

A Study of Population Change Led by China's Great Famine (1958-1961)*

——based on modified Lee-Carter model

Mi Hong^{1,2}、Yang Mingxu^{1,2}、Jia Ning²

(1.Population and Development Research Institute, Zhejiang University, Hangzhou ,31000, China

2、 School of Public Affairs, Zhejiang University, Hangzhou ,31000,China)

Abstract

Population study of hard times during 1958 and 1961 after Great Leap Forward is of great importance for China's contemporary population study. As for the exact number of abnormal deaths, there is no official report by Chinese government. Moreover, because of social disruption at that time, there is something wrong with the registration and preservation of demographic data. As a result, by now, there is only estimation and prediction of abnormal deaths during China's great famine, which can be done from an in-depth and micro perspective.

Although Lee-Carter Model is a very widely applied method to predict death rate in the world, it is found from the study conducted in this essay that the original model can not be mechanically used under China's situation. With modified Lee-Carter model, the study made an index curve which limits the death rate in the original model into "logistic" curve and made an adjusted parameter. Simulation results show that there is a big improvement in fitting goodness of parameters and actual prediction of modified model which can be further applied in the reverse

*This work is supported by the following funds:

- Zhejiang Provincial Natural Science Foundation of China (LZ13G030001);
- The Ministry of Science and Technology of China "Simulation Model Platform of Labor and Social Security Policy and Decision Support Systems"(2012BAK22B02);
- The Major Project of the National Social Science Fund (12&ZD099);
- The Major Social Science Project of Ministry of Education (12JZD035).

prediction of death rate.

Based on modified Lee-Carter prediction model of mortality, this study did reverse prediction with reference to China's official age-specific mortality data to re-construct and calculate single-age mortality and normal deaths if there were no great famine in the last century. In the end, with regard to total deaths estimation calculated by scholars home and abroad, we did estimation research of abnormal deaths led by famine.

It is shown in our result that population dynamics, especially abnormal deaths, during great famine, mainly occurred between 1958 and 1961. If there were no famine, normal deaths would be 43.39 to 43.85 million. With famine, abnormal deaths are between 16.24 and 23.37 million.

1、 Introduction

Periodicity of famine is China's feature. From BC 108 to AD 1911, there are 1828 times of known famine in total. According to Kane, the author of *China's Big Famine(1959-1961)*, in the early 30s of this century, it was found that most of the middle-aged people the author talked to in about 63% of China's counties had at least three times of starving experience. However, there is a surprising insufficiency of China's famine study.

Since the founding of new China, during 1958 and 1961, Big Leap Forward Movement and industrial development at the cost of sacrificing agriculture led to national food shortage and famine which was called by those in the countryside with the experience of famine hard times and lean year and three years of natural disaster by authority before 1980, then changed into three years of hard times. Scholars abroad called it three years of famine and great leap forward famine.

Studies done by scholars home and abroad are mainly based on census data in 1953, 1964 and 1982, western life table and data released by national ministry of health and there is big difference in the conclusion of each study with varying methodology. Therefore, scholars home and abroad in the field of history, demography and statistics are all concerned about such questions as how is the population change with big famine and what the exact number of abnormal death is. Making clear of the above questions is of great significance to the lesson learned from big famine and demographic, social and economic development since the founding of new China.

2、 Literature review

East Asia and the Pacific Department of the U.S. hold a careful attitude to China's population decline during great famine. In *Background Notes: China, October 1997*, it said that there were millions of people died of starvation during this period. *Cambridge Chinese History* (1990) estimated that there were 16 to 27 million deaths. People died in 1960 was over 10 million.

According to the revised data estimated by J.Bannister(1985), chief of China's section of Census Bureau of the U.S., there were 29.871 million abnormal deaths with 3119.5 fewer births and there was a total population decline of 61.066 million.

Ansley.Coale(1984), president of American Population and Demographic Committee and professor of Princeton University, with three times of census data and 1 % sampling survey data of fertility in 1982, calculated the newly-born population during census intervals and deaths over the years, and then got annual linear death toll with linear trend and ultra pure linear deaths by compared with actual deaths. It is shown in his revised result that there were 24.81 million abnormal deaths during 1959 and 1961 with 30.683 million fewer birth and 55.493 million people loss and 26.8 million abnormal deaths during 1958 and 1963.

The revised calculation of G.Calot(1985), the former director of French national institute of population is that the abnormal deaths during 1958 and 1962 are 28.509 million with 331.9785 million fewer deaths and the total people loss is 60.488 million.

Penny. Kane(1993) introduced research findings of some scholars in his book *China's Great Famine During 1959 and 1961*. Eder believed that the abnormal deaths between 1960 and 1961 are 23 million. Mosle estimated that the abnormal deaths in 1960 are between 11 million and 30 million. Besides, Hill calculated that the abnormal deaths during 1958 and 1962 are 30 million with 33 million unborn or delayed-born babies.

The former professor of Xi'an Jiaotong University, China, Jiang Zhenghua et. (1986,1988) published two essays on death toll during great famine. Firstly, they worked out life table in 1981 based on death data in 1982, then they established dynamic parameter estimation model. With national age-specific and gender-specific data in 1953, 1964 and 1982 census, they obtained fertility rate and death rate during the first and the third national census and estimated the deaths over the years. Finally, they estimated normal deaths according to changes in life expectancy. With the

deaths over the years minus normal deaths, they got the abnormal deaths of 17 million during 1958 and 1963.

Cao Shuji(2005), historical geographer and professor in history department of Shanghai Jiaotong University, rebuilt abnormal deaths during 1958 and 1963 around China with the method in the field of demography and historical geography. Firstly, with China's census data in 1953, 1964 and 1982, he calculated rate of population increase before and after famine based on local chronicles to confirm population number before and after famine in each place and calculate abnormal deaths. According to him, during 1958 and 1961, the total abnormal deaths around China is about 32.458 million.

The former director of National Statistics Bureau Li Chengrui (1997) pointed out that there were weaknesses in Coale's study. With the amendment of Coale's contradiction, Li Chengrui calculated the abnormal deaths during famine was 22 million. Meanwhile, he showed the compared with studies conducted by western scholars, the methodology applied by Jiang Zhenghua tended to be more scientific.

Luo Sheng (1988), from University of Pennsylvania, wrote in his doctoral dissertation that since there was no official announcement of age-and-gender-specific death rate and death toll during 1953 and 1982, not even age structure, general method is not applicable to rebuild annual life table during this period. General method refers to the calculation of life table and age- and -gender-specific distribution with age structure in 1953, 1964 and 1982 census and annual total population. First of all, with Brass-Logit system and model life table, he set up "error reduction" iterative algorithm and set three different population distribution and born mode to rebuild three kinds of age-specific survival probability rate during 1953 and 1982 and compare with the result of Coale and Bannister. Then, with population distribution obtained by three census, he assumed that sex ratio at birth between 1953 and 1964 was 105 and 106 between 1964 and 1982 to rebuild annual life table and age-and -gender-specific population distribution. On the basis of the above research results, he calculated that the abnormal deaths during 1959 and 1963 were 28.89million[10]. However, Luo did not show the exact counting process to demonstrate the calculation of normal deaths. Since there is no calculation result of total deaths and normal deaths, his result is not convincing.

3、 Methodology

3.1 Data

There is not sufficient research on population death in China, which is on the one hand because China's population study is mainly carried out under the background of population control with the focus on fertility, and on the other hand there has been no consistent, normative and systematic declared death data. Although death registration is in the charge of some government departments in China, there is no announced data available for research and reference. There was official record of death rate since 1954 before which there had only been estimation value. The items on age-specific death toll started to be included in the survey since the third census in 1982, followed with continuous statistics of age-specific death in 1994.

Age-specific death rate applied in this study is from *China's Statistical Yearbook of Population (1994-2006)*, *China's Statistical Yearbook of Population and Employment (2007-2010)* and the data of 2000 is from the website of China's Bureau of Statistics. Total population structure and other data is from census in 1953, 1964 and 1982 and sampling survey in 1987, 1995 and 2005. There are following problems to be solved as for this data set^[11]:

1. Samples are different. There is changing sampling in some years, milli sampling and survey data in some years.
2. Missing of Age-specific death data differs from age cohorts. For example, there is no data of 1995 and death data in some years is based on one-year age-specific grouping to 100 years old. Some are grouped by 5 years each. And the last cohort in most years is 90+ with 85+ in 1996 and 100+ in 2005.

In this study, there is model fitting by single age cohort. Compared with data in 5-year cohort, old-age death rate in single age cohort, especially data from sampling survey is higher than that in low age cohort.

The solutions are as follows.

1. In this study, we assume that changing sampling and milli sampling are with the feature of random sampling.
2. Death function is used to make up for old-age death data.
3. There is smooth disposal of death rate in single age cohort, of which firstly, median filter is

applied to insure proper noisy point in death rate curve and then smoothness happened with death function.

4. Interpolation fitting is used in missing data of several years.

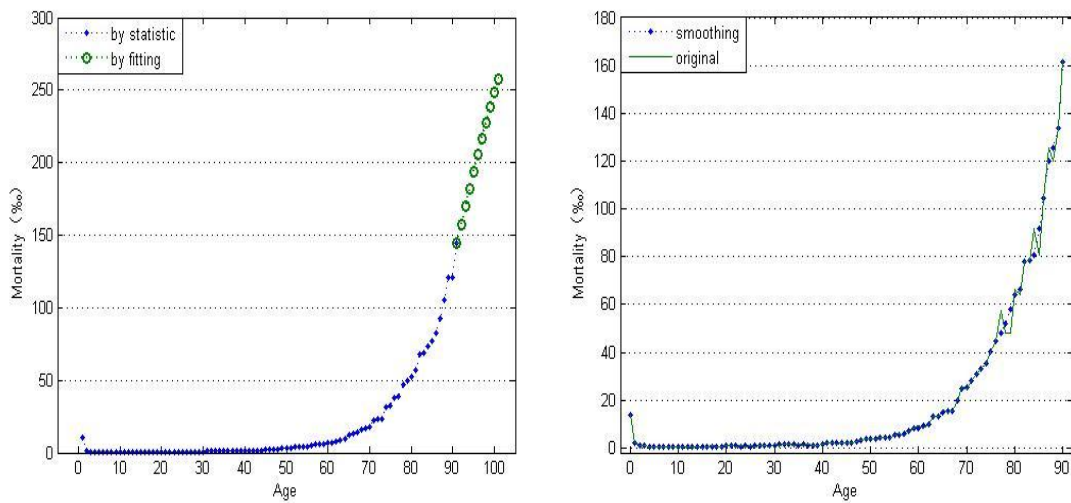


Figure 3.1 Data completeness and smoothness of some years

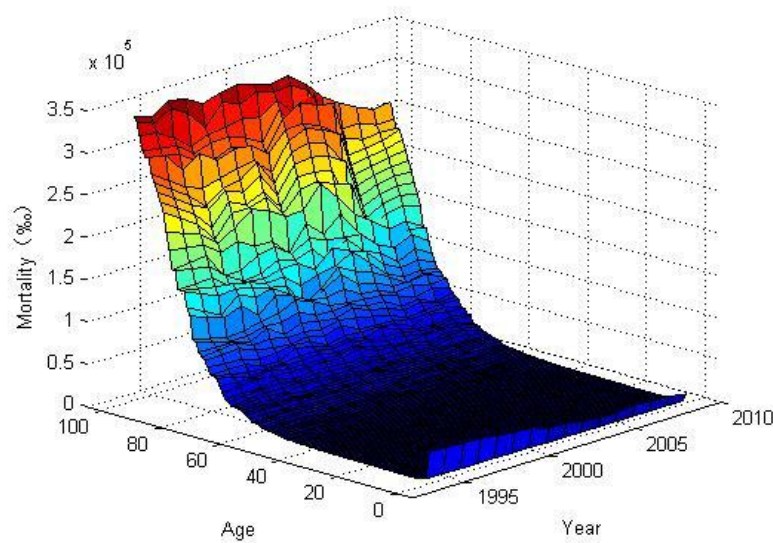


Figure 3.2 Age-specific death rate during 1994 and 2000

3.2 Method

Lee-Carter model is currently widely applied around the world to predict death rate. It is found in this study that for China’s situation, the original model can not be mechanically used since it can only do forward prediction with no reversed prediction of death rate,

Therefore, there is modification of the model to make it better used in the fitting of death rate

in China and the reverse estimation of death rate in the past years.

3.2.1 Classic LEE-CARTER Model

Lee-Carter model is extropolation one in which there is logarithm analysis of death rate and prediction with ARIMA time order. Specifically, it is based on age-specific mortality which is made log to show it in a linear combination with 3 parameters and then to predict future changing tendency of mortality with random time order. It is shown as follows:

$$\ln(m_{x,t}) = \alpha_x + \beta_x k_t + \varepsilon_{x,t} \quad (3.1)$$

Among which $m_{x,t}$ is rough mortality at time t and age x (from statistical materials in each year)

α_x is base number of age-specific mortality by time and mean value of natural logarithm of age-specific mortality. It reflects base number of log change of age-specific mortality and depicts time trend of age pattern of mortality.

β_x is the influence of k_t on the mortality of each age cohort, which reflects the influence of k_t on mortality log in each age cohort and relative changing coefficient of mortality in each age cohort.

k_t is changes of death level at time t and it is index of death level. It reflects the relative intensity of past death rate and the indictor of total mortality.

$\varepsilon_{x,t}$ is residual whose mean value is 0 and variance is δ_ξ^2 .

To get the single solution, there are two standardized restricted conditions:

$$\left\{ \begin{array}{l} \sum_{all\ x} \beta_x = 1 \\ \sum_{t=i}^j k_t = 0 \end{array} \right. \quad (3.2)$$

Among which i, j is the beginning year and ending year of mortality.

3.2.2 Weakness of Lee-Carter model

1、convergence

With Lee-carter model, the prediction of future mortality based on past experience is always lower than the actual result because of born error from it queue effect. Let us specify it with some specific age.

For some specific age, Lee-Carter mortality prediction model , in (3.1), when x is fixed, $m_{x,t}$ will be decreasing with time passing by to zero. It is because:

Firstly, change (3.1) into

$$m_{x,t} = \exp(a_x + \beta_x k_t) \quad (3.3)$$

For some specific single age, say 0 age cohort , $x=0$, there is :

$$m_{0,t} = \exp(a_0 + \beta_0 k_t) \quad (3.4)$$

Since mortality rate $m_{0,t}$ is the function of k_t changing with time t , a_0 and β_0 are not changing for time t , in forward prediction, changes of parameter k_t which is positively related to matched mortality tends to be reducing. With social development, the total decline of mortality is for sure with $m_{0,t}$ changing into 0.

Because of the above, Lee-carter model can only be used in the prediction of next few years with sufficient historical data(such as in the U.S. and Japan) in stead of the prediction of many years with a small amount of data. By now, in China, there is only 20 years of mortality data which can be used in the fitting of Lee-Carter model.

2、reverse prediction

In common sense, Lee-Carter mortality prediction model is based on available data to predict future mortality. And the estimation value tends to be decreasing. Then why not use it in the calculation of historical mortality? For example, in forward prediction, we used data set during 1994 and 2009 to predict 2020, then we can do reverse prediction in 1980 with the data from 2009 to 1994. However, Lee-Carter model can not be used in reverse prediction for the following reasons.

In reverse prediction of historical data with available data, k_t is increasing, which is different

from $\lim_{t \rightarrow -\infty} m_{x,t} = \lim_{x \rightarrow -\infty} \exp(a_x + \beta k_x) \rightarrow$

but $\lim_{t \rightarrow +\infty} m_{x,t} = \lim_{x \rightarrow +\infty} \exp(a_x + \beta k_x) \rightarrow +\infty$, however, as long as

$\lim_{t \rightarrow +\infty} m_{x,t} = \lim_{x \rightarrow +\infty} \exp(a_x + \beta k_x) > 1$, it is wrong since mortality can not be over 1 .

3.2.3 Modified model

To do reverse prediction of historical mortality by the model, firstly, convergence of the final item must be solved. Logistic curve is a good choice. In fact, even though there is the convergence of the final item, prediction of historical data is different from that of the future since “0” can be set as the limit of the mortality rate of some age in the next few years, which will not be the problem for the forward prediction of next few years. However, even there is reverse convergence in the model and limit of Logistic curve is 1, that is, mortality rate by reverse prediction is close to 1, it is not realistic to recent several decades. Therefore, there must be an adjustment of parameters which can play the following role.

1. To limit function convergence

Adjustment of mortality limit of next few years is like setting 0 as mortality limit in classic Lee-Carte model. For historical reverse prediction, mortality rate tends to be backing up. If based on Lee-Carter model, limit of mortality rate would be 1. However, the actual situation is because there is general analysis of mortality rate in the census and statistics of some years, we can have a general understanding of mortality level at that moment to estimate a limit range which will not be close to 1.

2. To improve fitting goodness of parameters
3. To vary changing tendency in different age cohorts

The changing tendency of each age cohort year by year is oneness and sameness. Whereas, there should be diversity in yearly changing tendency of each age cohort, which is the acceleration of increase or decrease should be different.

To solve the above problem, there is the modification of the model which is shown in the following:

$$m_{x,t} = \frac{1}{\gamma + e^{a_x + \beta_x k_t + \varepsilon_{x,t}}} \quad (3.5)$$

Among which, a_x is mortality rate of mean log in terms of time. β_x is changing rate of mortality rate. k_x is random time effect which complies with random walking with drifting or ARIMA process. γ is adjustment parameter to adjust convergence effect of the model. For different age cohort, there will be good fitting effect and mortality convergence of β_x with different γ , which is important to reverse prediction.

To get the solution of parameters, there is no consideration of residual and γ is regarded as constant. Therefore,

$$\frac{1}{m_{x,t}} - \gamma = e^{a_x + \beta_x k_t} \quad (3.6)$$

$$\text{Take the logarithm: } \ln\left(\frac{1}{m_{x,t}} - \gamma\right) = a_x + \beta_x k_t \quad (3.7)$$

The above reflects the assumption of the new model, which is the log value of mortality reciprocal, tends to be changing in curves. In order to get the only solution, as shown in Lee-Carter model, there are two limit conditions of the model which is the same as (3.2).

The parameter solution of the model is as follows:

The first step: $\gamma=1$, to complete the first estimation of β_x and estimation of α_x and k_t

The second step: Based on the above parameter, there is the calculation of error. Then adjust parameter γ_x and re-estimate β_x .

1. Estimation of α_x

The estimation presupposition of α_x and k_t is $\gamma=1$, then model (3.5) is changed into:

$$m_{x,t} = \frac{1}{1 + e^{a_x + \beta_x k_t}} \quad (3.8)$$

The above is similar as Lee-Carter model, in which the changing curve of mortality is changed from exponential type into “S” type and the convergence of final item is limited to 1 by divergence.

Sum of t and equation (4.7) is:

$$\sum_{t=i}^{t=j} [\ln(\frac{1}{m_{x,t}} - 1)] = \sum_{t=i}^j a_x + \sum_{t=i}^j \beta_x k_t \quad (3.9)$$

Among which i and j are the beginning year and the ending year of actual observation data of mortality. With the satisfaction of restricting conditions, since $\sum_{t=i}^j k_t = 0$ there is:

$$a_x = \frac{\sum_{t=i}^{t=j} [\ln(\frac{1}{m_{x,t}} - 1)]}{n} \quad (3.10)$$

n is observation year.

2. Estimation of k_t :

The sum of single age cohort in some specific year is:

$$\sum_{all\ x} [\ln(\frac{1}{m_{x,t}} - 1) - a_x] = \sum_{all\ x} \beta_x k_t \quad (3.11)$$

When t is not changed, say, it is the same year, k_t value in different age cohort is unchanged.

Moreover, because $\sum_{x=0}^{100} \beta_x = 1$ there is:

$$k_t = \sum_{x=0}^{100} [(\frac{1}{\ln m_{x,t}} - 1) - a_x] \quad (3.12)$$

3. Initial estimation of β_x :

According to the above estimation value, there is equation with least square method:

$$\frac{[1 / m_{x,t}] - 1}{\exp(a_x)} = \exp[\beta_x k_t] \quad (3.13)$$

With fitting method of non-linear equation, for a specific age cohort $\exp(a_x)$ is a constant term.

To take the left part of the equation as independent variable and k_t as dependent variable, there is the estimation of β_x in each age cohort. Therefore, fitting of a_x , k_t , β_x under the condition of $\gamma = 1$ is completed.

Choosing the proper γ value and match with β_x to satisfy fitting goodness and set a

reasonable limit to the changes of historical mortality is important to the accuracy of reverse prediction.

With plentiful experiments, there is the judgment of fitting goodness of β_x with R-square index. On the one hand, there is a great need for R-square whose value is close to 1, on the other hand, Sum (R-square) should be big.

1. to satisfy reasonable setting of mortality limit in different age cohort

There should be the setting of mortality limit of older age and younger age with estimation parameter γ to do error self adjustment. Besides, according to historical data and the research of some scholars, we can estimate the mortality level of older and younger age in some specific year for reference.

Estimation steps of γ :

The first step: put a_x 、 β_x 、 k_t obtained by $\gamma=1$ into equation (3.5) and set the left part in the equation as exact mortality $m_{x,t}$ of fitting year. The right part of the equation

is $\frac{1}{\gamma + e^{a_x + \beta_x k_t}}$. The left side equals the right side and then there is the calculation of error term γ which is $\gamma_{x,t}$ changing with x and t .

$$\gamma_{x,t} = \frac{1}{m_{x,t}} - \exp(a_x + \beta_x k_t) \quad (3.14)$$

The second step: to equalize it in terms of t and there is $\bar{\gamma}_x = \frac{\sum_{t=i}^j \gamma_{x,t}}{n}$, then to do median filter and smooth disposal and there is $\bar{\gamma}_x$ which is put into equation (3.8) to replace γ_x .

The third step: based on new γ_x , to re-estimate β_x and there is the parameter γ_x changing with x , which shows different convergence in different age cohort with different changing pattern.

If parameter γ is set as 1 permanently, then there is sharp change with k_t , and mortality limit in 0 age cohort and old age cohort is gradually close to 1, which leads to the same problem as in the forward prediction of Lee-Carter model with mortality rate close to 0 in 0 age cohort.

4、 Reconstruction result of mortality rate

There is analysis by time order in Lee-Carter model, which needs historical data of mortality in 30 years or more. However, there is fewer than 20 years of mortality data in China. Meanwhile, mortality data quality also restricted the application of the model. Since in the model there is the assumption that β_x will not change with time. But there may be interaction between age and time, especially obvious in the single age cohort of this study. Many research tried to reduce the interaction effect between age and time in β_x . In simulate age-specific mortality time order in the U.S., Lee and Miller (2000) found that there is a significance change of age pattern in the decrease of American population mortality. They suggested that the relevant data can be considered into two time period, of which one is from 1900 to 1950 and the other one is from 1950 to 1995. By this method, they think that fitting effect of β_x is good.^[12] Tuliapurkar et.al (2000) also suggested that mortality rate of the U.S. after last 1950s should be considered by staging to enhance fitting goodness of the model.^[13]

Carter and Prskawetz (2001) favorably selected 30 24 -year time record of age-specific mortality from 1953 in Australia, with SVD method, to generate 30 sets of a_x, β_x and k_t . By this method, they observed β_x ^[14] in different period.

The above study shows that parameter fitting should be based on effective analysis of basic data since mortality in different period shows varying changing pattern. In terms of China's 20 years of basic data, it is not scientific to do prediction or reverse prediction of mortality rate in 30-50 years. Therefore, 5-year reverse prediction was done in this study, followed with a reconstruction into basic data. Then with available data, there is optimization disposal, by which iterative prediction occurred to make the reverse prediction in different periods be with its own fitting parameter and ensure the stability of reverse prediction.

So in this study, 5-year prediction method is applied. Firstly, mortality rate during 2009 and 1994 is used to reconstruct data between 1993 and 1989. Then there is the integration of data from 2009 to 1989 to reconstruct mortality data between 1988 and 1984. By this method, , it was back to 1950 with some statistical data during intervals. In this chapter, there will be the detailed introduction and the fitting result of parameters of the first backing process.

1. Estimation result of a_x , k_t , β_x , γ_x

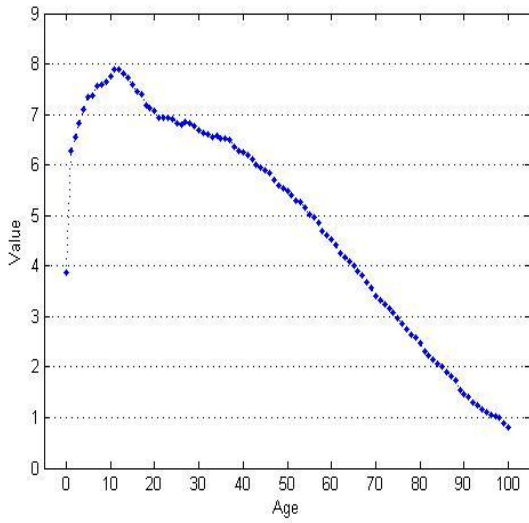


Figure 4.1 Fitting value of a_x

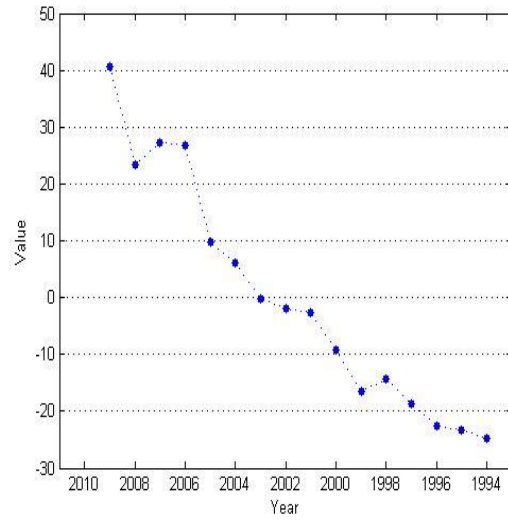


Figure 4.2 Fitting value of k_t

a_x reflects the base of log change of each age-specific mortality. It describes relative time trend of mortality age pattern. Being different from “U” shape trend in forward prediction of the classic model, in reverse prediction, since the parameter is in the position of denominator, there is reversed “U” shape trend in graph4-3.

In the new model, k_t is negatively related to mortality. From graph 4-4 it can be seen that k_t value is small if it were in the first several years, which reflect mortality rate tends to be increasing at reverse time.

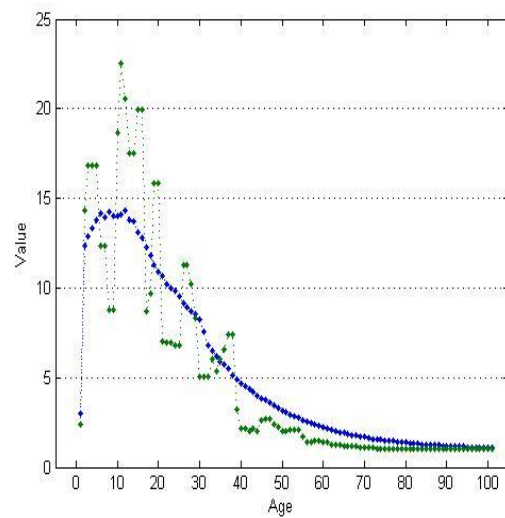
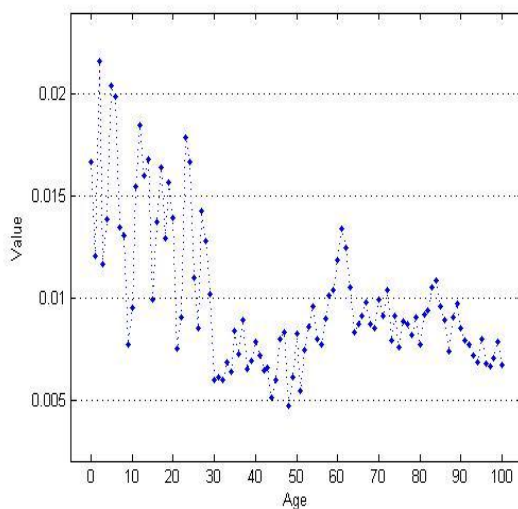


Figure 4.3 Fitting value of β_x

Figure 4.4 Fitting value of γ_x

It is seen from graph 4.3 that value was bigger in young-age cohort, which is mainly because there is a high mortality rate in newly-born babies and it is sensitive to mortality changes. β_x value is quite low in 10-and 11-year age cohort, which is because of the restriction of basic statistical data, resulting in the poor fitting effect of these two age cohort. β_x over 85 years old tends to be decreasing, which is due to that mortality in old-age cohort shows few changes with time and actual mortality is not sensitive to parameter changes.

Smooth fitting of γ_x is for its trend. It can be seen from the fitting value that it shows “U” shape, which is the same as a_x . γ_x shows no big changes since its value will determine prediction convergence limit of each age cohort and mortality prediction limit of adjacent age is not big.

2. Fitting effect

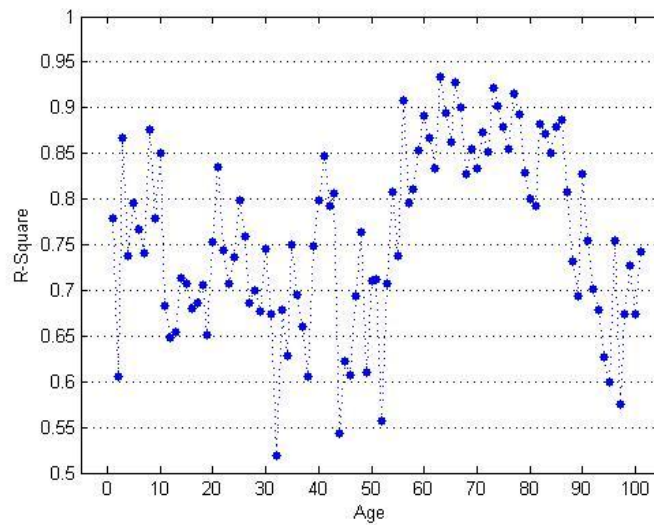


图 4-5 Fitting effect of β_x

It can be seen from the above graph that there are 95 points of R-Square>0.6 and 70 points of R-Square>0.7.

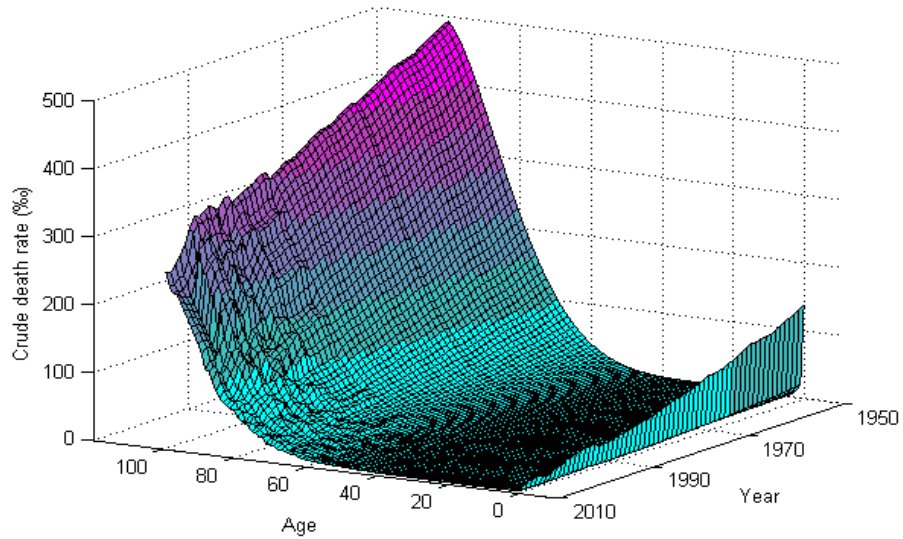


Figure 4.6 Mortality of single age cohort (1953—1993)

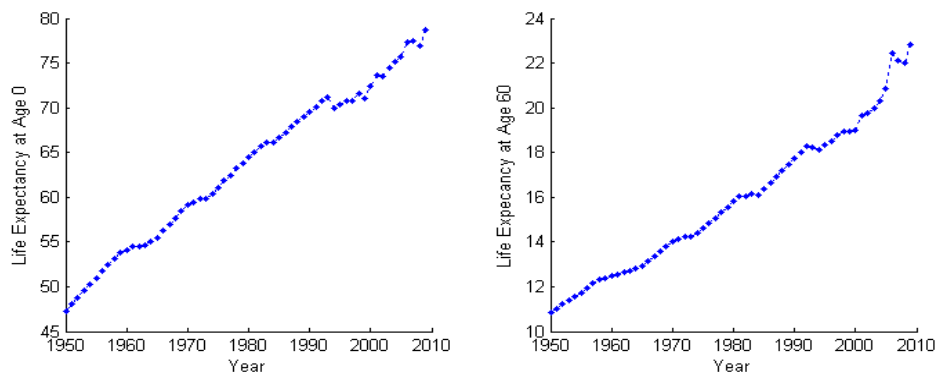


Figure 4.7 Changing trend of life expectancy (0 age cohort and 60 age cohort)

5、 Analysis of China’s abnormal deaths during 1958 and 1961

5.1 Concepts and research boundary

1. Total deaths (actual deaths) refers to the total population died during this period, including normal deaths and deaths of starvation.

2. Abnormal deaths refers to those died of starvation and subsequent illness.

$$\text{Abnormal deaths} = \text{Total deaths} - \text{Normal deaths} \quad (5.1)$$

or:

$$\text{Abnormal deaths} = (\text{Total mortality} - \text{normal mortality}) * \text{Actual population} \quad (5.2)$$

3. Population loss refers to those unborn people because of famine.

4. Normal deaths refers to those with no experience of famine.

To calculate normal deaths, first of all, there must be normal mortality of each single age cohort, then with population prediction method, there is the reconstruction of population number of yearly single age cohort. With the above procedure, we can get normal deaths in each year.

The achievement of this study is the calculation of normal deaths in single age cohort without famine.

Since there is no study of actual mortality in single age cohort and actual deaths, the calculation of abnormal deaths is the estimation of some scholars' research results, which is based on their estimation of total deaths to evaluate their estimation of abnormal deaths.

5.2 Calculation procedures

1. Reconstruct age-specific population structure of 1953.

With reference to age structure of census in 1953 and reconstruction method by Mi Hong(1996), there is the calculation of single age cohort structure of 1953. ^[15]。

2. Based on the situation at that moment, we set normal infant fertility and sex ratio at birth during 1953 and 1964.;

In this study, we take mean value of census in 1953 and 1964 as the birth rate of 1954 to 1963, as total birth rate and 105 as sex ratio at birth as.

3. Then put reconstructed population structure, annual birth rate, sex ratio and age-specific mortality into population prediction model to calculate population structure between 1953 and 1964.

4. With reference to the estimation of actual total population of 1958—1963 by some scholars, we separate the predicted population structure based on age cohort. Then combine with normal mortality of single age cohort to get normal deaths with the consideration of different actual total population.

5. Refer to available research results on total deaths by some scholars to estimate abnormal

deaths.

5.3 Calculation results

5.3.1 Total population

Table 5-1 Prediction in this study vs. reconstructed annual total population by other scholars (unit: people)

Year	Our study	Coale	Bannister	Calot	J.Zhenghua
1953	5.88E+08				
1954	6.02E+08				6.03E+08
1955	6.14E+08	6.14E+08	6.07E+08	6.14E+08	6.15E+08
1956	6.26E+08	6.26E+08	6.19E+08	6.27E+08	6.26E+08
1957	6.37E+08	6.39E+08	6.33E+08	6.45E+08	6.39E+08
1958	6.49E+08	6.45E+08	6.39E+08	6.58E+08	6.48E+08
1959	6.61E+08	6.52E+08	6.51E+08	6.63E+08	6.55E+08
1960	6.73E+08	6.49E+08	6.53E+08	6.53E+08	6.52E+08
1961	6.86E+08	6.46E+08	6.48E+08	6.50E+08	6.49E+08
1962	6.99E+08	6.55E+08	6.53E+08	6.64E+08	6.68E+08
1963	7.12E+08	6.87E+08	6.74E+08	6.80E+08	6.87E+08

1. Our reconstructed total population refers to the number with no famine.

2. Reconstructed population by other scholars refers to the actual population after famine.

Since 1958, our predicted total population started to exceed that by scholars, which is due to excessive deaths with famine.

5.3.2 Calculation of annual normal deaths and total deaths

With the consideration of morality in each age cohort, we sum the deaths in single age cohort.

Calculation procedures are as follows:

1. According to our predicted annual population structure, separate annual total population calculated by the scholars into population of each single age cohort.

2. Calculate deaths in each single age cohort.

Death in single age cohort =total population in each single age cohort ×morality in single age cohort

3. sum deaths in single age cohort based on age.

$$D_t = \sum_{i=0}^{100} P_{i,t} \times q_{i,t} \quad (5.3)$$

Among which D_t is t total deaths in year t and $P_{i,t}$ is population number at age I in year t .

$q_{i,t}$ is morality at age i in year t .

Annual total population is divided by total deaths and then there is total morality of each year, which is shown in the following table.

Table 5-2 1953~1964 Normal deaths and total morality of each year (with no famine)

Year	Total population	Normal deaths	Total morality rate (‰)
1953	594350000	11669052	19.63
1954	604377430	11535132	19.09
1955	614701515	11426000	18.59
1956	625567021	11338941	18.13
1957	636915594	11251785	17.67
1958	648748888	11103846	17.12
1959	660895459	11040751	16.71
1960	673276278	10962556	16.28
1961	685899625	10744194	15.66
1962	698873105	10532090	15.07
1963	712432481	10339914	14.51
1964	726569955	10091475	13.89

Table 5-3 Estimated total mortality of each year by scholars (unit: ‰)

Year	Bannister	Coale	Calo	J.Zhenghua
1954	24.2	29.1	19.9	
1955	22.33	22.4	22.3	17.23
1956	20.11	20.8	16.8	16.77
1957	18.12	19	13.2	16.55
1958	20.65	20.4	15.9	17.25
1959	22.06	23.3	19.2	18.96
1960	44.6	38.8	40.7	31.25
1961	23.01	20.5	27.0	24.57
1962	14.02	13.7	18.2	18.08
1963	13.81	13	21.2	16.72
1964	12.45	13.5	20.8	13.03

The calculated total normal deaths during 1958 and 1963 by our study is 6472335 which is not equal to those obtained by the calculation of Coale and Jiang Zhenghua et al. Since the assumption is with no famine, our calculated normal deaths is big with a big population base.

The actual normal deaths should be based on actual total population between 1958 and 1963 with reference to annual age-specific mortality.

With total population data by the scholars, there is the sum of single age cohort deaths.

Table 5-4 Abnormal deaths (unit: people)

Year	Coale	Bannister	Calot	J. Zhenghua
1958	11237012	11126074	11460631	11290479
1959	11127944	11095876	11315064	11168540
1960	10716662	10781932	10797299	10780114
1961	10307882	10340929	10375891	10365035
1962	10133692	10106619	10270911	10327841
1963	10102460	9919865	10010648	10115408
总计	63625652	63371294	64230444	64047418

Total population calculation of each year and deaths between 1958 and 1963: (annual total deaths = annual total population × annual actual mortality)

Table 5-5 Annual total deaths (unit: people)

Year	Coale	Bannister	Calot	J. Zhenghua
1958	13162488	13192253	10515799	11101065
1959	15200454	14350030	12736320	12323810
1960	25163740	29101500	26633807	20470000
1961	13236440	14904728	17567878	15953792
1962	8974185	9159266	12136458	11840411
1963	8925930	9310702	14437451	11303222
总计	84663237	90018478	94027713	82992301

5.2.3 Calculation of abnormal deaths

With no estimation of actual deaths during famine, there is the reference of available results to get abnormal deaths, total deaths and total morality rate

Table 5-6 Results by previous scholars (unit : Ten thousand people)

	Coale	L.Chengrui	J.Zhenghua
Total deaths	8620	8620	8299
Normal deaths	5940	6462	6602
Abnormal deaths	2680	2158	1700

With our calculated normal deaths, the following is the correction of abnormal deaths by previous scholars.

Table 5-7 Correction of research results by previous scholars (unit: Ten thousand people) (1958-1963)

	Coale	L.Chengrui	J.Zhenghua
Total deaths	8466	8466	8299
Normal deaths	6363	6462	6405
Abnormal deaths	2103	2004	1894

If there is only the calculation of abnormal deaths between 1958 and 1963, then,

Table 5-8 Correction of research results by previous scholars (unit: Ten thousand people) (1958-1961)

	Coale	Calot	J. Zhenghua
Total deaths	6676	6745	5984
Normal deaths	4339	4395	4360
Abnormal deaths	2337	2350	1624

Why in Coale's research, there are more abnormal deaths in 1958—1961 than that in 1958—1963? We found that the birth rate risen from 22‰ in 1961 to 47‰ in 1963, and the death rate declined from 20.5‰ to 13‰ at the same time in Coale's research, even lower than the real death rate estimated by our paper. So there will be more people alive in 1962 and 1963 and will offset the abnormal deaths if take 1962 and 1963 as the famine years.

Therefore, we think it will be more scientific to take 1958—1961 as the famine years.

Table 5-9 Abnormal deaths (Adjusted the famine time)

		Coale	J.Zhenghua
Previous	1958-1963	2680	1697
Now	1958-1961	2337	1624

6、 Conclusion

Firstly, there is the modification of Lee-Carter prediction model of morality. Simulation results show that modified model did well in the aspect of precision the fitting of China's morality data and it can also be applied in the reverse prediction of morality (reverse prediction of morality in the historical year).

Based on modified Lee-carter prediction model of morality, this study re-constructed normal morality (if there were no great famine) of each single-age cohort in 1953. It shows that the life expectancy for Chinese people in 1953, 1959 and 1964 is 48.43 years old, 52.23 years old and , 55.24 years old respectively.

With analysis of available data and reverse prediction data of this study, it is stated that since 1962, with the increase of birth rate and decrease of morality rate, there was few abnormal deaths, and there was tendency of population compensation. To do the research of population dynamics, especially abnormal deaths during great famine, the most rigorous time period should be between

1958 and 1961.

After the calculation of single age-cohort normal morality and total morality during famine, with population prediction method, it is calculated that the normal deaths in the 4 years (1958—1961) was 43.39 to 43.85 million and these numbers are close to Li Chengrui and Jiang Zhenghua's estimation.

Then there is the evaluation of abnormal deaths research done by Coale and Jiang Zhenghua. During 1958 and 1961, if it were based on Jiang Zhenghua's calculation of total deaths, the abnormal deaths would be 16.24 million. It would be 23.37 million if it were based on Coale's calculation. There is 70 million difference, which is mainly because big difference lies in the calculation of total deaths by Coale and Jiang Zhenghua. But if it were based on Coale's calculation, there would be 21.03 million abnormal deaths during 1958 and 1963, which is much smaller than 26.8 million and 23.37 million of 1958 to 1961. Therefore, according to the above results, this study is in favor of what Li Chengrui did, believing that Coale's method and results show great contingency and randomness, but Jiang Zhenghua's tends to be more scientific.

Meanwhile, there are limitations in this study and in-depth research is called for.

First of all, there is only analysis of total morality in this study with no research on gender-specific morality, which can be fulfilled in the follow-up research.

Next, due to word limit, while analyzing population dynamics during famine period, the researchers did not calculate actual total deaths from micro perspective. As for actual total deaths, there is only reference to previous researcher's achievements. Therefore, the calculated abnormal deaths can only refer to those of one time period. In terms of actual population during famine year and actual total deaths, there has been no exact number by now. There is a big difference between official statistics and scholars' calculation. This study is just an estimating and revising research of methods and results by previous scholars and more in-depth researches are needed.

7、References

- [1] Penny. Kane. 1959-1961China' s Great Famine. *Science Literature Press*. 1993 (In Chinese)
- [2] Cao shuji. The reason for the death during 1959~1961. *Chinese Population Science*.2005(1) (In Chinese)
- [3] Roderick MacFarquha. The Cambridge History of China(1945~1965). Chinese Social Science Press. 1990.
- [4] J.Banister. Surprise and comfirmation to the Chinese census. *International Union For the Seientifie Study of population. Florence*.1985.
- [5] A.J.Coale. Rapid Population Change in China , 1952-1982. Washington.D.C: *National Academy Press*, 1984.
- [6] G.Calot. The suggestion to anlysis the Chinese census data.*International Union For the Seientifie Study of population. Florence*. 1985.
- [7] J.Zhenghua , L.Nan. The dynamic correction for Chinese population data .*Academic journal of Xi'an Jiaotong University* ,1986(3):63-64. (In Chinese)
- [8] J.Zhenghua , L.Nan. The method and result for correctting the Chinese pupolation data. *Alamnce of China's Population, Economic and management Press*, 1987. (In Chinese)
- [9]L.Chengrui . The pupolation change led by the great famine[J]. *The Research for the history of CCP* ,Vol.2 1997:1-14. (In Chinese)
- [10] Luo Sheng. Reconstruction of life tables and age distributions for the population of China,by year,from 1953 to 1982. University of Pennsylvania(Doctoral thesis), 1988.
- [11]Z.Wei, Ch.Binzheng. The forecast for Chinese urban resident's mortality [J]. *Statistics and Management*, 2009(4):736-744. (In Chinese)
- [12] Lee.R D and Miller T. Assessing the performance of the Lee-Carter Approach to Modeling and Forecasting Mortality. *The 2000 meetings of the Population Association of America*. Los Angeles, 2000.
- [13] S.Tuljapurkar, L.Nan, C.Boe. A Universal Pattern of Mortality Decline in the G7 Countries.*Nature*, 2000(405):789-792.
- [14] Carter.L R, Prskawetz.A. Examining structural shifts in mortality using the Lee-Carter method. Mpidr Working Paper WP, 2001(07) : 57-68.
- [15] M.Hong,Z.Yougan. A Reverse forecast method in rebuilding Chinese population during 1946~1949.*Academic Journal of Shanxi Normal University*. 1996(S1): 124-134.
- [16] Lee.R D, Carter L. Modeling and Forecasting U.S mortality. *Journal of the America Statistical Association*,1992(87): 659-671.

- [17] Delwarde.A, Denuit.M, Eilers.P. Smoothing the Lee-Carter and Poisson log-bilinear Models for Mortality Forecasting: a penalized log-likelihood approach . *Statistical Modeling*, 2007,7(1):29-48.
- [18] Cairns.A.J.G, Blake.D, Dowd.K, Coughlan.G.D, Epstein.D, Ong.A, Balevich. I. A quantitative comparison of stochastic mortality models using data from England and Wales and the United States[R]. Edinburgh:Working paper, Heriot-Watt University and Pensions Institute Discussion Paper, 2007(10).
- [19] Grdada, C. O. The Great Famine in China, 1958-1962: A Documentary History. *Popul Dev Rev*2013(39), 333-346
- [20] Zhao, Z. W. & Reimondos, A. The Demography of China's 1958-61 Famine: A Closer Examination. *Population* .2012(67), 329-356 .
- [21] Moyn, S. Mao's Great Famine The History of China's Most Devastating Catastrophe, 1958-1962. *Nation* 2012(295), 32-36 .
- [22] Edgerton-Tarpley, K. Eating Bitterness: New Perspectives on China's Great Leap Forward and Famine. *J Asian Stud* 2012(71), 529-531.
- [23] Sorace, C. Catastrophe and Contention in Rural China: Mao's Great Leap Forward-Famine and the Origin of Righteous Resistance in Da Fo Village. *J East Asian Stud* 2011(11), 163-165.
- [24] Chen, Q. B. Excessive Investment, Compulsory Saving, and China's Great Famine of 1959-1961. *J Contemp China* 2011(20), 849-860.
- [25] Song, S. G. Mortality consequences of the 1959-1961 Great Leap Forward famine in China: Debilitation, selection, and mortality crossovers. *Soc Sci Med* 2010(71), 551-558.