Analysis and Verification of Probability in Chinese Double Chromosphere Lottery

Yuhan Chen

Collège du Léman
**Corresponding author. Email: yuhan_chen@cdl.ch*

ABSTRACT
Probability theory is an important mathematical branch to study the law of random phenomena. Its application widely exists in our life. As a typical random experiment of classical probability model, lottery is an excellent material for the study of probability theory. Similarly, the study of various probability problems in lottery can also guide the real lottery activities. This paper will explore the probability information hidden in the double chromosphere, one of the two state-run lotteries, from the perspective of probability theory and answer the important question that many lottery people are concerned about—what is the specific odds of winning the lottery? Is it realistic to buy lottery tickets to get rich? Finally, the relevant conclusions are verified by publishing the data of double chromosphere lottery.

Keywords: *Lottery, Double chromosphere, Probability, Classical probability model.*

1. INTRODUCTION

Lottery tickets are printed with numbers or graphics, purchased voluntarily by people, and can prove that the buyer has a written certificate to obtain rewards according to specific rules. It is an entertainment game of fair competition based on equal opportunities. Lottery tickets first appeared in ancient Rome 2000 years ago. The birth of China's welfare lottery can also be traced back to the period of economic reform in the 1980s[1].

The birth of lottery is accompanied by opportunity and uncertainty, which is precisely the important reason why lottery attracts countless people to participate. According to the Ministry of Finance of China, the total lottery sales in China reached 77.4 billion dollars and 64.9 billion dollars respectively in 2018 and 2019. Although the lottery sales were influenced by COVID-19 in 2020, it still reached 51.4 billion dollars. The main attraction of the lottery is that a person may become a millionaire just after spending only a few dollars[2]. As a result, an increasing number of people dream of getting rich by buying lottery tickets. But what are the specific odds of winning the lottery? Is it realistic to get rich by buying lottery tickets? Can we answer these questions within the scope of science? The answer is yes. Probability theory is an important subject to scientifically solve the problem of random phenomena. This paper will explore the probability information hidden in the double chromosphere, one of the two state-run lottery tickets, from the perspective of probability theory, and verify the results through public data. The second chapter explains the background knowledge used. The third chapter calculates the probability information of double chromosphere in detail. The fourth chapter verifies the probability calculation through the official public double chromosphere data. The fifth chapter is the conclusion of the paper.

2. BACKGROUND

2.1. Double Chromosphere

The Double Chromosphere is one of the most popular playing ways of China Welfare Lottery. There is a group of red balls which contains 33 numbers and a group of blue ball which contains 16 numbers. Gamblers need to choose a number from 1 to 33 in the group of red balls and a number from 1 to 16 in the group of blue ball. Those six red ball numbers cannot repeat each other. Then, six red ball numbers and one blue ball number make up a single betting number. If the number you selected is the same as the winning number, you can win the prize. The order of the selected number can be different from the winning number. Each betting costs two yuan. See Table 1 for specific winning rules[3].

Y. Jiang et al. (Eds.): ICEDBC 2022, AEBMR 225, pp. 1040–1045, 2022.
https://doi.org/10.2991/978-94-6463-036-7_153

Table 1. Award distribution table

Award level	Wining conditions		Award distribution
	Red ball	Blue ball	
The first prize	6 winning numbers	1 winning number	When the prize pool fund is less than 100 million yuan, the total amount of the first prize award is 75% of the current high prize award and the sum of the accumulated funds in the prize pool. When the prize pool fund is equal or more than 100 million yuan, the total amount of the first prize includes two parts, one is 55% of the current high prize award and the sum of the accumulated funds in the prize pool. The other part is 20% of the high prize of the current period. The minimum prize is equivalent to the second prize. The highest prize for single betting is 10 million yuan
The second prize	6 winning numbers	0 winning number	The total amount of award is 25% of the current high prize. The maximum limit of single betting is 5 million yuan.
The third prize	5 winning numbers	1 winning number	The award for single betting is fixed at 3000 yuan.
The fourth prize	5 winning numbers	0 winning number	The award for single betting is fixed at 200 yuan.
	4 winning numbers	1 winning number	
The fifth prize	4 winning numbers	0 winning number	The award for single betting is fixed at 10 yuan.
	3 winning numbers	1 winning number	
The sixth prize	2 winning numbers	1 winning number	The award for single betting is fixed at 5 yuan.
	1 winning number	1 winning number	
	0 winning number	1 winning number	

2.2. The classical probability

If the probability model satisfies the following two conditions， it is called classical probability model:

1) All events are equally likely.

2) There are a finite number of basic events that can occur.

The basic formula of classical probability is:

$$P(A) = \frac{f}{n} \tag{1}$$

In fact, A is a specific event, n is the total number of basic events, f is the frequency of event a, and P (A) is the probability of event A. In the double chromosphere example, A can be set as the winning event, n is the number of all ball selection situations, and f is the number of all winning situations. P (A) represents the winning probability[4-5].

3. PROBABILITY ANALYSIS

3.1. Analysis of winning probability of double chromosphere

Let N be the sample space of all possible ball selection of the double chromosphere. A is the winning event. A0, A2..., A6 is a complete event group of N (where a1-a6 represents the winning event from the first prize to the sixth prize, and A0 represents the non-winning event). There are the following conclusions.

$$A = A1 \cup A2 \cup A3 \cup A4 \cup A5 \cup A6$$

$$N = A0 \cup A1 \cup A2 \cup A3 \cup A4 \cup A5 \cup A6$$

According to the combination formula:

$$n = C_{33}^{6}C_{16}^{1} = 17721088$$

Then according to formula (1):

First prize probability: $P(A1) = \frac{f1}{n} = \frac{C_6^6C_{27}^0C_1^1C_{15}^0}{C_{33}^6C_{16}^1} = \frac{1}{17721088}$

Second prize probability : $P(A2) = \frac{f2}{n} = \frac{C_6^6C_{27}^0C_1^0C_{15}^1}{C_{33}^6C_{16}^1} = \frac{15}{17721088}$

Third prize probability: $P(A3) = \frac{f3}{n} = \frac{C_6^5C_{27}^1C_1^1C_{15}^0}{C_{33}^6C_{16}^1} = \frac{162}{17721088}$

Fourth prize probability : $P(A4) = \frac{f4}{n} = \frac{C_6^5C_{27}^1C_1^0C_{15}^1+C_6^4C_{27}^2C_1^1C_{15}^0}{C_{33}^6C_{16}^1} = \frac{7695}{17721088}$

Fifth prize probability : $P(A5) = \frac{f5}{n} = \frac{C_6^4C_{27}^2C_1^0C_{15}^1+C_6^3C_{27}^3C_1^1C_{15}^0}{C_{33}^6C_{16}^1} = \frac{137475}{17721088}$

Sixth prize probability : $P(A6) = \frac{f6}{n} = \frac{C_6^2C_{27}^4C_1^1C_{15}^0+C_6^1C_{27}^5C_1^1C_{15}^0+C_6^0C_{27}^6C_1^1C_{15}^0}{C_{33}^6C_{16}^1} = \frac{1043640}{17721088}$

∵ A1, A2,..., a7 belong to the same complete event group and are independent of each other

∴ according to the formula of probability addition:

$$P(A \cup B) = P(A) + P(B) - P(AB) \qquad (2)$$

Total winning probability:

$$P(A) = P(A1) + P(A2) + P(A3) + P(A4) + P(A5) + P(A6) = \frac{1188988}{17721088} \approx 6.71\%$$

As can be seen from the above formula, even the probability of winning the lottery is very small, let alone the probability of winning the first prize. But even if the extremely low probability of one in ten million still can not stop people's dream of winning the prize. More people buy lottery tickets for a long time in the hope that they can get paid. Next, we will analyze from the perspective of probability how long to buy the lottery before we have a chance to win the first prize.

3.2. The number of times you need to buy a double chromosphere to win a grand prize

Let X1, X2, X3..., Xn be independent and identically distributed random variables, indicating the probability of winning the first prize when buying a lottery. The probability distribution of Xi is shown in the table below.

Table 2. Probability distribution of random variables of first prize in each lottery

xi	0	1
p	$\frac{17721087}{17721088}$	$\frac{1}{17721088}$

Where 0 and 1 represent not winning the first prize and winning the first prize respectively.

Let the random variable $X = \sum_{i=1}^{n} Xi$ indicates the number of times to win the first prize when buying n lottery tickets. X obeys binomial distribution, i.e. $X \sim B(n, p)$, where $p = \frac{1}{17721088}$.

Assuming that the probability of winning the grand prize is greater than 90%, $P(X >= 1) > 90\%$ is required. According to the binomial distribution probability function.

$$p(k, n, p) = C_n^k p^k (1-p)^{n-k} \qquad (3)$$

$\because p = \frac{1}{17721088}$, $P(X >= 1) = \sum_{i=1}^{n} C_n^i p^i (1-p)^{n-i}$

∴ Required calculation $\sum_{i=1}^{n} C_n^i p^i (1-p)^{n-i} > 90\%$, as well as $(1-p)^n < 10\%$

∴ get n>=40804312

If you want to win the lottery, the probability of winning the lottery is greater than 90%. If each lottery ticket is random, you should buy at least 40804312. If you buy lottery tickets for 50 years, the number of drawing lottery per year is 156 times. Then buy at least $\frac{40804312}{50\times156} \approx 5232$ each lottery.

In the case of buying only one note in each issue, if you insist on buying the lottery for 50 years, the probability of winning the first prize is

n=50*156=7800, $P(X >= 1) = \sum_{i=1}^{7800} C_{7800}^i p^i (1-p)^{7800-i}$=0.044%

It can be seen that whether you want to win the first prize one time or insist on buying the lottery to win the first prize, it is very difficult. So what is the possibility of not pursuing the first prize but making profits through the double chromosphere lottery?

3.3. Double chromosphere revenue expectation

Let Y represent the random variable of profit when buying a lottery ticket, and the probability distribution of Yi is shown in the table below.

Table 3. Probability distribution of random variables of lottery profit per lottery ticket

y_i	0	6641037	143799	3000	200	10	5
p	$\frac{16532100}{17721088}$	$\frac{1}{17721088}$	$\frac{15}{17721088}$	$\frac{162}{17721088}$	$\frac{7695}{17721088}$	$\frac{137475}{17721088}$	$\frac{1043640}{17721088}$

Among them, the first prize and the second prize are floating bonuses, so the average value of each profit period from 17 years to now is calculated by point estimation as the profit expectation. The first prize is 6641037 and the second prize is 143799.According to the expectation formula

$$E(X) = \sum_{i=1}^{\infty} x_i p_i \qquad (4)$$

$$E(Y) = \sum_{i=0}^{6} y_i p_i$$

$$= 0 \times \frac{16532100}{17721088} + 6641037 \times \frac{1}{17721088} + 143799 \times \frac{15}{17721088} + 3000 \times \frac{162}{17721088}$$

$$+200 \times \frac{7695}{17721088} + 10 \times \frac{137475}{17721088} + 5 \times \frac{1043640}{17721088} = \frac{17415972}{17721088}$$

$$\approx 0.983$$

The expected profit of each lottery ticket is 0.983 yuan, while it takes 2 yuan to buy a lottery ticket. The expected rate of profit is

$$\frac{0.983}{2} = 49.1\%$$

4. DATA VALIDATION

4.1. Classical probability verification

The author collected the data of 17001-22029 double chromosphere lottery from 17 to 22 years[6], and collected the number of prize numbers. The results are shown in Figure 1 and Figure 2.

Although the occurrence frequency of each number is not exactly the same, it is relatively similar on the whole. The assumption that double chromosphere is a classical probability model with equal probability of each basic event is credible.

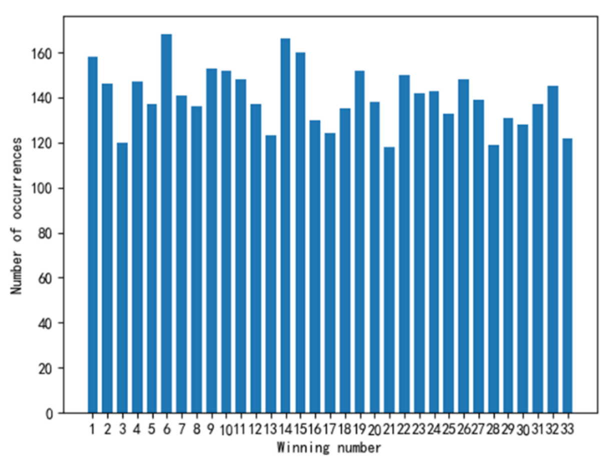

Figure 1 frequency of occurrence of red ball numbers of double chromosphere in phase 17001-22029

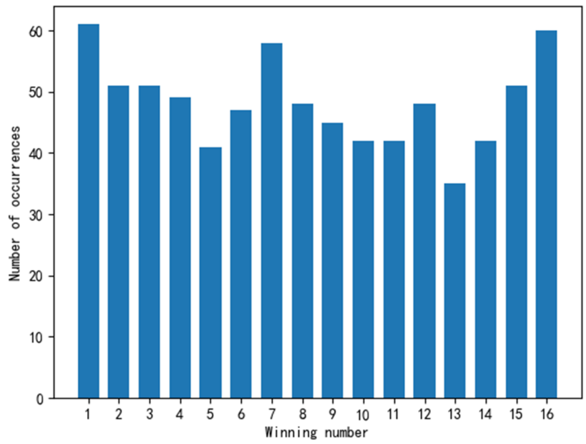

Figure 2 frequency of occurrence of green ball numbers of double chromosphere in phase 17001-22029

4.2. verification of winning probability

The author collected the lottery data of 17001-22029 double chromosphere from 17 to 22 years to verify the winning probability. According to the data, the sales volume in five years was 277640252354.00 yuan, including 13882012677 lottery tickets of double chromosphere, including 7687 tickets of the first prize.

The probability of winning the first prize is: $\frac{7687}{138820126177} \approx 0.000005537\%$

It is very close to the theoretical first prize probability of $\frac{1}{17721088} \approx 0.0000056\%$

There are 120391 tickets for the second prize, and the probability of the second prize is $\frac{120391}{138820126177} \approx$ 0.000088%.

It is very close to the theoretical second prize probability of $\frac{15}{17721088} \approx 0.0000846\%$ calculated previously.

In addition, the expected rate of return of 49.1% calculated in Section 3.3 is also consistent with 49% of the officially released sales as a bonus.

5. CONCLUSION

In this paper, we have made a detailed exploration of the double chromosphere game and its hidden probability information. Firstly, the double chromosphere purchase problem is modeled, combined with the classical probability, and the winning probability, rate of return, expected profit and so on are calculated in detail. Finally, the relevant probability calculation is verified by the official data. It comes to the conclusion that buying lottery tickets is a project with a very low winning probability and a very low rate of return[7-8]. As an entertainment, we should not be too persistent in winning the lottery[9].

REFERENCES

[1] Rong.Guo. Huilan.Wang. (2018). Probability analysis of the double chromosphere lottery winning numbers. Journal of Taiyuan Normal University, Vol.17, No.1

[2] Mihua.Chen. (2007). The charm of probability-Start with winning the lottery. Mathematics Teaching Newsletter,267,58-61.

[3] Fiedler K, Unkelbach C. Lottery attractiveness and presentation mode of probability and value information[J]. Journal of Behavioral Decision Making, 2011, 24(1): 99-115.

[4] Morris H, Mark J. Schervish. (2002) Probability and statistics. New York: Pearson Education,182-200

[5] Zhilian.Guo, Meilong.Yang. (2020). Analysis of the calculation method of classical probability. China Academic Journal Electronic Publishing House, 140. http://www.cwl.gov.cn/

[6] Mumpower J L. Lottery Games and Risky Technologies: Communications About Low-Probability/High-Consequence Events[J]. Risk Analysis, 1988, 8(2): 231-235.

[7] Farrell C, Cowley E, Edwardson M. Strategies to improve the probability of winning a lottery: Gamblers and their illusions of control[J]. ACR European Advances, 2005.

[8] Alarie Y, Dionne G. Lottery decisions and probability weighting function[J]. Journal of Risk and Uncertainty, 2001, 22(1): 21-33.