Analysis and forecast for growth of lottery sales based on optimal control model

Zhang Liu1, 2, Canhua Li1, Jian Wu1

1 Computer and Information Engineering College, Jiangxi Agricultural University, Nanchang 330045, China
2 School of Mathematics and Statistics, Wuhan University, Wuhan 430072, China

Corresponding author and e-mail: Zhang Liu, liuzhang1006@163.com.

Abstract. Based on the optimal control theory, this paper analyzes the relationship among the growth of lottery sales and the accumulation of capital and the rate of accumulation. By the introduction of Hamilton function, we established a mathematical model and calculate the satisfied equation. Using a simple model, we realized data fitting for the corresponding formula by Matlab through dealing with the sales and the funds of prize pool of the super lotto from 2014 to 2018. We made some forecasts according to the results.

1. Introduction
The issuance of lottery tickets is now a common form of fund-raising. It is usually based on legal funds and fairness, and redistributes social idle funds. Since the recovery of the lottery industry in China, it has been franchised by the China Welfare Lottery Distribution Center and the China Sports Lottery Distribution Center. According to the data of the Ministry of Finance, the cumulative sales of lottery tickets in China for the whole year of 2018 was 511.472 billion yuan, a year-on-year increase of 20%, and the lottery volume increased. Fucai and Sports Lottery have played a big role in the development of social welfare and sports, and have also made great contributions to China's philanthropy. With the continuous development of the lottery business, the mathematics problems contained in it have gradually attracted the interest of many experts and scholars. For example, in 2002, the National College Student Modeling Competition was based on the question of the probability of winning the lottery in the "mathematics in lottery". Meirui et al. (2010) used ruin probability to analyze the risk problem in lottery; Zou and Yin (2006) studied the probability of consecutive numbers in computer welfare lottery; Hu and Qian (2007) discussed the mathematical analysis and investment strategy of lottery tickets; Zhang and Zeng (2006) studied the optimal measurement and its application in the consumer behavior of lottery; Xue et al. (2006) used the “Xinjiang Style” computer welfare lottery to conduct an in-depth study on the mathematical analysis and application of lottery tickets; Li and Li (2002) investigated the composition and purchase behavior of Beijing sports lottery buyers, and studied the favorable and unfavorable factors for the purchaser and made recommendations through the analysis of the results.

In recent years, there have been more research issues in the field of lottery. For example, Jia and Xie (2009) applied multiple regression analysis to analyze the correlation between the welfare lottery sales in China's various provinces and the five major categories of economic life's predictors; Wei and
Lei (2013) analyzed the sales of sports lottery and the spatial heterogeneity of its driving factors in China by using exploratory spatial data analysis (ESDA) method and geographic weighted regression (GWR) model through various index data that affect social development; Han (2015) used the panel data-based method to establish a dual logarithmic linear regression model and a general linear regression model to quantitatively analyze the factors affecting the overall lottery sales and Internet lottery sales in China. Zhang (2017) studied the factors affecting the sales of welfare lottery tickets in China under the capping system, and proposed some reasonable countermeasures; Liu (2017) explores the distribution strategy of Chinese sports lottery bonuses and how to help the lottery department design the optimal winning probability and winning prizes by constructing the general equilibrium model of the lottery market to promote the scientific development of the lottery industry; Li (2019) studied the interactive effects of China's lottery market and stock market from the perspective of behavioral finance, and combined with theoretical and empirical results, proposed policy recommendations for improving related fields. By now, the research focus on the lottery market is mainly focused on the sales plan of lottery tickets and the direct consumption behavior of lottery players. The analysis of the lottery market is a comprehensive multidisciplinary issue involving mathematics, economics, sociology and more. This paper starts from the current most popular and high-selling lottery ticket in the Chinese market----Super Lotto, and analyzes the relationship between the growth of lottery sales, accumulated funds and the cumulative rate. In addition, this paper also introduces the Hamilton function to establish a mathematical model, calculates the equation that it satisfies, and uses the simplified model to realize the data fitting of the equations by Matlab. Through the collection and consolidation of the sales volume and prize pool funds of the Super Lotto between 2014 and 2018, the future sales were analyzed and forecasted. Finally, based on the relevant results, reasonable comments and suggestions was proposed for the healthy development of the Chinese lottery market.

2. General model

The sales amount of a certain lottery ticket at time $t$ is recorded as $x(t)$, and the portion for accumulating funds (referring to the balance of the current prize pool funds) is recorded as $y(t)$, and the proportion of accumulated funds in the lottery sales is recorded as $u(t) = x(t)/y(t)$ (Accumulation rate), the optimization control model is established below. The purpose of modeling is to find the optimal accumulation rate $u(t)$, so that the total sales of the lottery market $x(t)$ will grow bigger and faster.

The growth rate $z(t)$ of the lottery sales revenue depends on the total income $x(t)$ and the cumulative rate $u(t)$ of the lottery sales at that time, so their relationship can be expressed as

$$z(t) = f(t, x(t), u(t)).$$

(1)

Considering that for a period of time $T$ (such as one year or three years), to make $x(t)$ from the initial value $x(0)$ to the largest possible value $x(T)$, the initial condition is

$$x(0) = x_0,$$

$$\max x(t) = \{\max x(t) : 0 \leq t \leq T\}$$

(3)

Then the problem of the study can be summarized as how to find the extreme value of (3) under the conditions of (1) and (2).

The above problem is equivalent to its dual problem: under the fixed endpoint condition $x(T) = x_1$, the time $T$ used to make the total revenue of the lottery from $x_0$ to $x_1$ is the smallest, so
\[ x(0) = x_0, \quad x(T) = x_f, \quad G(u(t)) = \int_0^T dt. \]

The method of finding the extremum problem by function can express the above problem as follows. To find the optimal cumulative rate under conditions (1) and (4) to make the minimum of \( G \) in (5).

The Lagrange multiplier method is used to transform conditional extreme value into the unconditional extreme value, and the Lagrange multiplier factor \( \lambda(t) \) is introduced to obtain the function,

\[ L(x(t), u(t)) = \int_0^T \left[ 1 + \lambda(t) \cdot \left( f(t, x(t), u(t)) - z(t) \right) \right] dt. \]

Introducing the Hamilton function

\[ H(t, x(t), u(t)) = 1 + \lambda(t) \cdot f(t, x(t), u(t)), \]

so equation (6) can be written as

\[ L(x(t), u(t)) = \int_0^T \left( H - \lambda(t) \cdot z(t) \right) dt. \]

According to the necessary condition that the function (8) reaches the extreme value, the Euler equation is

\[
\begin{align*}
\frac{\partial(H - \lambda \cdot z)}{\partial x} - \frac{d}{dt}\left( \frac{\partial}{\partial z}(H - \lambda \cdot z) \right) &= 0, \\
\frac{\partial(H - \lambda \cdot z)}{\partial u} - \frac{d}{dt}\left( \frac{\partial}{\partial u}(H - \lambda \cdot z) \right) &= 0, \\
\frac{\partial H}{\partial x} + \lambda'(t) &= 0, \\
\frac{\partial H}{\partial u} &= 0,
\end{align*}
\]

that is,

\[ \lambda'(t) = -\lambda(t) \cdot \frac{\partial f(t, x, u)}{\partial x}, \]

so the whole problem is equivalent to solving the following equivalent equations

\[
\begin{align*}
\lambda'(t) &= -\lambda(t) \cdot \frac{\partial f(t, x, u)}{\partial x}, \\
\lambda(t) \cdot \frac{\partial f(t, x, u)}{\partial u} &= 0, \\
z(t) &= x'(t) = f(t, x, u), \\
x(0) &= x_0, \quad x(T) = x_f.
\end{align*}
\]

Solving the equations can find the optimal function \( u(t) \) and the state curve \( x(t) \).

3. Data and empirical analysis

3.1. Data collection and processing

This article takes the Super Lotto of China Sports Lottery as the research object. The sample taken by empirical modeling is the real-time sales data of Super Lotto from 2014 to 2018. The data comes from the Chinese lottery website. The lottery draw is for three issues a week (every Monday, Wednesday, and Saturday). Due to the different interval between each lottery, the single sales are affected by many...
factors. In order to accurately study the nature of the problem, we selected the one-month time as the smallest research unit from a macro perspective, and calculated the average value of multiple lottery sales in a month. After sorting and calculation, the data of the prize pool at the end of each month during the 2014-2018 period and the average sales of the monthly lottery are given. See Table 1 below. The trend chart of the change is shown in Figure 1-2.

| Table 1. Average monthly sales of Super Lotto and the prize pool balance during the 2014-2018 period. |
|---|---|---|---|
| month | Prize pool fund balance(yuan) | Average sales (yuan) | month | Prize pool fund balance(yuan) | Average sales (yuan) |
| 2014-01 | 324,041,671 | 127,895,072.46 | 2016-07 | 3,206,441,463 | 179,966,295.00 |
| 2014-02 | 325,669,431 | 128,793,210.11 | 2016-08 | 3,449,353,717 | 180,868,686.64 |
| 2014-03 | 293,099,517 | 132,254,991.71 | 2016-09 | 3,502,839,878 | 182,755,096.08 |
| 2014-04 | 427,750,617 | 129,852,745.38 | 2016-10 | 3,742,471,490 | 180,594,100.07 |
| 2014-05 | 576,508,734 | 156,968,476.69 | 2016-11 | 3,857,570,669 | 188,990,544.69 |
| 2014-06 | 729,191,341 | 175,685,479.15 | 2016-12 | 3,235,556,086 | 200,199,759.54 |
| 2014-07 | 929,910,085 | 184,106,520.77 | 2017-01 | 3,296,873,997 | 192,723,541.73 |
| 2014-08 | 595,465,040 | 175,753,372.62 | 2017-02 | 3,427,151,759 | 187,535,170.91 |
| 2014-09 | 759,165,112 | 165,263,532.46 | 2017-03 | 3,606,190,192 | 192,786,536.69 |
| 2014-10 | 980,248,197 | 184,053,199.69 | 2017-04 | 3,661,625,450 | 203,691,244.15 |
| 2014-11 | 1,251,566,141 | 214,223,099.92 | 2017-05 | 3,602,783,336 | 222,328,107.86 |
| 2014-12 | 1,135,404,076 | 209,299,295.64 | 2017-06 | 3,790,728,402 | 216,539,231.17 |
| 2015-01 | 922,796,421 | 194,949,233.77 | 2017-07 | 3,970,858,480 | 204,014,110.00 |
| 2015-02 | 1,052,353,522 | 182,887,707.56 | 2017-08 | 4,125,812,710 | 203,501,900.46 |
| 2015-03 | 1,230,022,010 | 158,525,391.85 | 2017-09 | 4,288,325,250 | 209,612,396.31 |
| 2015-04 | 1,201,607,522 | 162,773,880.92 | 2017-10 | 4,224,508,242 | 201,946,730.69 |
| 2015-05 | 1,275,356,607 | 180,416,114.77 | 2017-11 | 4,219,678,773 | 213,370,776.92 |
| 2015-06 | 1,369,791,223 | 173,538,394.23 | 2017-12 | 4,478,126,302 | 220,315,411.85 |
| 2015-07 | 1,470,621,295 | 155,740,439.00 | 2018-01 | 4,726,825,681 | 213,185,240.50 |
| 2015-08 | 1,449,994,950 | 162,825,559.50 | 2018-02 | 5,016,451,627 | 213,498,688.56 |
| 2015-09 | 1,512,320,825 | 163,098,424.62 | 2018-03 | 5,297,754,793 | 217,227,654.31 |
| 2015-10 | 1,707,307,787 | 165,095,955.31 | 2018-04 | 5,368,233,194 | 244,673,841.92 |
| 2015-11 | 1,964,244,295 | 171,323,855.00 | 2018-05 | 5,703,821,164 | 248,867,827.31 |
| 2015-12 | 2,174,681,996 | 176,055,021.38 | 2018-06 | 5,958,857,007 | 220,105,063.77 |
| 2016-01 | 2,308,385,677 | 181,598,760.08 | 2018-07 | 6,175,967,288 | 207,896,708.77 |
| 2016-02 | 2,423,340,376 | 179,281,063.40 | 2018-08 | 6,308,725,026 | 207,638,184.31 |
| 2016-03 | 2,586,646,951 | 180,529,878.46 | 2018-09 | 6,552,450,316 | 208,220,686.54 |
| 2016-04 | 2,815,398,495 | 203,681,497.23 | 2018-10 | 6,594,664,860 | 215,470,012.50 |
| 2016-05 | 3,132,059,059 | 208,380,693.69 | 2018-11 | 6,827,035,600 | 221,977,626.00 |
| 2016-06 | 3,255,540,118 | 183,587,770.31 | 2018-12 | 7,118,917,925 | 213,793,699.21 |
Figure 1. The balance of the lottery prizes at the end of each month from 2014 to 2018 (the horizontal axis is the month and the vertical axis is the month-end balance).

Figure 2. The average monthly sales volume from 2014 to 2018 (the horizontal axis is the month and the vertical axis is the monthly average sales).

3.2. Empirical simplified model
The following questions are discussed by a simplified model. In order to make the function $f(t,x(t),u(t))$ discussed above have a concrete and concise expression, it is possible to set the sales growth of the lottery ticket to have such a rule: when the accumulation rate $u$ is small, the relative growth rate $z/x$ of sales increases with the increase of $u$. This is because the accumulation of funds in the prize pool promotes the consumption of the lottery. However, as $u$
becomes larger, the growth rate of \( \frac{z}{x} \) is slower and slower. When \( u \) increases to a certain extent, \( \frac{z}{x} \) becomes smaller. This is because the total amount of lottery consumption is limited, and the increase in the amount of input will make the accumulated prize pool less (the number of first prizes will increase), and this result is consistent with the actual average monthly sales (Figure 2). Therefore, the above model can be simplified to

\[
\frac{z}{x} = u(r - su),
\]

(12)

where \( r, s \) is a constant and \( x \) is the total monthly sales. \( z \) is the growth rate and \( u \) is the cumulative rate. The sales data of 2014-2018 is counted, and the values of \( r \) and \( s \) can be calculated by statistical methods.

The next step is to further simplify the equation to \( y = -sx^2 + rx \). Using the known data, we can find the values of 60 sets of \( y \) and \( x \), and then use Matlab to achieve least squares fitting. Then the running result is obtained: \( r = 0.0163, \ s = -0.0149 \); combining (1), (4) can obtain \( f = u(s - ru)x \), thus we obtain the equations

\[
\begin{align*}
\lambda'(t) &= -\lambda(t) \cdot u(r - su), \\
\lambda(t) \cdot (r - 2su)x &= 0, \\
z &= u(r - su)x, \\
x(0) &= x_0, \quad x(T) = x_i.
\end{align*}
\]

Solve it we can get

\[
\begin{align*}
u(t) &= r/2s, \\
x(t) &= x_0 \exp \{r^2 t/4s\}, \\
T &= 4s (\ln x_i - \ln x_0)/r^2.
\end{align*}
\]

(14)

That is, when the accumulation rate is a fixed value \( r/2s \), then the lottery sales revenue increases in the form of an approximate exponential function, and the shortest time is \( T \).

### 3.3. Empirical Analysis: Prediction and Analogy

According to the conclusions in (12-14), combined with the data in Table 1, we can predict the sales of Super Lotto in the future. Using Matlab programming, we can predict the relevant value of monthly average sales for the first nine months of 2019, see Table 2. Note that the actual sales data for the first nine months of 2019 are also given in Table 2, we compare the two and present the relative errors in Table 2.

| month | Prize pool fund balance(yuan) | Forecast of average sales per month(yuan) | The true value of the average monthly sales per month (yuan) | Relative error |
|-------|-------------------------------|------------------------------------------|----------------------------------------------------------|---------------|
| 2019-01 | 7261573924                    | 214865506.06                             | 215615698.38                                             | 0.3%          |
| 2019-02 | 7260411038                    | 216714259.02                             | 231554865.11                                             | 6.4%          |
| 2019-03 | 6999024180                    | 220735472.18                             | 266684995.15                                             | 17.2%         |
| 2019-04 | 5385687606                    | 231912314.46                             | 272816492.85                                             | 14.9%         |
| 2019-05 | 4808403292                    | 246667997.12                             | 281250266.08                                             | 12.2%         |
| 2019-06 | 4372011525                    | 263076639.05                             | 283034572.00                                             | 7%            |
| 2019-07 | 3755102164                    | 275946565.77                             | 275733593.79                                             | 0.07%         |
| 2019-08 | 2965514601                    | 278208715.43                             | 278215424.67                                             | 0.02%         |
| 2019-09 | 2618860320                    | 27958464.14                              | 289671755.15                                             | 3.4%          |
It can be seen from the forecast results that the sales in the first three months of 2019 are basically in line with our forecast, especially the relative error of the forecast results in the first month and the next few months is basically within 3%. However, with the difference in prediction time and the influence of various random factors in reality, the relative error of prediction will be significantly larger in some months. Therefore, after the model is established, it is necessary to continuously replenish new data and update the values of \( r \) and \( s \) in the model to realize dynamic prediction of the model, thereby maintaining high prediction accuracy. On the other hand, we also forecast the sales of the Super Lotto for a long time in the future. We forecast the sales data of the Super Lotto in October-December 2019 and December of the next three years, see Table 3 below. In Figure 3, we present a comparison of projected sales and real sales for the Super Lotto during 2019. As can be seen from the figure, the sales forecast of the Super Lotto is increasing as a whole, and the increasing trend is not completely exponential growth, but according to the law of the expression (12-14); and the real monthly average sales are affected by a variety of random factors, which makes analysis errors between the predicted and actual values. However, in the long run, the forecast data of this method is basically consistent with the real situation, indicating that the forecasting effect of this method is relatively good, which can provide a reasonable decision basis for the benign operation and supervision of China's lottery business.

**Table 3.** Forecasts sales for October-December 2019 and the next three years.

| Time     | 2019-10 | 2019-11 | 2019-12 | 2020-12 | 2021-12 | 2022-12 |
|----------|---------|---------|---------|---------|---------|---------|
| Forecasts| 281663836.40 | 285321152.50 | 290718042.18 | 323843696.56 | 348386681.91 | 366567393.29 |
| asts     | 40      | 50      | 18      | 56      | 91      | .29     |

**Figure 3.** Comparison of projected sales and real sales for the Super Lotto during 2019.

4. Related conclusions analysis and recommendations

Recently, lottery has become the sixth largest industry in the world. The progress of the lottery business is not only related to the development of China's economy, but also closely related to sports, politics, environmental protection, charity, social welfare, etc. From the perspective of statistical mathematics, analyzing and forecasting the trend of China's lottery market is also a realistic issue that we are concerned about. Based on the above analysis, we put forward some opinions and suggestions on the sound development of the Chinese lottery market.

Reasonably adjust the amount of bonuses and the proportion of awards. It can be seen from the above that the growth rate of lottery sales \((z/x)\) is closely related to the cumulative rate \((u)\) of the
lottery prize pool. That is, how much funds are left in the prize pool during each lottery draw, as the prize pool funds enter the next lottery fund flow, and how much funds are returned to the lottery pockets through the award. It is necessary to properly stimulate the consumption of lottery players and reserve reserves for future capital turnover. This paper calculates that when the cumulative rate is a fixed value $r/2s$, that is, $\mu = -57338$, the growth rate of sales is the largest.

Appropriate increase in bonuses to stimulate consumption. Super Lotto has added an extra bonus every year from 2014 to 2018, respectively

From the period of 14052 to 14066 (2014-5-5 to 2014-6-9), the additional bonus amount is 150 million yuan.

From the period of 15051 to 15072 (2015-5-4 to 2015-6-24), the additional bonus amount is 360 million yuan.

From the period of 16042 to 16061 (2016-4-11 to 2016-5-28), the additional bonus amount is 500 million yuan.

From the period of 17044 to 17063 (2017-4-17 to 2017-6-3), the additional bonus amount is 500 million yuan.

From the period of 18041 to 18060 (2018-4-9 to 2018-5-26), the additional bonus amount is 600 million yuan.

Comparing the time interval of the above additional bonuses with the corresponding sales, it can be clearly seen that the sales during the extra bonus period is significantly higher than the usual sales, and the sales tend to increase slowly. In addition, the enthusiasm of the lottery has gradually increased, so the appropriate extra bonus is beneficial to the increase in lottery sales and the healthy development of the lottery market.

Contact the actual and adjust the strategy. Sorting the lottery sales data in recent years, we can find that the sales of lottery tickets are not only affected by the prize pool funds and the amount of lottery, but also interfere with many realities, such as newly issued laws and regulations, political patterns, festivals and holidays. Therefore, the sales strategy of lottery tickets cannot be uniform, but it is necessary to adjust strategies according to the actual situation to achieve greater growth in sales and better serve social welfare undertakings.

Acknowledgement
This work is supported by Jiangxi Science and Technology Plan Fund Project (GJJ20180201).

References
[1] MEI R., ZHANG C., ZHANG W.H. Analysis of Risk Issues in Lottery Using Ruin Probability. Journal of Hebei North University (Natural Science Edition). 2010,(1).
[2] Zou Z.M., Yin T.K. The Probability of Serial Numbers in Computer Welfare Lottery. Mathematical Statistics and Management. 2006, (4).
[3] Hu R.S., Qian G.H. Mathematical Analysis and Betting Strategies of Lottery Tickets. Journal of Changshu Institute of Technology. 2007, (10).
[4] Zhang D.M., Zeng Z.L. Optimization Measurement and Its Application in Comparative Study of Consumer Behavior in Lottery. Statistics and Decision Making. 2006, (15).
[5] Xue C.G, Dong X.Y., Zhang G.Y. Mathematical analysis of lottery tickets and its application statistics and decision making. 2006, (1).
[6] Li Zhi, Li Yong. Analysis and research on the current situation of computer sports lottery buyers in Beijing. Journal of Beijing Sport University 2002, (2).
[7] Jia C., Xie Z.J. Analysis of Factors Affecting China's Welfare Lottery Sales and Forecast Based on Residual Principal Component Analysis. Mathematical Statistics and Management, 2009, 2: 191-203.
[8] Wei D.Y., Lei W. Research on Spatial Heterogeneity of China Sports Lottery Sales and Its Driving Factors Based on ESDA-GWR. Journal of Tianjin Physical Education Institute, 2013, 28(4): 304-309.
[9] Han J.J. Analysis of the Factors Affecting the Sales of Lottery Tickets and the Strategy of the Development of Lottery Industry in China. China University of Mining and Technology, 2015.

[10] Zhang C.G. Research on the Factors Affecting the Lottery Sales in China under the Capping System. Southwest Jiaotong University. 2017.

[11] Liu S.W. Research on Distribution Strategy of Chinese Sports Lottery Bonus under the Background of Uncertainty. Shandong University. 2017.

[12] Li Y.Z. Research on the Interaction Effect between China's Stock Market and Lottery Market Based on Behavioral Finance. Shanghai Normal University. 2019.

[13] Jiang Q.Y. Mathematical Model. Higher Education Press. 1993

[14] Shen J., Zhou T.H. Mathematical Prediction Methods in Lottery Tickets. Journal of Liaocheng Teachers College (Natural Science Edition), 2002, (1).

[15] Yu Q., Zhou F.Y., Zhang Y.S. Mathematical Model of Lottery Analysis. Journal of Qiannan Normal College for Nationalities. 2002, (6).