Logistic-Based Population Projection Model

Zhenjing Di[1] and Hui Xu[1*]
[1]Department of Mathematics, Faculty of Science, Yanbian University, Yanji, China.
*Corresponding author email id: 361471648@qq.com

Date of publication (dd/mm/yyyy): 24/12/2021

Abstract **– The population issue has always been a core issue of social and political stability and economic development, and is a key constraint on socio-economic development and the enhancement of comprehensive national power. A reasonable projection of the population of a region is conducive to the correct understanding of social development trends, and to the publication and implementation of national policies. Therefore, the analysis and optimisation of existing population forecasting models and the improvement of the accuracy of forecasting models are crucial to accurately grasp the population size and structure of each region, to promote population growth in line with the development of social productivity and to optimise the rational allocation of social and economic resources. This paper uses traditional logistic models as the basis, and uses non-linear least squares estimation and linear least squares estimation to estimate and optimise the model parameters. By comparing the predicted and actual values of the three models from 2009 to 2018, the optimal model is derived and applied to the population forecast of a region in Jilin Province, providing a scientific basis for coordinating the sustainable development of the region's population and resources.**

Keywords **– Logistic Models, Non-Linear Least Squares, Linear Least Squares, Population Projections, Sustainable Development.**

I. INTRODUCTION

Population is one of the most important concerns in the world. Scientific understanding of the laws of change in population numbers, establishing population models and making accurate forecasts are the prerequisites for effective control of population growth. Since the founding of New China, many experts and scholars have investigated and researched the population problem. 1957, Ma Yinchu published “A New Theory of Population [1]”, which discussed the population problem and proposed that population growth needed to be compatible with the national economy and that too rapid population growth would exert great pressure on the economy and society. In 1982, Cui Qiwu et al[2] argued from the adsorption theory of chemical kinetics that population growth and resource limitation do not satisfy a linear relationship in nature, and proposed a generalized model of population growth in the article “A new mathematical model of population growth-an expansion of the classical logistic and exponential equations In 1990, Wang Shousong constructed a dynamic model of population growth, the generalized logistic model, by establishing a linear relationship between the relative growth rate of the population, the population size and the growth rate in A generalized logistic model of single population growth[3]. In 1997, Xu W. et al[4] used In 2006, Wang Yanchen et al. expressed the population growth rate as an exponentially decreasing function in “Population forecasting and improvement of LOGISTIC model[5]”, and obtained an improved model with better fitting effect than the traditional logistic model. In 2015, Li Yuan et al[6] improved the method of determining the parameters in the Logistic model based on the least squares method, and determined the parameters through curve fitting, providing ideas for the improvement of the traditional Logistic model.

In 1950, five years after the establishment of the United Nations, the total world population was about 2.6 billion. The world population grew to 5 billion in 1987 and reached 6 billion in 1999. The global population

Copyright © 2021 IJISM, All right reserved

passed the 7 billion mark in 2011[7]. The growing population provides sufficient labour for productive life and to a certain extent solves the productivity problem. However, the growing population has a significant impact on the stability of ecosystems and the development of human society. Therefore, it is important to establish prediction models and grasp population growth patterns to promote the coordinated and sustainable development of population and economic, social, resource and environment, and to help the implementation of relevant national population policies.

II. Overview of Traditional Logistic Models

At present, there are many models for studying population aggregates, mainly Malthusian models, grey system GM(1,1) models, logistic models, linear regression models, Leslie models, etc. Among them, one of the more commonly used models is the logistic model. One of the more commonly used models is the logistic model.

The traditional logistic model was first proposed by the Dutch bio-mathematician Verhulst in 1838[8]. This model compensates to some extent for the shortcomings of the Malthusian model, taking into account factors such as population constraints, and is one of the most important methods for studying population growth patterns in a finite space.

Assume that the population growth rate r(X) is a linear function of x, $r(x) = r - sx$. It is assumed that the maximum number of people that can be accommodated by natural resources and environmental conditions is x_m, i.e. the growth rate is $r(x_m) = 0$ when $x = x_m$.

By hypothesis, there are

$$r(x) = r(1 - \frac{x}{x_m}) \tag{1}$$

There will be

$$\begin{cases} \frac{dx}{dt} = r(1 - \frac{x}{x_m})x \\ x(t_0) = x_0 \end{cases} \tag{2}$$

The solution is

$$x(t) = \frac{x_m}{1 + (\frac{x_m}{x_0} - 1)e^{-r(t - t_0)}} \tag{3}$$

where x_0 is the population size at the initial moment, t_0 is the initial moment and t is the termination moment. Logistic models reflect the balance between ecosystems, overcoming the tendency of other models to predict infinite population growth and providing a better description of the pattern of population or other biological population growth when the population grows to a certain number, which is more in line with the actual situation [9]. However, traditional logistic models regard the relationship between population growth rate $r(x)$ and total population x as a linear relationship in its simplest form, and the resulting predictions are still subject to a certain amount of error from the actual values. To this end, this paper improves the logistic model according

Copyright © 2021 IJISM, All right reserved

to existing estimation methods and obtains two optimised estimation methods for parameters x_m and *r*, thus establishing a more accurate population projection model and providing a basis for the implementation and improvement of population policy in Jilin.

III. Two Optimisations of the Logistic Model

A. *Non-Linear Least Squares Estimation*

Parameter estimation using nonlinear least squares, simplifying the original problem and then solving the objective function optimally so that the value of the function is minimized:

$$x^* = \arg\min_x \{F^2(x)\} \tag{4}$$

$$F^2(x) = \frac{1}{2}\sum_{i=1}^{m}(f_i^2(x)) = \frac{1}{2}\| f(x)\|^2 = \frac{1}{2} f(x)^T f(x) \tag{5}$$

where function $F^2(x)$ is the objective function to be solved, $f(x)$ consists of the difference between the model function and the measured value and x^* is the local minimum of the objective function.

According to the data, the Gaussian Newton method is a simple algorithm for solving non-linear least squares problems [10] , the basic idea of which is to perform a first-order Taylor expansion of the non-linear function $F(x)$ [11]:

$$F(x+\Delta x) = F(x) + J(x)\Delta x \tag{6}$$

where is the Jacobi matrix of the $J(x)$ function, i.e. the first derivative of the F function.

In order to solve for the appropriate value of the variable Δx to minimise the value of the function $\| F(x+\Delta x)\|^2$, a least squares problem can be constructed which converts the problem of minimising the function $F(x)$ from x into a problem of minimizing the function $F(x+\Delta x)$ from Δx, while the objective function becomes $\frac{1}{2}\| F(x) + J(x)\Delta x\|^2$:

$$\Delta x^* = \arg\min_{\Delta x} \frac{1}{2} || F(x+\Delta x)||^2 = \arg\min_{\Delta x} \frac{1}{2} || F(x) + J(x)\Delta x||^2$$

Expanding the objective function:

$$\frac{1}{2}\| F(x) + J(x)\Delta x\|^2 = \frac{1}{2}(F(x) + J(x)\Delta x)^T (F(x) + J(x)\Delta x)$$
$$= \frac{1}{2}(\| F(x)\|_2^2 + 2F(x)^T J(x)\Delta x + \Delta x^T J(x)^T J(x)\Delta x)$$

In order to find the extreme value of this objective function, equation (8) has to be derived and made equal to zero.

$$2F(x)^T J(x) + 2J(x)^T J(x)\Delta x = 0 \tag{7}$$

$$J(x)^T J(x)\Delta x = -J(x)^T F(x) \tag{8}$$

From equation (8), a linear equation for Δx can be obtained, which is called the Gaussian Newton equation.

Copyright © 2021 IJISM, All right reserved

For a specific algorithmic implementation of the above process, Matlab software can be programmed to enable non-linear least squares estimation. The demographic data collected is taken as the initial condition and the parameters x_m and r from equation (3) are fitted with the remainder of the data.

A. *Linear Least Squares Estimation*

Firstly, the Logistic Equation is expressed as,

$$\frac{1}{x}\cdot\frac{dx}{dt}=r-sx \quad s=\frac{r}{x_m}$$

Difference equations can be obtained using the difference method,

$$\frac{x(k)-x(k-1)}{\Delta t}\cdot\frac{1}{x(k)}=r-sx(k), k=2,3,\cdots,22$$

where the step $\Delta t=1000$ is set and the parameters r and s are fitted.

IV. Comparison of the Two Optimisation Models

The two models were analysed to predict the population of a region in Jilin Province from 2009 to 2018, and the predicted values were compared with the actual values, and the fit of each model was derived from indicators such as error.

Table 1. Analysis of population projection data for model 1, 2009-2018.

Year	Actual Value / Millions	Predicted Value / Millions	Error	Squared Error	Absolute Error	Error Rate
2009	7.565065	7.422054	0.143011	0.020452	0.143011	0.018904
2010	7.588921	7.455295	0.133626	0.017856	0.133626	0.017608
2011	7.617663	7.486772	0.130891	0.017132	0.130891	0.017183
2012	7.569037	7.516564	0.052473	0.002753	0.052473	0.006933
2013	7.526708	7.544747	-0.018039	0.000325	0.018039	0.002397
2014	7.545472	7.571397	-0.025925	0.000672	0.025925	0.003436
2015	7.538335	7.596587	-0.058252	0.003393	0.058252	0.007727
2016	7.534284	7.620386	-0.086102	0.007414	0.086102	0.011428
2017	7.489211	7.642864	-0.153653	0.023609	0.153653	0.020517
2018	7.512896	7.664086	-0.151190	0.022859	0.151190	0.020124

Use the data in the table to find the relative standard deviation S_1 and the mean absolute percentage error ($MAPE_1$) to examine the effect of the model predictions:

$$S_1=\sqrt{\frac{1}{10}\sum_{t=1}^{10}e_t^2}=0.107919$$

$$MAPE_1=\frac{1}{10}\sum_{t=1}^{10}\frac{|e_t|}{x}=0.012626$$

where, denotes the error e_t.

Copyright © 2021 IJISM, All right reserved

Table 2. Analysis of population projection data for model 2 for the period 2009-2018.

Year	Actual Value / Millions	Predicted Value / Millions	Error	Squared Error	Absolute Error	Error Rate
2009	7.565065	7.422054	0.143011	0.020452	0.143011	0.018904
2010	7.588921	7.455295	0.133626	0.017856	0.133626	0.017608
2011	7.617663	7.486771	0.130892	0.017133	0.130892	0.017183
2012	7.569037	7.516563	0.052474	0.002754	0.052474	0.006933
2013	7.526708	7.544746	-0.018038	0.000325	0.018038	0.002397
2014	7.545472	7.571396	-0.025924	0.000672	0.025924	0.003436
2015	7.538335	7.596585	-0.058250	0.003393	0.058250	0.007727
2016	7.534284	7.620385	-0.086101	0.007413	0.086101	0.011428
2017	7.489211	7.642862	-0.153651	0.023609	0.153651	0.020516
2018	7.512896	7.664084	-0.151188	0.022858	0.151188	0.020124

Use the data in the table to find the relative standard deviation S_2 and the mean absolute percentage error ($MAPE_2$) to examine the effect of the model predictions:

$$S_2 = \sqrt{\frac{1}{10}\sum_{t=1}^{10} e_t^2} = 0.107915$$

$$MAPE_2 = \frac{1}{10}\sum_{t=1}^{10}\frac{|e_t|}{x} = 0.012625$$

where, denotes the error e_t .

A comparative analysis of $S_i(i=1,2)$ and $MAPE_i(i=1,2)$ shows that Model 2 is a slightly better fit and that the model corresponds well to actual population growth.

IV. Population Projections

A collection of statistical yearbooks from Jilin Province yielded 34 years of data for a region of Jilin Province as shown in Table 3.

Table 3. Table of 34 year population statistics for a district in Jilin Province.

Year	Population/ Million	Year	Population/Million
1985	5.895000	2002	7.125055
1986	5.951000	2003	7.182348
1987	6.027000	2004	7.240845
1988	6.119000	2005	7.314959
1989	6.265822	2006	7.392561
1990	6.378237	2007	7.459463
1991	6.426346	2008	7.525303

Copyright © 2021 IJISM, All right reserved

Year	Population/ Million	Year	Population/Million
1992	6.457351	2009	7.565065
1993	6.510368	2010	7.588921
1994	6.574999	2011	7.617663
1995	6.672912	2012	7.569037
1996	6.767781	2013	7.526708
1997	6.837875	2014	7.545472
1998	6.868673	2015	7.538335
1999	6.912278	2016	7.534284
2000	6.996354	2017	7.489211
2001	7.057321	2018	7.512896

The improved logistic model described above was used to forecast the population trends in a region of Jilin for the next thirty years. The logistic model with linear least squares estimated parameters and was used for the projections to obtain a maximum population accommodation of $x_m = 7.8413$ million, and a population growth rate of $r = 0.0066$, to obtain the projected population data as shown in Table 4.

Table 4. Population projection data table.

Year	Population/ Million	Year	Population/ Million
2019	7.684109	2035	7.885047
2020	7.703007	2036	7.892274
2021	7.720832	2037	7.899070
2022	7.737640	2038	7.905460
2023	7.753486	2039	7.911468
2024	7.768420	2040	7.917117
2025	7.782491	2041	7.922427
2026	7.795747	2042	7.927418
2027	7.808233	2043	7.932109
2028	7.819990	2044	7.936517
2029	7.831059	2045	7.940661
2030	7.841478	2046	7.944554
2031	7.851285	2047	7.948212
2032	7.860513	2048	7.951650
2033	7.869196	2049	7.954879
2034	7.877364	2050	7.957913

Based on the above projections, again using Matlab software, the population growth trends predicted by the

Copyright © 2021 IJISM, All right reserved

model were programmed and plotted as shown in Figure 1.

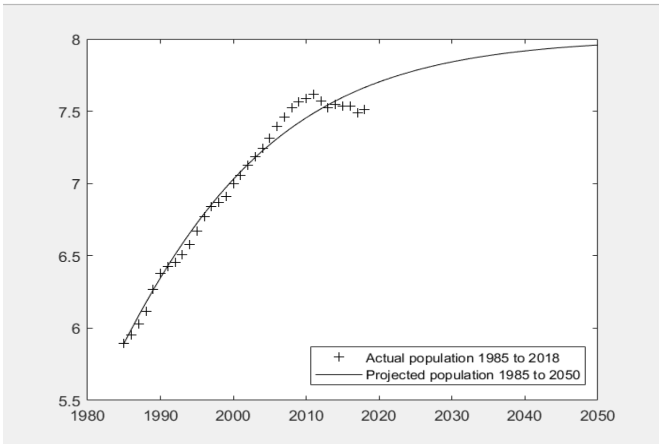

Fig. 1. Graph of projected population growth trends.

V. CONCLUSIONS

In this paper, on the basis of the traditional logistic model, the estimation methods of the parameters and in the model are further improved and refined. The three optimised logistic models were obtained by using the non-linear least squares estimation method and the linear least squares estimation method for parameter estimation, respectively, and the three models were analysed and compared to examine the accuracy of the models by deriving the relative standard deviation S_i (i = 1,2,3) and the mean absolute percentage error ($MAPE_i$ (i = 1, 2, 3)). Finally, the optimal prediction model is applied to the prediction of the population of a region in Jilin Province, and the population size of the region in the next 30 years is derived through Matlab programming plots, providing a direction for the relevant departments to plan the development of the region and helping to build China into a strong, democratic, civilised, harmonious and beautiful modern socialist country by 2050.

Population is one of the most important issues in the development of society today. China's population problem is no longer limited to the question of how many people there are, but is moving in a diversified direction with the continuous development of society and economy [14]. In this regard, accurately predicting the future demographic development trend of each region in China is one of the important issues in adhering to the concept of sustainable development and mastering the structure of labour force distribution. The existence of a large surplus labour force in China will be a long-term phenomenon, and there is a long way to go to achieve the goal of a modern and powerful country by 2050, as enterprises can gain a foothold in the world competition and control the unemployment rate [15]. The optimised forecasting model provides new ideas for regional population forecasting, and plays a non-negligible role in stabilising the social structure, promoting a rational distribution of labour, and providing a side-by-side understanding of regional population flows as well as the issuance and implementation of various national policies, reflecting the sustainable development of population, resources and society, and working to adjust China's population structure to a more rational development direction.

REFERENCES

[1] LI Yaxin. 2019. Reflections on current population issues and future trends in China. Regional governance 42:220-222.
[2] Cui Qiwu. 1982. A new mathematical model of population growth - an extension of the classical logistic and exponential equations. Journal of Ecology 4: :403-415.
[3] Wang Shousong.1990. Generalized logistic model for the growth of single populations. Journal of Biomathematics 1:21-25.
[4] Xu Wenke,Zhang Shisheng.1997. Two-way differential fitting of logistic models of biological populations and improvement of initial prediction values. Agricultural Systems Science and Synthesis 1:1-3+9.

Copyright © 2021 IJISM, All right reserved

[5] Wang Yanchen, Duan Junsheng, Wang Yan. 2006. Population forecast and improvement of LOGISTIC model. Statistics and Decision Making 22:136-137.
[6] Li Yuan,Yang Chen-Chen, Wang Xuefeng, Shen Shiyun.2015. Improved logistic population model based on least squares. Technology Vision,31:7-9.
[7] https://www.un.org/zh/sections/issues-depth/population/index.html
[8] Wu Xinyuan.1990. A numerical method for fitting logistic Stimulus curves.Journal of Biomathematics 1:26-32.
[9] Yan Huizhen.2008.Application of logistic model in population prediction.Journal of Dalian University of Technology 4:333-335.
[10] Håkan Ramsin, Per-Åke Wedin. A comparison of some algorithms for the nonlinear least squares problem. 1977, 17(1):72-90.
[11] K. Levenberg,A method for the solution of certain nonlinear problems in least squares, Quart. Appl. Math. 2 (1944) 164–168.
[12] Jilin Statistical Yearbook, China Statistics Press, 1985-2019.
[13] Jiang Qiyuan, Xie Jinxing, Ye Jun. 2018.Mathematical modeling. Higher Education Press - Beijing, 2018,5.
[14] Chen Ruyong.2001. Review of research on moderate population in China. Northwest Population 1:12-15.
[15] Jiang Zhenghua, Zhang Yingguang.2003. Population research and population work in the New Century and New Stage. Chinese Population Science 1:3-10.

AUTHOR'S PROFILE

First Author
Zhenjing Di, Study in the department of mathematics of Yanbian University (Yanji, China). The thesis is a project of Jilin Province. The paper is the result of the project "Algorithmic model of ordinary differential equations" (202010184009) of the Innovation and Entrepreneurship Training Program for University Students.

Second Author
Hui Xu, His major is mathematics. He is now an associate professor in the Department of Mathematics, School of Science, Yanbian University, and he is a master's tutor. His main research direction is mathematical education technology, mathematical modeling and intelligent algorithm.

Copyright © 2021 IJISM, All right reserved