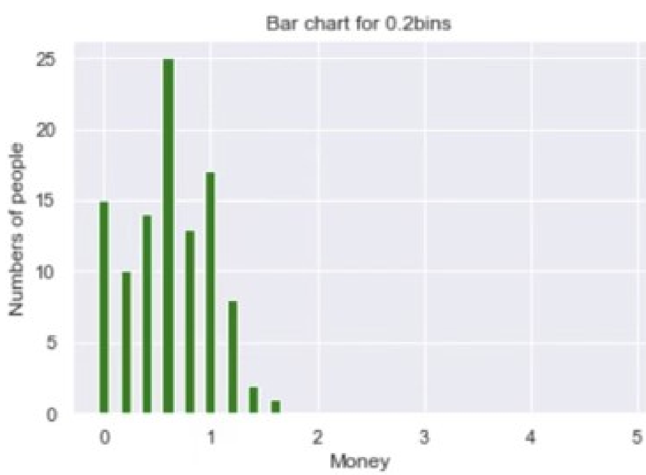

Figure 4: Distribution using 0.2 bin

Figures 5 shows the distribution of money at order 1 to 7 by the experiment data. The horizontal axis is the money value with the bin of 0.2 and the vertical axis is the frequency of people getting the money. Although the distributions are still irregular due to the limitation of sample size, we can see that the distribution tends to be right skewed as the increase of order number and the variation also increases.

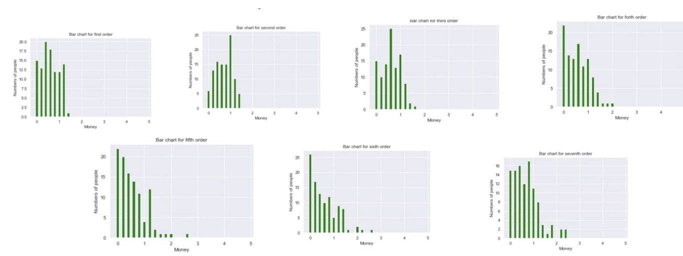

Figure 5: Distribution of money for each order

Based on the mechanism we designed on python, we are able to simulate the process by the program and compared the result to the data we derived from experiment. Initially, we simulate the red envelope grabbing using 105 trials, which is the same as the experiment. The purpose is to test whether the mechanism we write correspond to our experiment results. Figure 6 and 7 below are the distribution of the first money grabber of the experiment and the simulation. We also compared other data like the mean and variance of the experiment and simulation, which shows similar results. But as there is only 105 trials, the data requires to be test further.

Figure 6: Distribution of the experiment

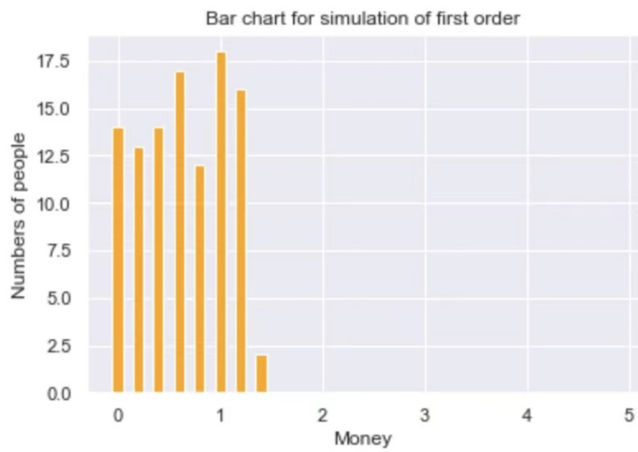

Figure 7: Distribution of the simulation

To simulate the process with a large sample size, we first choose to conduct the ten thousand trial simulation to test the experiment. Figure 8 is the bar chart that shows the mean distribution of the money at the seven different order. The plot generally shows that the expectation of money value at different orders are the same. The mean at the second order interestingly shows a higher value, which correspond to our finding of the experiment, but there still exists considerable difference. We decided to take 100000 trials in simulation, making the result more accurate. Figure 9 shows the mean value by 100000 trials. The mean value varies only from the minimum of 0.712287 to the maximum of 0.715236, which we consider sufficient to accept the hypothesis that the expectation is the same for all orders of grabbing, the expected value is about 0.714, calculated from dividing 5 yuan by 7 participants.

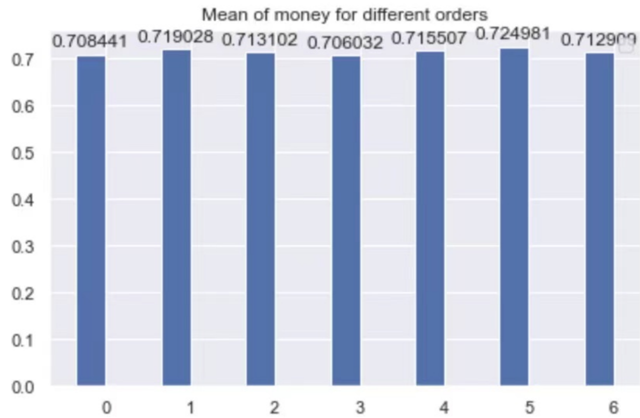

Figure 8: Mean of money for different orders in 10000-trial simulation

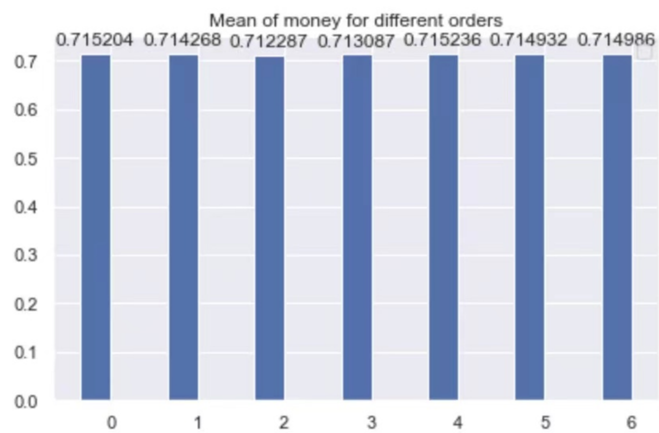

Figure 9: Mean of money for different orders in 100000-trial simulation

As the 100000-trial simulation satisfies our requirement, we plotted the variance of each order of the simulated result. From Figure 10, we can see that the variance increases with the order of grabbing. Comparing to the previous bar char that shows the variance of the experiment, the variance of simulation is much smaller in magnitude, which is reasonable as the sample size increases about 1000 times. Another significant finding is that the last two order has almost the same variance. The reason has been proved mathematically in the theory part of the report. Intuitively, the mechanism of the red envelope is to divide the remaining money randomly to two pieces with the minimum of 0.01 yuan. So the value of the last two grab is determined by one step, making the variance of the two value equal. Also, we noticed that the variance difference of the adjacent order increases significantly from the fifth grab to the sixth grab, this could also be explained by the mechanism. As during the last two grabbing, the WeChat stopped limiting the value to be no more than twice of the average of the remaining money, the variance would certainly increase without the limitation.

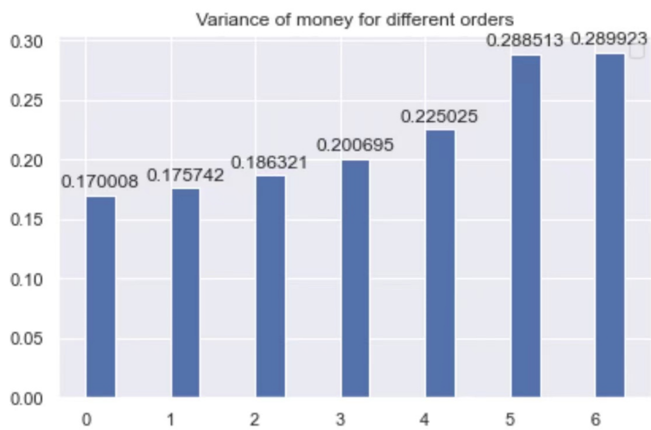

Figure 10: Variance of money for different orders in 100000-trial simulation

Figure 11 shows the distribution of each order using the result from the 100000-trial simulation. We can see that the distribution of the first order is close to a uniform distribution. As the increase of order, the distribution became more right skewed shape. The comparison also shows the increase of variance with the order. The last two figure shows that the variance are the same during the last two grabbing.

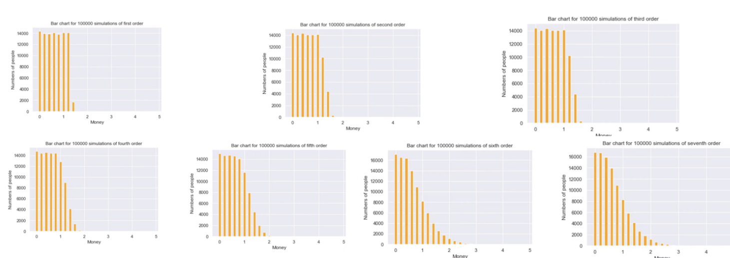

Figure 11: Distribution of different orders in 100000-trial simulation

We thought it will be interesting to analyze this mechanism commercially. For WeChat, the Red Envelope needs to be fun and engaging. This will mean they have to limit the upper number of each draw. This mechanism is similar to the mechanism of box opening in the gaming industry. (Rune Paweł, 2019)To increase user experience, the random item needs to avoid extreme cases such as constantly giving out little or large amounts to one person. The upper limit is only a simple trick in this domain.

5 Theorem

We attempt to prove $E[x_1] \approx E[x_2] \approx E[x_3] \approx \cdots E[x_n]$.
We know that $X_1 \sim U\left(0.01, \frac{2S}{n}\right)$,$E[x_1] = \frac{0.01+\frac{2S}{n}}{2} \approx \frac{S}{n}$
We also know that $X_k \sim U\left(0.01, 2\frac{s-\sum_{i=1}^{k-1} x_i}{n-k-1}\right)$,which could be approximately seen as $X_k \sim U\left(0, 2\frac{s-\sum_{i=1}^{k-1} x_i}{n-k-1}\right)$
Then $f(x_n \mid x_1, x_2 \ldots x_n) = \frac{n-k+1}{2\left(s-\sum_{i=1}^{k-1} x_i\right)}$,$x_n \in \left[0, 2\frac{S-\sum_{i=1}^{k-1} x_i}{n-k+1}\right]$
$E[X_n \mid x_1, x_2, \ldots x_{n-1}] = \int_0^{2\frac{s-\sum_{i=1}^{k-1} x_i}{n-k+1}} \frac{(n-k+1)\cdot x_n}{2\left(s-\sum_{i=1}^{k-1} x_i\right)} dx_n = \frac{s-\sum_{i=1}^{k-1} x_i}{n-k+1}$
Now use mathematical induction to prove:
When n=1, $E[x_1] \approx S/n$
We assume that when n=k,$E[x_k] \approx S/n$
When n=k+1,$E[x_{k+1}] = E[E[x_{k+1} \mid x_1, x_2, \ldots x_k]] = E[\frac{s-\sum_{i=1}^{k-1} x_i}{n-k}] = \frac{S}{n-k} - \sum_{i=1}^{k} E[X_i] = \frac{s}{n-k} - (k/n-k) \cdot \frac{s}{n} = S/n$
Therefore,$E[x_1] \approx E[x_2] \approx E[x_3] \approx \cdots E[x_n] \approx \frac{s}{n}$

Besides, we attempt to prove that $Var[X_1] < Var[X_2] < \cdots Var[X_{n-1}] \approx Var[X_n]$
Because we already know that $E[x_1] \approx E[x_2] \approx E[x_3] \approx \cdots E[x_n]$,then we only need to compare $E[x_n^2]$
Let S_m denote $\sum_{i=1}^{m} x_i$
According to previous proof, we could obtain $f_{s_m|_{m-1}(p|q) = \frac{n-m+1}{2(s-q)}}$,where $p \in \left[q, \frac{(n-m-1)q+2s}{n-k+1}\right]$
Therefore, $E\left[s_m^2\right] = \int_q \frac{f_{m-1}(q)\cdot(n-m+1)}{2(s-q)} \left(\int_p p^2 dp\right) dq$
We finally obtain $E\left[s_m^2\right] = E\left[s_{m-1}^2\right] \cdot \left(3(n-m+1) - 6 + \frac{4}{n-m+1}\right) + \frac{-2m^2+2mn+2m-2}{n^2-mn+n}$
Also, we could similarly get $E\left[x_m^2\right] = \int_q f_{s_{m-1}}(q) \cdot \frac{n-m+1}{2(s-m)} \left(\int_p p^2 dp\right) dq = \frac{4E\left[s_{m-1}^2\right]+4\left(1-\frac{2m}{n}+\frac{2}{n}\right)}{3(n-m+1)^2}$
According to these two equations, we could get $E\left[X_{m+1}^2\right] = \left(\frac{1}{3(n-m)^2} + 1\right) E\left[X_m^2\right]$,which can be seen as a proportional sequence regarding $E[x_m]$, where $E[x_1] = 4s^2/3n^2$
Therefore,$E\left[X_m^2\right] = \frac{4s^2}{3n^2} \prod_{i=1}^{m-1} \left(\frac{1}{3(i-m+1)^2} + 1\right)$ which could be easily observed that $E\left[X_1^2\right] < E\left[X_2^2\right] < \cdots < E\left[X_{m-1}^2\right] = E\left[X_m^2\right]$.
Therefore $Var[X_1] < Var[X_2] < \cdots Var[X_{m-1}] = Var[X_m]$

6 Conclusion

In this project, we first carried out the experiment of WeChat red envelope. In the experiment, a total of 105 red envelopes were uniformly distributed among 7 people, which means 15 red envelopes were distributed by each person. Each red envelope contains five yuan. We recorded the experimental data and found that: The money of red envelopes is uniformly distributed between 0.01 and two times the mean value of the remaining red envelopes. Based on this finding, we conducted simulation experiments using python.

In the python environment, we conducted 100000 experiments, and use NumPy to calculate and plot the relationship between variance, mean and order. We found: 1. The mean value of the amount of money does not change significantly with the order. 2. However, the variance changes significantly with the increasing order. The variance increases with the rank and reaches the peak in the sixth and seventh places (i.e. the last two places)

The above findings show that if you want to pursue stable results, you'd better grab the red envelope as soon as possible. If you want to win more money, you can grab the red envelope later, but you also need to take greater risks

7 Appendix

Contributions:
Tianxuan Sun: Data Analysis, Powerpoint
Xi Chen: Simulation using Python
Tianji Sun: Presentation
Jiayang Hong: Data visulization and mathmatical proof

8 Reference

Nielsen, Rune Grabarczyk, Paweł. (2019). Are Loot Boxes Gambling? Random Reward Mechanisms in Video Games. Transactions of the Digital Games Research Association. 4. 10.26503/todigra.v4i3.104.
Glossy (2021)WeChat's red envelope cover is fashion brands' shiny new ad space. https://www.glossy.co/fashion/wechats-red-envelope-cover-is-fashion-brands-shiny-new-ad-space/