

# Data Processing Method in the Application of the Mathematical Contest in Modeling

Xiaoling Xu<sup>1, a</sup>

<sup>1</sup>Xi'an Innovation College of Yan'an University, China

<sup>a</sup>17113525@qq.com

**Keywords:** Data processing, mathematical modeling, modeling contest, application

**Abstract.** Mathematical modeling is a process to analyze and solve problems, as training students' mathematical contest in modeling ability of solving practical problems more and more get the attention of the colleges and universities. Numerical processing method is not only applied in the process of the model, it also has strong for the establishment of the model of the directive. The research data of processing method in the application and mathematical modeling would provide the level of mathematical contest in modeling, improve university students' ability to deal with data, improve university students' ability to solve practical problems. This paper introduces the commonly used method of data modeling and numerical processing method is discussed with an example in the application of mathematical modeling.

## Introduction

Mathematics is the study of quantitative relationship between the real world and the space form of science, in the long history of its emergence and development, has been closely related to all kinds of application problems. Mathematics not only lies in the abstraction of the concept, the characteristics of the logical rigor, the clarity of the conclusions and system integrity, and lies in its application universality [1]. Since the 20th century, with the rapid development of science and technology and the increasing popularity of computer, people more and more precise, to the requirement of all sorts of problems make mathematics application is more and more extensive and in-depth, especially in the 21st century knowledge economy era, the status of the mathematical sciences tremendous changes will occur, it is from the backup to the forefront of national economy, science and technology [1].

Model is to facilitate the practical problems, to cut a part of the objective things, abstract and alternatives of the prototype [2]. It is in order to reflect the features of the prototype of a certain part. Mathematical model is of a real object, use mathematical tools to the modeling of a mathematical structure. Modeling is a process to analyze problems and solve problems, when we describe the actual object in the process of change in time or space, forecast its future state, often differential equation model is established to solve the problem.

## Commonly used Methods for Data Processing

**Simple statistical method.** Need in the process of mathematical modeling, calculating the data of the mean, median, standard deviation, variance, the maximum, minimum, confidence, etc. If you are not familiar with Excel data processing function, may be what to do. Excel can do all kinds of data statistics, computing, processing and statistical graphics drawing, play a major role in mathematical modeling [2]. Under the default installation, Excel does not directly provide data analysis toolkit, for the first time when using need to click "tools" -> "add-ins" choice "analysis tool library", "planning" and "conditional sum wizard", "ok" for installation. And Spas statistical analysis software, often used for processing large data.

**The interpolation and fitting.** Interpolation in the mathematical modeling is a common data processing method, interpolation is on the basis of discrete data for continuous function, makes the continuous curve through all of the given discrete data points [3]. Interpolation is important methods, the discrete function approximation using it can function in the value of a finite number of points,

estimate the approximation of function on other points. There are a two-dimensional interpolation method, the Lagrange interpolation, the piecewise linear interpolation, cubic spline interpolation, etc.

**Regression analysis and time series analysis.** Regression analysis is the study of the statistics of the relationship between variables a statistical method [4]. It is from a set of dependent and independent variables and the observed data, looking for a functional, the statistical correlation between the variable is relationship between expressed.

**Digital simulation.** In the process of mathematical modeling, model test is a very important step. Often need to model test to simulate the model solution compared with the actual data, data modeling is to realize the model means a very important method of testing [5].

**Modeling method of calculus** [6]. Equation everyone not unfamiliar, all equations is to study the relationship between the known and unknown, along with the birth of calculus, people started to research the dynamic continuity problems, such as objects under certain conditions the motion state of change, real trajectories of artillery, rocket flight path, after an object under the action of gravity to do free fall down along with the change of time, etc. We through the use of differential equation of knowledge study the law of the change.

## Application of Data Processing in the Mathematical Contest in Modeling

**Examples of mathematical contest in modeling 1.** Airlines often can make the passengers free ticket. Passengers booked ticket, likely because of a variety of reasons, not to fly, so, when the plane takes off, there will be some empty seats to waste. In order to reduce the loss, airlines tend to adopt the overbooking ticket means, namely: allow passengers flight reservation number more than seats on the plane [7]. But, in so doing, the passengers booking will happen again by not on the plane, was "crowd out". To be "crowd out" the passengers, the airline must offer certain compensation.

The question now is: how airlines should adopt the overbooking strategy, to minimize their losses, the biggest profits?

The cost of a flight, fly to their destinations include fuel, crew and ground staff wages, the airport management fees, aircraft maintenance and so on, almost all of these have nothing to do with passenger Numbers, therefore, as an approximation, we can assume that the cost of each flight is a constant [3]. Due to the airline's profit is equal to (after deducting the compensation) ticket fee income minus the flight, when the flight cost is constant, the airlines to achieve maximum profits, can need not consider to fly, as long as (after deducting the compensation) ticket fee income reached maximum.

Assume that,

$g$  - per ticket price (as approximation, we do not consider the level of the seat, considered  $g$  to be a constant).

$b$  - for each "crowd out" the compensation of passengers.

$M$  - the total number of seats on the plane.

$N$  - to make the passengers flight reservation number (as is oversubscribed, so there will be  $N \geq M$ ).

$\xi$  - the actual number of passengers to fly ( $0 \leq \xi \leq N$  is a random variable).

$\eta$  - (after deducting the compensation) ticket fee income ( $\eta$  is related with  $\xi$ , yes  $\xi$  is function, is also a random variable).

$$\eta = f(\xi) = \begin{cases} g\xi & \text{when } 0 \leq \xi \leq M \\ gM - b(\xi - M) & \text{when } M < \xi \leq N \end{cases}$$

Set  $\xi$  probability distribution for  $P\{\xi = k\}$ ,  $k = 0, 1, 2, \dots, N$ . At this time, can find out the mathematical expectation of a  $\eta$ , namely the airlines (after deducting the compensation) the average income for airfare.

$$E\eta = Ef(\xi) = \sum_{k=0}^N f(k)P\{\xi = k\}$$

$$= \sum_{k=0}^M gkP\{\xi = k\} + \sum_{k=M+1}^N [gM - b(k - M)]P\{\xi = k\}.$$

Calculated on a per ticket price  $g$  (after deducting the compensation) ticket fee for average income.

$$\begin{aligned} \frac{E\eta}{g} &= \sum_{k=0}^M kP\{\xi = k\} + \sum_{k=M+1}^N [M - \frac{b}{g}(k - M)]P\{\xi = k\} \\ &= \sum_{k=0}^M kP\{\xi = k\} + \sum_{k=M+1}^N kP\{\xi = k\} - \sum_{k=M+1}^N kP\{\xi = k\} + \sum_{k=M+1}^N [M - \frac{b}{g}(k - M)]P\{\xi = k\} \\ &= \sum_{k=0}^N kP\{\xi = k\} - \sum_{k=M+1}^N [(k - M) + \frac{b}{g}(k - M)]P\{\xi = k\} \\ &= E\xi - (1 + \frac{b}{g}) \sum_{k=M+1}^N (k - M)P\{\xi = k\} \end{aligned}$$

Actual number of passengers to fly  $\xi$  what is distributed? There are  $N$  passengers booked a ticket, every passenger booked ticket may come by plane or by plane, to establish each passengers booked on the actual probability are  $p$  by plane, and as the approximation, to establish whether each passenger plane were independent of each other. In this way,  $N$  passengers will come by plane is an independent repeat test sequence ( $N$  heavy Bernoulli trials),  $\xi$  is  $N$  trials in times of success, so obey parameter  $(N, p)$  binomial distribution, that is  $\xi \sim b(N, p)$ . The probability  $\xi$  distribution for is,

$$P\{\xi = k\} = C_N^k p^k (1 - p)^{N-k}, \quad k = 0, 1, 2, \dots, N$$

Mathematical  $\xi$  expectation for  $E\xi = Np$ , so

$$\begin{aligned} \frac{E\eta}{g} &= E\xi - (1 + \frac{b}{g}) \sum_{k=M+1}^N (k - M)P\{\xi = k\} \\ &= Np - (1 + \frac{b}{g}) \sum_{k=M+1}^N (k - M)C_N^k p^k (1 - p)^{N-k} \end{aligned}$$

Now the question is: in  $M$ ,  $b/g$  and  $p$  the value of the given conditions, and  $N$ , for achieving  $E\eta/g$ .

This is an optimization problem, because the variables in the sum limit, and combination of expression, it is difficult to use analytic method, however, can be used by numerical calculation, comparing the size of the method, the optimal solution of it.

**Examples of mathematical contest in modeling 2.** Known results of 12 teams in the league, asked to design a ranking method based on these results to football team [4].

This problem can have a variety of different approaches, regression analysis is one of a kind of practice.

Let  $m = 12$  team's strength are  $\beta_1, \beta_2, \dots, \beta_m$ , these are unknown constants,

Let  $y_i$  is the  $i$  time when the first game, the score manifested by the home team and the visiting team's strength difference between the two teams. For example, when the two teams score is 3: 2, You can define  $y_i = 3 - 2$  or  $y_i = \sqrt{3} - \sqrt{2}$  or  $y_i = \sqrt[3]{3} - \sqrt[3]{2}$  or  $y_i = \ln\left(\frac{1+3}{1+2}\right)$  and so on.

Let the first game, is 1 teams 2 teams, 1 team-based team, two teams for the visiting team; the first two games, three teams to four teams, three teams based team, four teams for the visiting team; the first three games, 1 vs 4 team, a team-based team, four teams for the visiting team. Then,

$$\begin{aligned} y_1 &= \beta_1 - \beta_2 + \varepsilon_1, \\ y_2 &= \beta_3 - \beta_4 + \varepsilon_2, \\ y_3 &= \beta_1 - \beta_4 + \varepsilon_3, \\ &\dots \end{aligned}$$

For each game, there are

$$y_i = \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_m x_{im} + \varepsilon_i$$

among them,

$$x_{ij} = \begin{cases} +1 & \text{The first game } y_i, j\text{-team competition as the home team} \\ -1 & \text{The first game } y_i, j\text{-team competition as the visiting team} \\ 0 & \text{The first game } y_i, j\text{-team competition as the visiting team} \end{cases}$$

$\varepsilon_i \sim N(0, \sigma^2)$   $i$  was the first game the result of random error,  $i = 1, 2, \dots, n$

As can be seen, this is actually a constant item  $\beta_0$  does not appear linear regression, regression equation

$$y = \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_m x_m + \varepsilon$$

Requirements from observations,  $x_{i1}, x_{i2}, \dots, x_{im}$ , (+1, -1 Or 0) and  $y_i$  (match result) departure, then  $\beta_1, \beta_2, \dots, \beta_m$ . Estimates calculated strength of each team, it can give strength according to the size ranking of each team.

When the actual calculation, but also taking into account the results of the competition reflects only the difference between the strength of the team, only know the relative size of the relationship, the lack of an absolute benchmark, in order to determine the numerical strength of each team, in fact, impossible. To solve this problem is also very easy, as long as the value given in advance a strength of the team, as a benchmark it [7]. For example, you can make  $\beta_m = 0$ , which is equivalent to remove the last entry in the regression equation, and linear regression, you can find the estimated value of other  $\beta$ .

**Examples of mathematical contest in modeling 3.** Demographers Malthus, 200 years ago investigated within the British-born population statistics make assumptions about the birth rate is a fixed value, so that the establishment of an idealized model of population projections [2, 5].

Keep moment  $t$  population of  $x(t)$ , here to take advantage of calculus, assuming  $x(t)$  is one differentiable function, according to Malthus's hypothesis, in period  $t$  to  $t + \Delta t$ , the amount of growth

$$x(t + \Delta t) - x(t) = rx(t)\Delta t$$

With the increase of population decline, will represent a  $r$  function, you can put  $x$  as a decreasing function  $r(x)$ , it may make  $r(x)$  as

$$r(x) = r_0 - sx \quad (r, s > 0)$$

Here's  $r_0$  is the inherent growth rates, in theory, is when the growth rate of  $x=0$ , then

$$\begin{cases} \frac{dx}{dt} = r(x)x \\ x(0) = x_0 \end{cases}$$

To determine the significance of the coefficients  $s$ , where the introduction of natural resources and environmental conditions that can accommodate a maximum population number  $x_m$  is called population capacity. When the  $x_m$  and when population is no longer growing, that  $x = x_m$  into the

formula too, so  $r(x_m) = 0$ , the  $s = \frac{r}{x_m}$  into the equation, so  $r(x) = r(1 - \frac{x}{x_m})$  then,

$$r(x) \rightarrow \begin{cases} \frac{dx}{dt} = rx(1 - \frac{x}{x_m}) \\ x(0) = x_0 \end{cases}$$

Formula moderate right-hand side of the factor  $rx$  reflects population growth trends, factor  $(1 - \frac{x}{x_m})$

is a reaction environment and resources for population growth effect. You can see  $x$  smaller,  $rx$

smaller,  $(1 - \frac{x}{x_m})$  becomes large, so that population growth is no longer determined only by the time, but the combined effect of the two factors.

According to this model, we can predict that by 2020 China Population  $x(2020) \approx 1.4$  billion after the forecast result to be the actual situation to give answers shown as Fig.1.

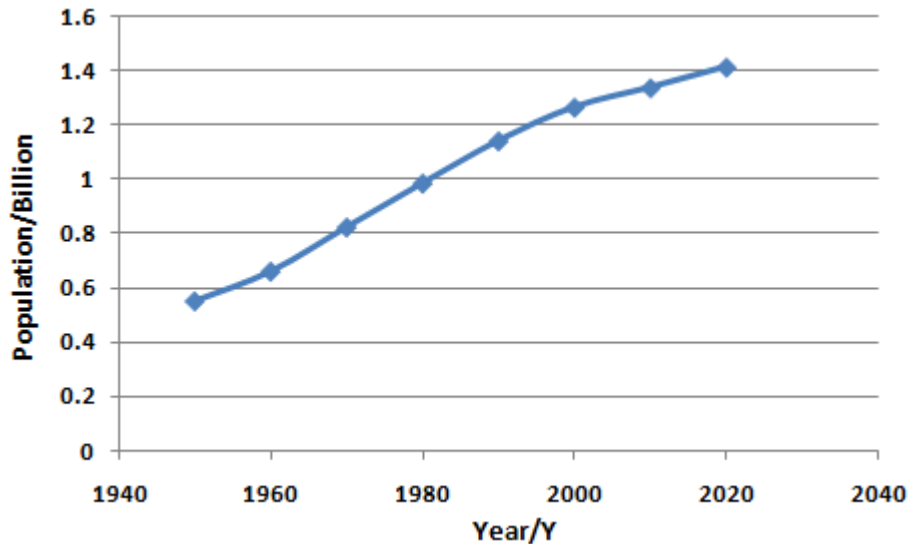


Fig.1 Chinese population forecast

## Conclusion

Data processing plays in mathematical modeling contest in an extremely important role, using accurate and reasonable method of data processing to complete seemingly complex practical problems, simplify, qualitative analysis quantitative analysis becomes more scientific and reliable. In the modeling process, should allow students to understand the importance of data processing, to master the correct method of data processing and mathematical software, mathematical modeling can greatly improve efficiency.

## References

- [1] G.J. Wang, Modern methods of mathematical modeling, Beijing: Science Press, 2006.pp.44-54.
- [2] Q.Y.Jiang and J.X. Xie, Mathematical model, Beijing: Higher Education Press, 2005, pp83-87.
- [3] Q.Ch. Yang, Mathematical Software and Mathematical Modeling, Hunan Institute of Humanities Science and Technology 2006,pp.38-42.
- [4] X.S.Yuang, Lingo and Excel in Mathematical Modeling, Beijing: Science Press 2007, pp. 129-134.
- [5] J. Chen, Application of ordinary differential equations in mathematical modeling, Jilin Institute of Education report, 2011.pp.55-59.
- [6] Sh.L. Yu and D.H. Zhao, Mathematical Software and Mathematical Modeling, College Mathematics, 2012.pp.324-332.
- [7] Y.J. Liang, Sh.T. Shi, Application of ordinary differential equations in mathematical modeling, Beijing: Science Press, 2013.pp.111-115.