# Reverse Engineering Algorithmic Mechanism Behind WeChat Red Envelope

Jiawei Liang
student ID:98277894
School of Information Science and Technology
ShanghaiTech University

***Abstract*—This study investigates the engineering algorithmic mechanism behind WeChat red envelope from two parts: the relationship between order and amount, the money composition of red envelope. The findings show that the order of snatching red envelope seems to fit the Uniform distribution and the composition of red envelope fits the normal distrubution when there is a huge amount of data.**

***Keywords*—WeChat red envelope, order and amount, Uniform distribution, Normal distribution**

## I. Introduction

Recent years, WeChat red envelope becomes more and more popular among relatives and colleagues especially during the festivals. In order to increase the enjoyment and playability, what is the algorithm used by Tencent team behind WeChat red envelope? I collected some data to find it.

The main results and contributions of this report are summarized as follows:

- **Red envelope data collection and processing.**
- **Obersevation,hypothesis and build the model.**
- **Based on hypothetical algorithm to simulate and compared with the pratical data.**
- **Theoretical anlysis and conclusion.**

## II. Data Collection and processing

The initial data [1] comes from 15 classmates and everyone sends a 10yuan WeChat red envelope with the number of it is 15. Use the excel and matlab to draw the below pictures. The first picture is a histrgram about money and frequecy. Every line of the second picture is a red envelope that divide into 15 in snatching order.

## III. Obersevation,hypothesis and build the model.

By obserseving the above picture, the first historgram seems to fit the normal distrubution PDF and every order's money seems to evenly distributed within the same range. Then, using the matlab to do the gaussian fitting test.The code is shown below:

gaussian fitting test code

```
x=[data];% the initial data
normplot(x);
[muhat,sigmahat,muci,sigmaci]=normfit(x);
        %muhat and sigmahat are the expectation
         and variance, muci and sigmaci are the
         confidence interval.
y=normpdf(x, muhat,sigmahat);
        %y is the PDF of x
plot(x,y,'*');
```

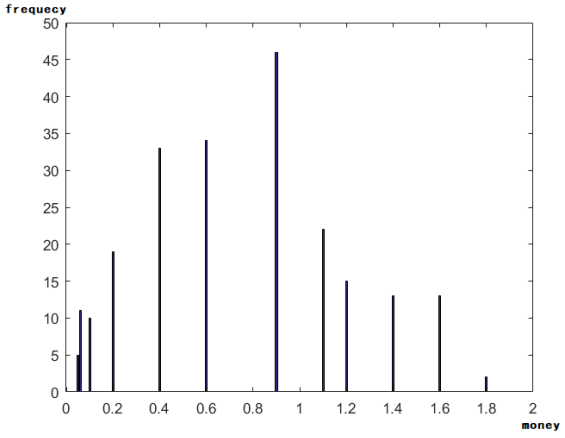


Fig. 1. Example 1 for Data visualization

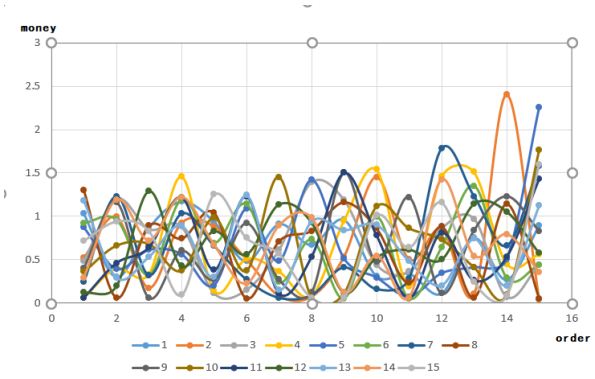


Fig. 2. Example 2 for Data visualization

The output is

$$muhat = 0.6667; sigmahat = 0.4627;$$
$$muci = [0.6059, 0.7275]; sigmaci = [0.4235, 0.5099]$$

From the Fig.3, we find the data mostly distributes around the line which means it fits normal distribution well. Then, the gaussian fitting test(Fig.4) also says the data could be normal distribution. From Fig.5, we can find the order has a little effect on the amount. Thus, we can assume the amount fits the normal distribution and the order fo snatching fits the uniform distribution and mainly discuss the distribution of amount.

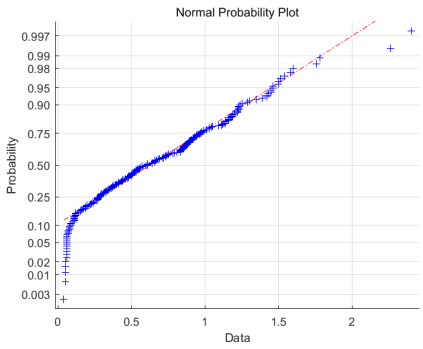


Fig. 3. normplot

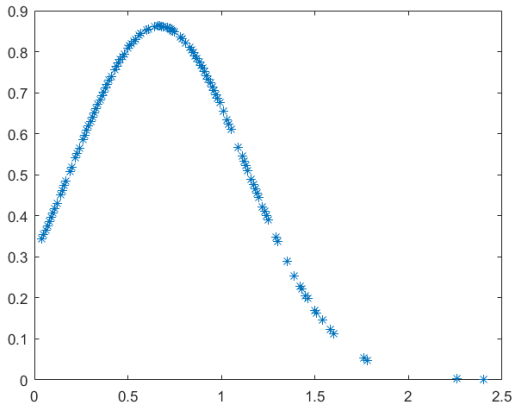

Fig. 4. gaussian fitting test

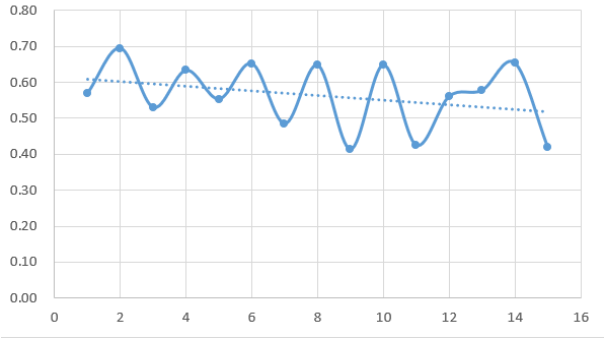

Fig. 5. every order's mean money

## IV. Based on hypothetical algorithm to simulate and compared with the pratical data.

I use the MATLAB to simulate the process of WeChat red envelope and the code [1] is shown below:

simulation code

```
function r=red_envelope(a,b)
r=[];
while (b>1)
    s=round(normrnd(0.6667,0.4627)*100)/100;
    if (0<s)&&(s<a)
       a=a-s;
            r=[r,s];
            b=b-1;
    end
end
r=[r,a];
```

The parameter $a$ of the function $red\_envelope$ is the amount of a red envelope and the parameter $b$ is the number of a red envelope divided. Then, we use the program to generate 50 sets of data with $a = 5, b = 6$ [2] and compare with the pratical data [3].

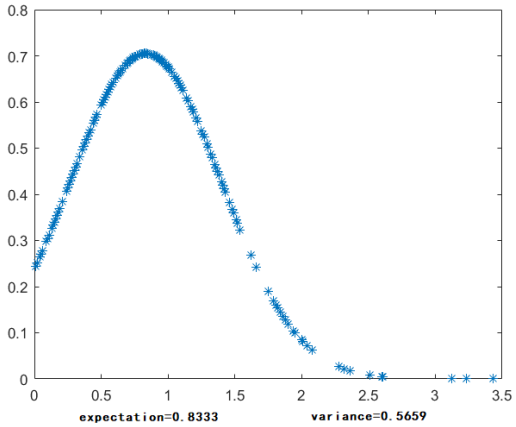


Fig. 6. simulation

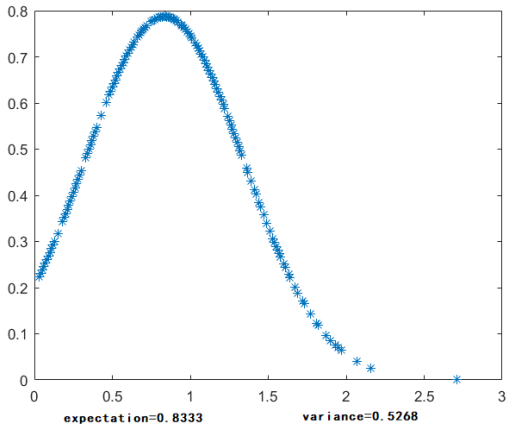


Fig. 7. pratical

From Fig.6 and Fig.7, the simulatation and pratical have the same expectation and similar variance (less than 7% error).Thus, the hypothesis is likely to be right.

## V. Theoretical anlysis and conclusion.

When I use the data to do the gaussian fitting test, I put all the data together regardless of order and times since we can prove that if $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$ are independent for $i \in [1, n]$, the distribution of $X_1, X_2, \cdots, X_n$ is $\mathcal{N}(\sum_{i=1}^{n} \mu_i, \sum_{i=1}^{n} \sigma_i^2)$. Thus, using all data together can make the result more accurate.

Accoding to my investigation, the order has a little effect on the amount and the amount fits the normal distribution, so it doesn't matter whether snatching the WeChat red envelope fast and most people get the amount around the expection. Since the amount of red envelope distribution makes most people get the amount around the expection, and few people get a huge amount, it can increase the playability and interesting and the WeChat red envelope is popular with people.

## ACKNOWLEDGMENT

During this project, I collaborated and discussed with my classmates Jin Yao and Tianyi Zhang. I also adopted part of contents from Zhihu [4] to Section III.The initial data comes from I with schoolmates Jiaxuan Wu, Sifan Wang, Keyi Yuan, Yihan Wang, Songjie Xie, Chenfang Li, Jiachun Jin, Pei Lin, Xiangyu Zeng, Yihang Li, Tianyang Zhao,etc. The others data used to verification comes from Jin Yao.

## REFERENCES

[1] S. W. K. Y. Y. W. S. X. C. L. J. J. P. L. X. Z. Y. L. T. Z. e. Jiawei Liang, Jiaxuan Wu, the excel initial data,https://github.com/Mine4ever/red-envelope-data.

[2] M. program, the excel generate data,https://github.com/Mine4ever/red-envelope-data.

[3] J. Yao, the excel others data,https://github.com/Mine4ever/red-envelope-data.

[4] "Discussions in zhihu," https://www.zhihu.com/question/22625187.