Prediction Of Population Size and Age Structure in China Based on Grey Prodiction and Leslie Model

Linhui Wang [1, a], Yutong Huang [1, b], Zhihong Liu [2, *]

[1] Jinan University855 East Xingye Road, Panyu District Guangzhou, Guangdong, 510000, China

[2] South China Agricultural University Wushan Road 483, Tianhe District Guangzhou, Guangdong, 510000, China

* Corresponding Author Email: liuzhihong1127@163.com, [a] wanglinhui0017@163.com,

[b] huangyutong1127@163.com

Abstract. Population is a major problem that restricts China's comprehensive, coordinated and sustainable development. Therefore, it is of great significance to establish a model that can accurately predict the population and age structure. This paper uses GM (1,1) to obtain the gender ratio of future newborn population and total population, quantifies the three-child policy as the increase in the fertility rate of women of childbearing age, uses the Leslie model to obtain the female population at all ages, and finally gets the national population and age structure in the next five years under the background of the three-child policy. The results show that the opening of the three-child policy positively impacts the increase of the newborn population and the growth rate of the total population in the early future. But as time goes by, it still cannot restrain the decrease in the newborn population and the net increase of the total population. In addition, the aging problem has not been effectively improved in the short term. In the next five years, the proportion of the aged population continues to rise, while the young population gradually decreases. In terms of the labor force coefficient, there will be a downward trend from 2021 to 2025 and a slight increase from 2025 to 2026.

Keywords: Three-child policy; GM (1,1); Leslie model.

1. Introduction

With the development of the times, the law of natural population development has significantly changed. Accurate population prediction is crucial to understanding China's social and economic status and the grasp of future development trends [1].

From "family planning" to "double two-child" to "single two-child" and then to "universal two-child", the birthright is gradually returning to the family. In recent years, the newborn population has decreased significantly. Based on national conditions, on May 31, 2021, China implemented the policy that a couple can have three children and supporting measures [2], known as the "three-child policy". Studying the fertility situation and the future trend of population policy under the Chinese three-child policy is a hot topic of general concern from all walks of life [3].

Foreign scholars have a long research history in future population prediction and innovate on the basis of previous research. In 1798, British scholars established the Malthus modelbased on the ideal population prediction model, depending on the geometric series of population growth, without resource restriction. In 1838, the Dutch scholar Verhulst introduced the competition item on the original formula of the Malthus model and established the Logistic model, which solved the problem that the population was fixed and made up for the Malthus model defects. However, the model cannot solve the influence of non-linear factors such as fertility policy and medical level on population growth in human society. In 1945, Australian scholar Leslie used the method of studying animal population change for population research and proposed the Leslie method, a discrete mathematical random model to predict the trend of population environment development.

Domestic scholars' research on the age structure of population and trend is becoming increasingly mature. *Double – exponent Model of Population Evolution and China's Population Prediction* adopts the double exponential decay model to predict the population and age structure by predicting the

natural growth rate of the population. The results indicate that China's population shows a growth trend, reaching 1.45 billion, 1.478 billion, 1.503 billion and 1.524 billion, respectively, in 2025, 2030, 2035 and 2040 [4]. *Stochastic Model for Population Forecast: Based on Leslie Matrix and ARMA Model* uses Leslie matrix and ARMA model to explore the changes of China's population and predicts the number of China's population at all ages until 2050 [5]. *The Dynamic Model of Grey System and Its Application to Population Forecasting* is based on the dynamic grey model to predict China's population. The prediction results show that China's population will not exceed 1.45 billion in 2025, and the final population will not exceed 1.5 billion [6].

In conclusion, foreign scholars have put forward many classical methods and models with reference significance. The research of domestic scholars on population prediction mainly focusses on the total population, but there is relatively fewer research in the field of age structure, and scholars have different views on the peak of the total population, which leaves space for our research. Based on the "three-child policy" environment and the research of previous scholars, this essay aims to use the appropriate model to predict the changes in total population and the age structure brought about by policy, so as to contribute to the population prediction theory of our country.

2. Population change in China based on the Leslie model

2.1. Model overview

Leslie model is a discrete matrix model based on age and gender. The Leslie matrix is constructed from the mortality rate and fertility rate of female age groups by using the female population of a given country or region as a column vector. Multiply the former by the latter to get a column vector of the female population of all ages for the next year. Finally, the size of the total population is calculated by the gender ratio.

2.2. Data collection and preprocessing

The prediction of the age structure of population in this paper takes the annual national age structure of population as the research object, and the experimental data are mainly from *China Population and Employment Statistical Yearbook* [7]. This paper collects four data on the number of females in each age group from 0 to 95, fertility rate, mortality rate and the proportion of female newborns in the newborn population in 2019.

2.2.1 Processing of missing mortality values

Due to the sampling survey, the mortality rates of females in the small number of children and labor force age groups are missing. Since the difference in male and female mortality rates is mainly in aging age groups [7], the mortality rates of males in the corresponding age groups are used to fill the missing values.

2.2.2 Survival rate calculation

The survival rate formula for all age groups is as follows:

$$s_i(t) = 1 - d_i(t)$$

Where $s_i(t)$ is the survival rate of female in group i in year t, $d_i(t)$ is the mortality rate of female in group i in year t.

2.2.3 Fertility rate calculation

(1) Fertility rate under the two-child policy

The relevant data are checked through the *National Statistical Yearbook* and *China Population and Employment Statistical Yearbook*. China implemented the "double two-child policy" in November 2011, and introduced the "single two-child policy" in December 2013. In October 2015, China implemented the "universal two-child policy". Therefore, the increased fertility rates in 2012,

2014 and 2016 compared with the previous year are calculated and summed up as the impact of the two-child policy.

In this paper, the fertility rate is defined as the ratio of the newborn population in that year to the number of females of childbearing age. Let the newborn population in year t is A_t and the number of females of childbearing ages in year t is Y_t, then the fertility rate b_t in this year is:

$$b_t = \frac{A_t}{Y_t}$$

The fertility rate of each year is as follow:

Tab. 1 Fertility rate (2011-2016)

Year	*Fertility rate (%)*
2011	2.975
2012	3.589
2013	3.568
2014	3.733
2015	3.093
2016	3.617

According to the figures in the table, the fertility rate increased by the two-child policy is 1.303%.

(2) Fertility intention under the two-child and three-child policy

The relevant data are derived from the survey report. According to *the investigation reports The Impact of implementing the Universal Two-Child Policy on Family Education*, 53.3 % of one-child families have no intention of having a second child. The survey results of the Jinan investigation team of the National Bureau of Statistics show that 85% of respondents do not intend to have a third child, 6.3% have the desire to have a third child, and 8.7% are not sure at present [8].

(3) Increased fertility rate in 2021

Assume the ratio of the surveyed fertility intention under the policy to the actual increased fertility rate is fixed. Then the influence brought by the opening of the three-child policy is compared with the fertility intention to have a second child under two-child policy and the eventual increased fertility rate, which is calculated according to the ratio formula:

$$\frac{x_1}{x_2} = \frac{z_1}{z_2}$$

Where x_1 is the fertility intention to have a second child in 2015, x_2 is the fertility intention to have a third child in 2021, z_1 is the increased fertility rate by the two-child policy , and z_2 is the increased fertility rate by the three-child policy.

Based on the formula and the survey data, the increased fertility rate in 2021 is 0.176%.

According to the above data, the introduction of a fertility policy can improve the overall fertility rate in the short term. Fertility willingness is affected by the economy, health and other factors, which further affects how the three-child policy can improve fertility. Therefore, we assume that implementing the three-child policy will increase the fertility rate in the short term. However, its effect will weaken over time, and the corresponding time function is set as follows:

$$y(t) = 1.76 - Kt \ \ (t = 1, 2, 3...)$$

Where K is the impact factor, t is the year after the three-child policy was implemented, and $y(t)$ is the increased fertility rate in this year. Considering the changes in the second-child fertility rate, the economic input and the fertility intention brought by raising the third child, we take $K = 0.1$. The increased fertility rates are as follows:

Tab. 2 Revised fertility rates for childbearing age groups in 2021 according to policy impact (Part)

Age	*2022*	*2023*	*2024*	*2025*	*2026*
15	1.126	1.123	1.120	1.118	1.115
16	3.695	3.687	3.678	3.670	3.662
17	9.907	9.884	9.862	9.840	9.818
...	...	...	...	...	...
47	0.414	0.413	0.412	0.411	0.410
48	0.265	0.265	0.264	0.264	0.263
49	0.149	0.148	0.148	0.148	0.147

2.3. Model building

2.3.1 Model hypotheses

(1) China is a closed system that does not take into account population migration.

(2) Within the predicted time, there will be no adverse factors such as wars, large-scale infectious diseases, natural disasters, etc.

(3) Females over 95 years old were included in the same group.

2.3.2 Notations

Tab. 3 Notations

Symbol	*Definition*
$x_i(t)$	Female population in age group i in year t, range of [0,95]
$b_i(t)$	The female fertility rate in the age group i from 15 to 49 in year t
$d_i(t)$	The female mortality rate in age group i in year t
$s_i(t)$	The female survival rate in age group i in year t
$w(t)$	The female births as a percentage of births in year t

Female population in age group 0 in year t is:

$$x_0(t+1)=\sum_{i=15}^{49} b_i w_i x_i(t)$$

Female population in age group $i+1$ in year $t+1$ is the number of survived in age group i.

$x_{i+1}(t+1)=s_i(t)x_i(t)$

2.3.3 Leslie matrix

$$L=\begin{pmatrix} wb_1 & wb_2 & ... & wb_m \\ s_1 & 0 & ... & 0 \\ 0 & s_2 & 0 & ... \\ 0 & ... & s_{m-1} & 0 \end{pmatrix} \quad m=95$$

Then we have $x(t+1)=Lx(t)$.

The total female population in year t is obtained by summing up the female population of each age group in year t, and the total population in year t is obtained by the gender ratio.

2.4. Model results

The Leslie model is used to calculate the female population of all ages in China in the next five years. The model results are shown in the following table:

Tab. 4 Chinese Female Population by Age Group in the next 5 Years (Part)

Age	*2022*	*2023*	*2024*	*2025*	*2026*
0	6632274	6393999	6145997	5913969	5715891
20	6570698	6147908	6842495	6697825	6972645
40	11037957	9637373	9773134	9907596	11578297
60	10520798	12712906	11157627	11270639	11392457
80	2743559	2773958	2932224	3108280	3475014

3. Analysis of gender ratio of newborn population and total population in China based on GM (1,1)

3.1. Model overview

In order to predict the gender ratio of newborn population and total population in China, the grey prediction model is adopted in this paper. One of the characteristics of the grey prediction method is that it requires less information and can transform the disordered discrete original sequence into an ordered sequence with high accuracy. This paper uses the model to predict the gender ratio of newborn population and the gender ratio of China's total population in the next five years based on the data from the opening of the second child policy to the present nine years.

3.2. Data preprocessing

The data of gender ratio of newborn and total population are mainly from *China Population and Employment Statistical Yearbook*.

Tab. 5 Gender ratio of newborn and total population in China (2011-2019)

Year	*Gender ratio of newborn population*	*Gender ratio of total population*
2011	117.78	105.17
2012	117.70	105.12
2013	117.60	105.22
2014	116.88	105.04
2015	113.51	105.02
2016	112.88	104.98
2017	111.90	104.81
2018	\	104.64
2019	111.30	104.46
2020	111.30	105.07

Since the data of gender ratio of newborn population released by the National Bureau of Statistics was missing in 2018, this paper uses the data interpolation method to fill the missing value. After calculation, the gender ratio of newborn population in 2018 is 111.7.

3.3. Model building

The original sequence of the gender ratio of newborn population to be predicted is:

$$y^{(0)}=(y^{(0)}(1),\ y^{(0)}(2),\ \ldots,\ y^{(0)}(10))$$

3.3.1 Data inspection

Find the grade ratio of this sequence:

$$\lambda(k)=\frac{y^{(0)}(k-1)}{y^{(0)}(k)}\quad k=1,\ 2,\ \ldots,\ 10$$

According to the test, the grade ratio of the gender ratio of newborn population in China falls in the tolerable coverage space$\left(e^{-\frac{2}{11}},\ e^{\frac{2}{11}}\right)$. Then, the sequence can be predicted by GM (1,1).

3.3.2 GM (1,1)

To accumulate each term of the original sequence:

$$y^{(1)}(n)=y^{(0)}(1)+...+y^{(0)}(n)=\sum_{i=1}^{n}y^{(0)}(i)$$

Compared with the original sequence, the stability of the accumulated sequence is increased, and the randomness is reduced.

The resulting accumulated sequence is:

$$y^{(1)}=(y^{(1)}(1),\ y^{(1)}(2),\ ...\ ,\ y^{(1)}(10))$$

The trend of the new series is expressed as:

$$\frac{dy^{(1)}}{dt}+ay^{(t)}=u$$

where a is the development coefficient and u is the grey action quantity.

$$\begin{bmatrix}a\\u\end{bmatrix}=(B^TB)^{-1}B^TY$$

$$Y=(y^{(0)}(2),\ y^{(0)}(3),\ ...\ ,\ y^{(0)}(10))^T$$

$$B=\begin{bmatrix}-\frac{1}{2}[y^{(1)}(1)+y^{(1)}(2)] & 1\\ \cdots & \cdots\\ -\frac{1}{2}[y^{(1)}(9)+y^{(1)}(10)] & 1\end{bmatrix}$$

The grey prediction model and corresponding predicted value are obtained:

$$\hat{y}^{(1)}(t+1)=\left[y^{(0)}(1)-\frac{u}{a}\right]e^{-at}+\frac{u}{a}$$

$$\hat{y}^{(0)}(t+1)=\hat{x}^{(1)}(t+1)-\hat{y}^{(1)}(t)$$

$\hat{y}^{(0)}(t)$ is the sequence of predicted gender ratio of newborn population each year, which gives the value from 2021 to 2026. Using the same model, the predicted gender ratio of total population is obtained.

3.4. Model Results

The prediction results are as follows:

Tab. 6 Gender ratio of newborn and total population in China (2021-2026)

Year	*Gender ratio of newborn population*	*Gender ratio of total population*
2021	109.2068	104.6382
2022	108.3037	104.5802
2023	107.4081	104.5222
2024	106.5199	104.4643
2025	105.6391	104.4063
2026	104.7655	104.4384

4. Accuracy test of model-predicted results

In order to test the accuracy of the model, the model in this paper is used to predict the total population after the "universal two-child policy" was implemented in 2015, and then compared with the actual population from *National Statistical Yearbook*. The experimental data in 2015 are from *China Population and Employment Statistical Yearbook.*

The increased fertility rate in 2016 is 0.521%, which is computed based on the data in 2011 and 2012 when the "double two child policy" was implemented. The predicted value is very close to the actual value of 0.524%.

Use GM (1,1) to obtain the gender ratio of future newborn population and total population from 2016 to 2019:

Tab. 7 Predicted gender ratio of total population and its relative error in China (2016-2019)

Year	***2016***	***2017***	***2018***	***2019***
Predicted	104.98	104.93	104.88	104.84
Real	104.46	104.64	104.81	104.98
Relative error	0.50%	0.28%	0.07%	0.13%

It can be seen from the above table that the relative errors between the prediction results and the actual results are not greater than 1%, implying that the gray prediction fitting degree is good.

Use the Leslie model to predict the total population from 2016 to 2019, and the corrected predicted results are:

Tab. 8 Predicted total population (ten thousand) and its relative error in China (2016-2019)

Year	***2016***	***2017***	***2018***	***2019***	***2020***
Predicted	137668	138416	139084	139688	140201
Real	137462	138271	139008	139538	140005
Relative error	0.15%	0.10%	0.05%	0.11%	0.14%

The relative errors are not greater than 1%. In conclusion, the prediction results of the model are reliable, and the model in this paper can be used to predict the population change after the implementation of the "three-child policy", which has certain practical significance.

5. China's future population forecast under the three-child policy

5.1. Prediction results and analysis of newborn population

Leslie Model and MATLAB are used to process the original data of *China's Population and Employment Statistical Yearbook in 2019*, and predict the female population of all ages in 2021. Based on the three-child policy, the fertility rate data in 2021 are revised. Then we use the Leslie model to continuously obtain the female population of all ages in 2022-2026 after implementing the three-child policy in 2021. Finally, the trends of China's newborn population, total population, and age structure in the next five years are summarized.

As the newborn population in 2021 is still unknown, the newborn population in 2021 is predicted through the combination of Leslie model and gray prediction model. The following is the forecast of newborn population from 2021 to 2016:

Let the number of newborn females in year t be $x^{(0)}(t)$, and the predicted gender ratio of newborn population in year t be $w(t)$, then the newborn population $A(t)$ in year t is:

$A(t) = x_0(t) \cdot (1 + w(t)\%)$

Therefore, the predicted newborn population in China is:

Tab. 9 Newborn population in China (2021-2026)

Year	*Predicted newborn population*
2021	13546436
2022	13815272
2023	13261672
2024	12692707
2025	12161429
2026	11704171

Compared with 2021, the number of newborns in 2022 has significantly increased. However, as time goes by, the newborn population in the next four years still shows a gradual decline.

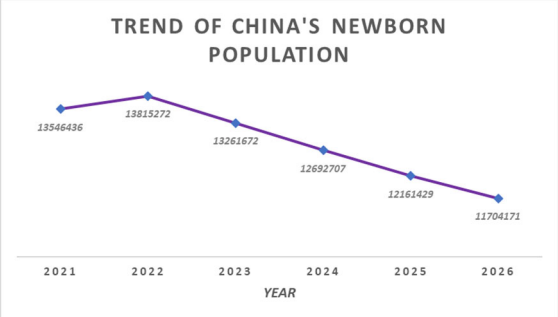

Fig. 1 Trend of China's newborn population from 2021 to 2026.

5.2. Prediction results and analysis of total population

The total population in previous years is estimated by means of a sample survey, so there is a margin of error. This paper revises the projected total population based on data from the seventh national census in 2020. The following is the prediction of China's total population in 2021 and the next five years:

Let the gender ratio of the predicted total population in year t be $n(t)$. The total population of the seventh census in 2020 is P_{2020}, then the total population of China $B'(t)$ in year t is:

$$x(t) = \sum_{i=0}^{95} x_i(t)$$

$$B(t) = x(t) \cdot (1 + n(t)\%)$$

$$B'(t) = B(t) \cdot \frac{P_{2020}}{B(2020)}$$

Therefore, the predicted total population in China is

Tab. 10 Predicted total population in China (2021-2026)

Year	*Predicted total population*
2021	1413791700
2022	1418195649
2023	1421795536
2024	1424414309
2025	1426131425
2026	1427021597

After the full implementation of the three-child policy in 2022, China's net population growth increased slightly, but not significantly. Compared with 2020 and 2021, the China's net population growth in 2022 is largely increased due to the increase of the newborn population.

Similarly, since the newborn population gradually decreases after that, the growth trend of China's total population will gradually slow down under the assumption that the mortality rate will not change much in the next five years. In 2026, the growth of China's population has been very gentle, and it can be known that the net increase of China's population will gradually tend to zero.

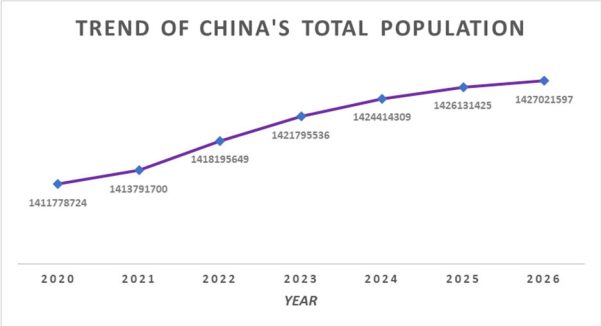

Fig. 2 Trend of China's total population from 2020 to 2026.

It can be concluded from the above results that although the opening of the three-child policy do have positive significance for the increase of the newborn population and total population in China in the early stage, it still cannot restrain the decrease of newborn population and the net increase of population as time went by.

The three-child policy still cannot save China's low fertility rate. The main reason is the low willingness of families to have three children. Therefore, it is not enough to implement the three-child policy only from the perspective of policy.

5.3. Prediction of China's future age structure of population

The population coefficient by age refers to the proportion of the total population of different age groups in the total population. According to age groups, the population aged 0-14 is the children population, aged 15-64 is the labor force, and aged 65 and above is the aging population. Among them, the coefficient of aging population is the most intuitive index to reflect the aging trend of population.

In this paper, the modified female demographic characteristics are selected to reflect the overall population characteristics, and the prediction results of population coefficient by age are shown in table IX and X.

Tab. 11 The population of Chinese females by age groups (2021-2026)

Year	*Children Population*	*Labor force*	*Aging population*
2021	107907234	480297462	102669062
2022	107508306	478041842	107672210
2023	106732093	476803324	111643624
2024	105710548	477334561	113611767
2025	105299274	476339336	116055721
2026	103712068	477254249	117361274

Tab. 12 The population coefficient of Chinese females by age groups (2021-2026)

Year	*Children(%)*	*Labor force(%)*	*Aging(%)*
2021	0.156189511	0.695202931	0.148607557
2022	0.156189511	0.689593803	0.155321318
2023	0.153531805	0.685871260	0.160596936
2024	0.151739761	0.685178855	0.163081384
2025	0.150924652	0.682733563	0.166341785
2026	0.148514922	0.683424593	0.168060485

In order to make the trend display more intuitive, the prediction results are visualized as following diagrams.

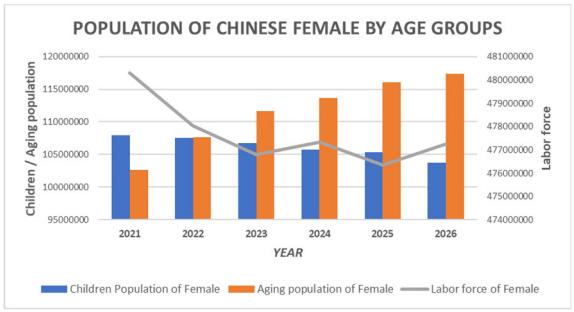

Fig. 3 Trend of female population in China (2021-2026).

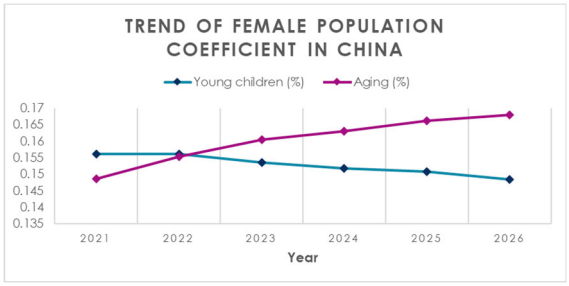

Fig. 4 Trend of female children and aging population coefficient in China (2021-2026).

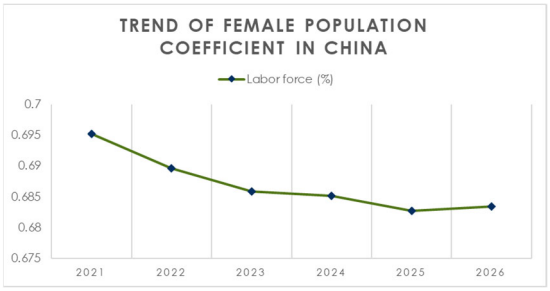

Fig. 5 Trend of labor force population coefficient in China (2021-2026).

The change in the female population reflects the change in the total population. In the next five years, China's labor force population and the population of young children will approximately continue to decline, especially after 2022, there will be an obvious trend of decline. The aging population has an apparent upward trend, and the aging degree is still increasing. By 2026, the proportion of female ageing population will be 16.8%, which has seriously exceeded the proportion of international aging population, that is, the elderly population over 60 accounts for 10% of the total population.

According to the model results, the aging of China's population has become an irreversible trend, and the disappearance of demographic dividends cannot be reversed. Even with the opening of the three-child policy, the aging problem has not been effectively improved. Under the condition of the three-child policy, we should introduce a series of relevant laws, regulations and policy subsidies, improve infrastructure, and encourage women of childbearing age to give birth actively to slow down the aging degree.

References

[1] Meng Lingguo, Li Chaoling, Hu Guang. China population: resources and environment [J], 2014,24(2):132-141.

[2] Zhang Boling. Let the three-child support policy come into effect [N]. People's Daily, 2021-08-18(7).

[3] Wang Jun, Li Xiangmei. Low fertility situation, population policy dilemma and outlet under Chinese three-child policy [J]. Youth Exploration, 2021, (5):50-6.

[4] Long S M, Liu Y J, Yang L, et al. Double-exponent Model of Population Evolution and China's Population Pridiction[J]. Journal of Pingdingshan University, 2010.

[5] Stochastic Model for Population Forecast:Based on Leslie Matrix and ARMA Model[J]. Population Research, 2011.

[6] Hao Y H, WANG. The Dynamic Model of Gray System and Its Application to Population Forcasting[J]. Mathematics In Practice and Theory, 2002.

[7] Liu Xiaojing. A study on the influence of causes of death on life expectancy in Chinese population [D]. Tianjin University of Finance and Economics,2019.

[8] Song Lei. Jinan: More than 80% of residents do not want to have a third child [J]. Sichuan Provincial Information,2021(08):58-59.