Team Control Number

201931497

Problem Chosen

A

2019

ShuWei Cup

Summary Sheet

(Your team's summary should be included as the first page of your electronic submission.)

Current status and optimization of medical system

Summary

In the era of rapid development of society, aging problem and medical needs have become the social focus. This paper predicted the aging, medical needs and major diseases by establishing different mathematical models, and analyzed the competition and cooperation strategies between public hospitals and private hospitals, finally gave reasonable advice.

The first step is to predict the degree of aging and medical needs, we use the proportion of the elderly population to the national population to measure the aging and establish a grey prediction model, which is reasonable after passing the residual test and level deviation value test. We predicted that the degree of aging in China in the next five years will be: 0.118, 0.121, 0.125, 0.129, 0.133.

For the prediction of medical needs, we analyzed the impact of five indicators on medical needs. After detecting the multiple contributions and variable significance test of the model, we excluded part of the variables. Then a multiple linear regression model is obtained. It is estimated that the per capita health expenditure will reach 4416.3 yuan in the future.

The second step is to predict the most important diseases in the future. We collected data on eight major diseases including cancer and cardiovascular disease in Beijing and established polynomial fitting model. We found that heart disease is the most important disease in the future, which is estimated that the number of patients will reach 408,500 in 2019.

The third step is to establish a queuing theory model. We simplified the complex Markov queuing network into an $M/M/S/\infty$ model. Then we designed the relevant experiments to simulate and compared the performance of the hospital patient random queuing system and the shortest queue queuing system. As a result, it was found that queuing according to the shortest queue can greatly reduce the queuing time and the average queue length.

The final step, we taking into account their respective advantages and disadvantages, we cluster hospitals in Chengdu by K-means algorithm, then divide the hospital into three hospital groups. The hospitals in each group can cooperate with each other and benefit each other.

Key word: *Grey model, multiple linear regression, queuing theory model, K-means algorithm*

Content

1. Introduction 1
1.1 Background 1
1.2 Work 1
2. Problem analysis 1
2.1 Analysis of question one 1
2.2 Analysis of question two 2
2.3 Analysis of question three 2
2.4 Analysis of question four 2
3. Symbol and Assumptions 2
3.1 Symbol Description 2
3.2 Fundamental assumptions 3
4. Model 3
4.1 Prediction ageing trends and residents' medical needs 3
4.2 The prediction of Beijing's future major disease 8
4.3 Markov Queueing Network Model 11
4.4 Competitive partnership between public hospitals and private hospitals 15
5. Sensitivity Analysis 17
5.1 Sensitivity analysis of development coefficient 18
5.2 Sensitivity analysis of Gray efficiency 18
5.3 Sensitivity analysis of Arrival rate and Service rate 18
6.Strengths and Weakness 19
6.1 The advantages of the model 19
6.2 The shortcomings of the model 19
7. **Conclusion** 19
References 20
Appendix 22

1. Introduction

1.1 Background

With the acceleration of China's aging process and the rapid development of the economy, the medical problem has become a social hot issue that cannot be ignored. China's current real medical needs and potential medical needs in the future are staggering. This undoubtedly makes medical resources increasingly tense. Using reasonable methods to accurately predict the medical needs of our residents is the first step that formulation of national policies and the resolution of China's medical resource allocation.

Since 2009, the Central Committee of the Communist Party of China has introduced a series of new medical reforms to further guide and standardize the development of social capital hospitals and private hospitals. However, private hospitals still have a large gap with public hospitals in terms of access, operation, talent and scientific research. At the same time, public hospitals face the real problem of bed tension and difficulty in coping with medical peak demand, which undoubtedly makes the competition and cooperation between public hospitals and private hospitals a breakthrough point.

1.2 Work

Based on the above background and analysis of the topic, the following issues need to be addressed:

Task 1: Based on the data of residents' income, population age structure and economic development level in the relevant statistical analysis data of the National Bureau of Statistics, establish corresponding mathematical models to make a reasonable prediction of China's aging trend and residents' medical needs.

Task 2: Select a province and use mathematical models to analyze the most common diseases in the province in the future, with a view to recommending the overall development of key public hospitals in the province.

Task 3: Propose a common queuing theory and its associated optimal queuing method to provide solutions for real problems in hospitals

Task 4: In the context of complex cooperation and competition between public hospitals and private hospitals, propose an optimal cooperation and competition strategy among multiple hospitals.

Task 5: Write a 1-2 page proposal to the relevant medical authority

2. Problem analysis

2.1 Analysis of question one

Question 1 requires a reasonable predictions of the aging trend of China and residents' medical needs according on relevant statistics such as residents' income, population age structure, and economic development level. We should first select indicators to quantify the aging trend and the medical needs of residents. Because of the ambiguity of the factors affecting aging, we use gray prediction methods. in order not to miss the important factors, we use the Multiple linear regression models for prediction.

2.2 Analysis of question two

Question 2 requires us to analyze the most common disease in a certain province in the future and making recommendations for public hospitals. Firstly, it is necessary to find the disease incidence rate data of a certain province. It is obvious that this is a time series data, so the disease data with higher incidence rate is screened, and then we use polynomial fitting for prediction.

2.3 Analysis of question three

Question 3 requires the use of queuing theory algorithm to propose the best queuing method. We need to select the appropriate queuing theory based on the characteristics of queuing at the time of hospital inspection, and then calculate the queuing method based on the obtained results.

2.4 Analysis of question four

Question 4 requires us to propose competition and cooperation strategies between public hospitals and private hospitals. First of all, it is necessary to consult the data to understand the advantages and disadvantages of public hospitals and private hospitals. On this basis, the relevant competition and cooperation strategies are proposed based on the clustering model.

3. Symbol and Assumptions

3.1 Symbol Description

Symbol	Definition	Units
a	Development coefficient	—
b	Gray effect	—
ε	Relative residual	—
λ	Level ratio	—
y	per capita health expend	Yuan
x_1	GDP	Billon

ρ	Ratio deviation	—
λ_n	Arrival rate	People/min
μ	Server rate	People/min
L_q	Average queue length	People
W_q	Average waiting time	min

3.2 Fundamental assumptions

1. Hypothetical conditions for multiple linear regression.(1) Zero mean: $E(\boldsymbol{u}) = \boldsymbol{0}$;(2) Same variance and no autocorrelation: $Cov(u_{i,}u_j) = \sigma^2$;(3) Normality: $u_i \sim N(0,\sigma^2)$
2. Assume that the patient arrives and leaves the Poisson distribution;
3. Assume that the patient meets the principle of first-come, first-served;
4. Ignoring the output stream of one service station will become the input stream of another service station, and the arrival of each service station still meets the Poisson distribution.

4. Model

4.1 Prediction ageing trends and residents' medical needs

In this question, it is required to make a reasonable prediction of China's aging trend and residents' medical needs according to the data of residents' income, population age structure and economic development level in the relevant statistical analysis data of the National Bureau of Statistics. First, the proportion of the number of elderly people in the total number of people reflects the degree of aging, and the gray model GM (1, 1) is used to predict the trend of aging. Second, we choose multiple regression method to predict medical needs.

4.1.1 Establishment of grey prediction model

(1) Model establishment

For non-stationary stochastic processes, Models based on static analysis methods may lose practical significance since make the confidence interval of prediction too large. The grey model can generate a dynamic or non-dynamic white module of a known data sequence according to a certain plan, and develop or find its inherent law from the chaotic raw data to predict future values. For the prediction of population aging trend, through the existing census data, a grey prediction model is established to find out the evolution law of the inherent aging trend, and then the future aging trend can be predicted. To this end, the model will be built below.

Normally, After the random non-negative sequence has been accumulated for many times, most of them can be approximated by an exponential curve. Now establish a whitening gray prediction model with m=1 column, whitening GM (1,1) differential equation[1]:

$$\frac{dx^{(1)}}{dt}+ax^{(1)}=b \tag{1}$$

In the middle, a is the development coefficient; b is the gray effect

According to the principle of least squares, Find the parameters a 、 b :

$$a=\frac{CD-(n-1)E}{(n-1)F-C^2},\ b=\frac{DF-CE}{(n-1)F-C^2} \tag{2}$$

$$C=\sum_{k=2}^{n} z^{(1)}(k);\quad D=\sum_{k=2}^{n} x^{(0)}(k);\quad E=\sum_{k=2}^{n} z^{(1)}(k)x^{(0)}(k);\quad F=\sum_{k=2}^{n} z^{(1)}(k)^2$$

A predictive model can be obtained by solving differential equations:

$$\hat{x}^{(1)}(k+1)=\left[x^{(0)}(1)-\frac{b}{a}\right]e^{-ak}+\frac{b}{a}\ (k=0,1,2...,n) \tag{3}$$

Restore the model to get the predicted value:

$$\begin{cases} \hat{x}^{(0)}(1) = x^{(1)}(1) = x^{(0)}(0) \\ \hat{x}^{(0)}(k+1) = x^{(1)}(k+1) - x^{(1)}(k) \end{cases} \tag{4}$$

(2) Data source and indicator definition

The data of this question are all from the National Bureau of Statistics' census and statistical bulletin. According to international definitions, when a country or region with an elderly population aged 65 and over accounts for more than 7% of the total population, it means that the country or region is ageing. Based on this, this paper will measure the aging trend by the proportion of aging.

(3) Model test

①Residual test: calculate relative residuals

$$\varepsilon(k)=\frac{x^{(0)}(k)-\hat{x}^{(0)}(k)}{x^{(0)}(k)},k=1,2,...,n \tag{5}$$

If $|\varepsilon(k)|<0.05$, it is considered that a higher requirement is reached; $|\varepsilon(k)|<0.1$, it is considered to meet the general requirement.

②Level deviation value test: calculation Level deviation value

$$\lambda(k)=\frac{x^{(0)}(k-1)}{x^{(0)}(k)},k=1,2,...,n\ ,\ \rho(k)=1-\frac{1-0.5a}{1+0.5a}\lambda(k) \tag{6}$$

If $|\rho(k)|<0.01$, it is considered that a higher requirement is reached; $|\rho(k)|<0.05$, it is considered to meet the general requirement.

③Test results and analysis

The data of 1998-2018 is substituted into the gray GM (1,1) model established above to obtain the model value, the relative residual value and the step deviation value:

Table1: GM (1,1) prediction model residual test and grade ratio deviation test

Year	1998	1999	2000	...	2016	2017	2018
Original A value	0.0671	0.0690	0.0696	...	0.1085	0.1139	0.1194
Model A value	0.0671	0.0651	0.0671	...	0.1078	0.1110	0.1143

Residual	0	0.0039	0.0025	...	0.0007	0.0029	0.0051
Relative residual	0	5.652%	3.592%	...	0.645%	2.546%	4.271%
Ratio deviation	0.027%	0.009%	0.019%	...	0.035%	0.047%	0.046%

It can be seen from the above table that the GM (1,1) prediction model established by using the data from 1998 to 2018 has reached a higher requirement through the relative residual test, and is far higher than the higher standard for the ratio test. It can be seen that the GM (1,1) prediction model has higher model accuracy and reliability, and can predict the next population aging trend.

(4) Forecast results and analysis

Based on the gray prediction GM (1,1) model established above, we will predict the aging trend from 2019 to 2023. The forecast results are as follows:

Table2: Prediction results using GM (1,1) prediction model

Year	2019	2020	2021	2022	2023
Predicted ageing ratio	0.1178	0.1213	0.1249	0.1287	0.1326
growth rate	3.06%	2.97%	2.97%	3.04%	3.03%

It can be seen from Table 2 that the aging ratio of the next five years predicted by the GM (1,1) prediction model has reached the international definition of aging, that is, the ratio is higher than 7% for aging. The growth trend shows that the aging is showing a relatively positive growth, and the growth rate is maintained between 2.97% and 3.03%. It is predicted that by 2023, the number of elderly people in China will account for 13.26% of the national population, undoubtedly, a steady growth trend will be a tricky business. In order to more intuitively show the trend of aging, the graph of the aging trend will be given below:

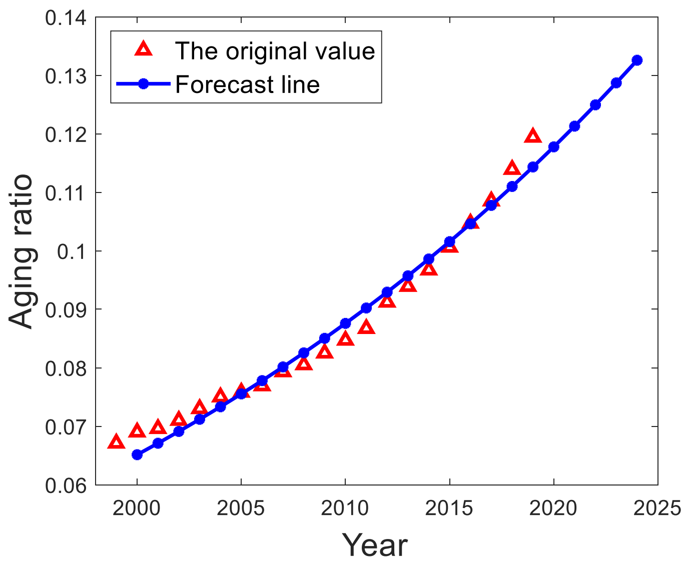


Figure1: GM (1,1) model predicts the trend of aging

It can be seen intuitively from the figure that the proportion of China's aging population to the total population of the country, that is, the trend of population aging is increasing, and the growth rate is relatively stable. Combined with the reality, it is more objective and reasonable, and the forecast is in line with the reality.

4.1.2 Multiple Linear Regression Model for Forecasting Medical needs

(1) Explanatory variables and the interpreted variables

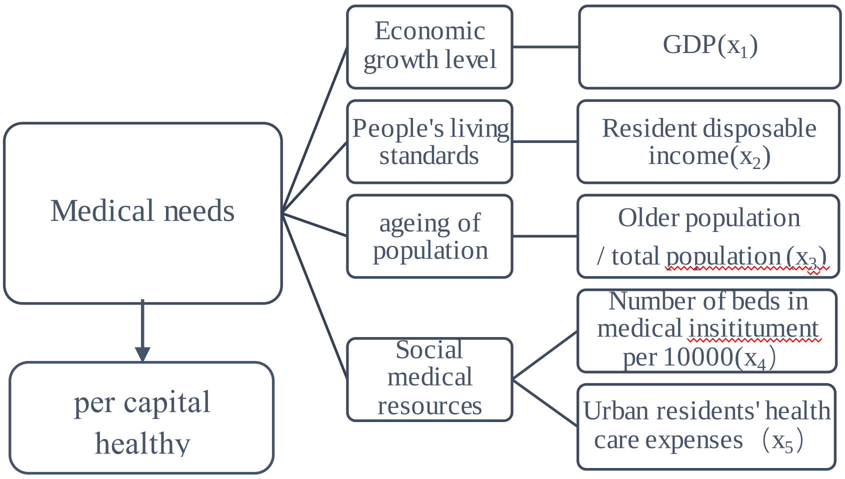


Figure2: Selection of explanatory variables

In order to fully reflect the medical needs of Chinese residents, we select the indicators shown in the above figure to measure the medical needs and the influencing factors of residents' medical needs, and use y and x to represent different variables.

(2) Data source

We use relevant data published in “China Statistical Yearbook” on population, national economic accounting, health, and people's lives.

(3) Establishment of multiple linear regression model

For a multiple linear regression model with k-1 explanatory variables, the general form is:

$$y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + u_i \tag{7}$$

The parameters $\beta_i (j = 1, 2, ..., k)$ in the model are partial regression coefficients， the sample size is n.

We use the OLS method to estimate and seek the minimum square sum and minimum：

$$\min \sum e_i^2 = \sum (Y_i - \hat{Y}_i)^2$$

$$\min \sum e_i^2 = \sum [Y_i - (\hat{\beta}_1 + \hat{\beta}_2 X_{2i} + \hat{\beta}_3 X_{3i} + ... + \hat{\beta}_k X_{ki})]^2$$

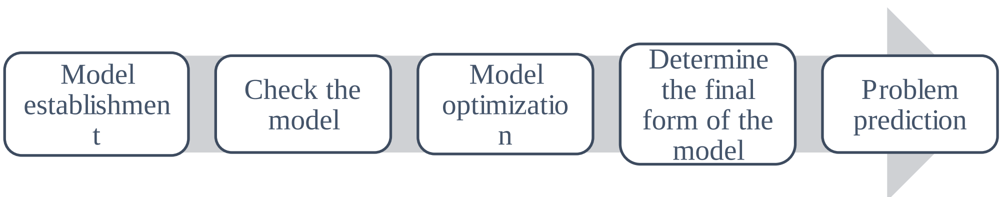


Figure3: Flow chart for establishing the model

The data used in this question has a common time trend. In order to eliminate the influence of time trend and consider the economic significance represented by the selected indicators of the model, we take the logarithm of each variable directly and add the time variable T. The least square method is used to establish multiple regression model.

Draw a scatter plot: Assume that the explanatory variable is linearly related to the interpreted variable, and draw a scatter plot between the interpreted variable $\ln y$ and the explanatory variable $\ln x_1$, $\ln x_2$, $\ln x_3$, $\ln x_4$, $\ln x_5$. (The graphics are shown in the Appendix as Figure 1.1)

As can be seen from the figure, there is a linear relationship between the explanatory variable and the interpreted variable.

A multiple linear regression model was established for the interpreted variable $\ln y$ and each explanatory variable（x_1, x_2, x_3, x_4, x_5） as well as the time variable T.

Table 3.1: The results of Stepwise linear regression

Variable	coeffcient	Std.	t-Statistic	Prob
C	-2.731642	3.836895	-0.711941	0.5279
$\ln x_1$	0.0289356	0.244443	1.183731	0.3218
$\ln x_2$	0.344148	0.375095	-0.917494	0.4265
$\ln x_3$	-0.347239	0.154151	1.902644	0.1532
$\ln x_4$	1.893033	0.905871	-0.383320	0.7270
$\ln x_5$	1.877925	0.534151	3.515719	0.0390
T	0.091057	0.074409	1.223747	0.3457

Table 3.2: The results of Model evaluation

Important Indicators	
R-Squared 0.993548	Adjusted R-squared 0.993168
F-statistic 2617.119	Prob(F-statistic) 0.00000

It can be seen from the P value presented by the regression results that the explanatory variables are not significant, however, the value of R^2 is close to 1. Therefore, there may be multiple collinear problems between explanatory variables in the model.

Then we use stepwise regression to eliminate the effects of multiple collinearity. When the first four variables are step wisely regressed, the variables are significant, and the fitting effect is good. After the addition, all the variables are not significant and R^2 is close to1, It can be judged that serious multi-collinearity occurs, considering that the important variables will not be affected after the rejection, so we use the first four explanatory variables to build the model.(See the appendix for the regression process as 1.2)

Table 4.1: The results of multiple linear regression

Variable	coeffcient	Std.	t-Statistic	Prob
C	-4638.785	331.6129	-13.98855	0.0000
$\ln x_1$	0.000947	0.000375	2.522363	0.0244
$\ln x_2$	0.015478	0.002547	2.666431	0.0447
$\ln x_3$	0.563895	0.041627	13.546440	0.0000
$\ln x_4$	1.893033	0.238118	7.9499520	0.0005

T	-92.51352	7.733920	11.962050	0.0000

Table 4.2: The results of Model evaluation

Important Indicators	
R-Squared 0.963874	Adjusted R-squared 0.973168
F-statistic 6135.503	Prob(F-statistic) 0.00000

The regression results after the variables were removed showed that the P values of each variable were significant, the R^2 showed that the degree of fitting was better,

and the F variable was very significant, so the equation was obtained:

$$\ln y = -4638.785 + 0.000947\ln x_1 + 0.015478\ln x_2 + 0.563895\ln x_3 + 1.893033\ln x_4 - 92.51352T$$

(4) Medical needs prediction

the data on per capita health expenditure in 2018 has not been published, so we use the data of the explanatory variables in 2018 to predict the medical needs of residents in 2018, and predict the future years based on the trend of regression images.

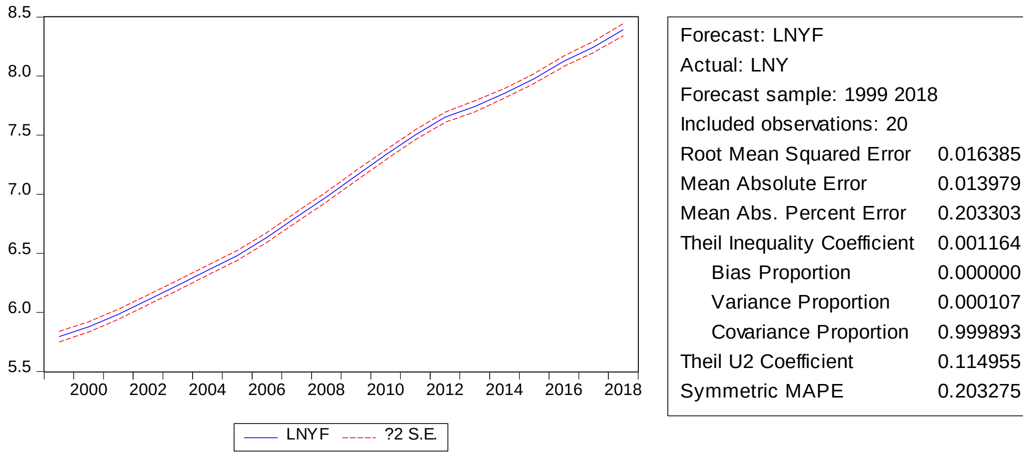


Figure4: Result of prediction

it's easy to see： $\ln Y_F = 8.392446$, Y_F =4413.60 , Y_F indicates the per capita health expenditure in 2018, and $\ln Y_F$ is the logarithm of per capita health expenditure predicted by the established multiple linear regression model.

According to the established equation, the coefficients of all explanatory variables are positive, which indicates that the medical needs have the same trend as the selected indicators. Whenever GDP increases by 0.000947%, the personal hygiene costs of residents will increase by 1%, as will the impact of the remaining variables. According to the conclusion of the above questions, China's aging population will continue to increase. At the same time, China's economy will maintain its growth state. Therefore, medical demand will continue to grow in the future. It is estimated that the per capita health expenditure may reach 4,500 yuan in the future. It is particularly important to optimize the medical service system.

4.2 The prediction of Beijing's future major disease

The second problem is mainly the prediction of future diseases in a certain province， we chose Beijing as the research object. We got data on AIDS, malignant tumor/cancer, cardiovascular disease and cerebrovascular disease, lung cancer, thyroid cancer, colorectal cancer, and liver cancer through "Beijing Municipal Healthy Commission Information

Center"[2]. Next we will use a simple and efficient polynomial fitting model for analysis and prediction.

4.2.1 Establishment of least squares polynomial fitting model

The general form of the fitting polynomial:

$$y = a_0 + a_1 x + a_2 x^2 + \cdots + a_k x^k \tag{8}$$

y is approximated by a linear combination of coefficients of a k-th order polynomial in the above formula.

In order to study some complex functions, we usually use high-order polynomials to approximate representation function:

$$f(x) \approx p_n(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + ... + a_n(x - x_0)^n$$

Note: Here, a linear combination of coefficients of the nth-order polynomial is used to approximate representation f(x).

The square of the result of the polynomial function is subtracted from the true function value of f(x) to represent the error relationship between f(x) and the polynomial function:

$$Loss = \sum_{i=1}^{n}\left(f(x_i) - \left[a_0 + a_1(x_i - x_0) + a_2(x_i - x_0)^2 + ... + a_n(xi - x_0)^n\right]\right)^2 \tag{9}$$

The principle of fitting is to make the loss function as small as possible. In other words, Fit the data to a polynomial $y = a_0 + a_1 x + a_2 x^2 + \cdots + a_k x^k$ to minimize the error.

Then uses RMSE(Mean square), SSE（Variance of sum）, and R-square (Coefficient of determination) to test the degree of fit.

4.2.2 Disease prediction in Beijing

First, based on the collected data, we found that the number of patients with heart disease and cancer far more than the number of other diseases, so we will mainly predict the incidence of these two diseases, considering data acquisition of cancer is difficult and the amount of the data is small, so we have collected four common cancers: lung cancer, thyroid cancer, colorectal cancer, and liver cancer.

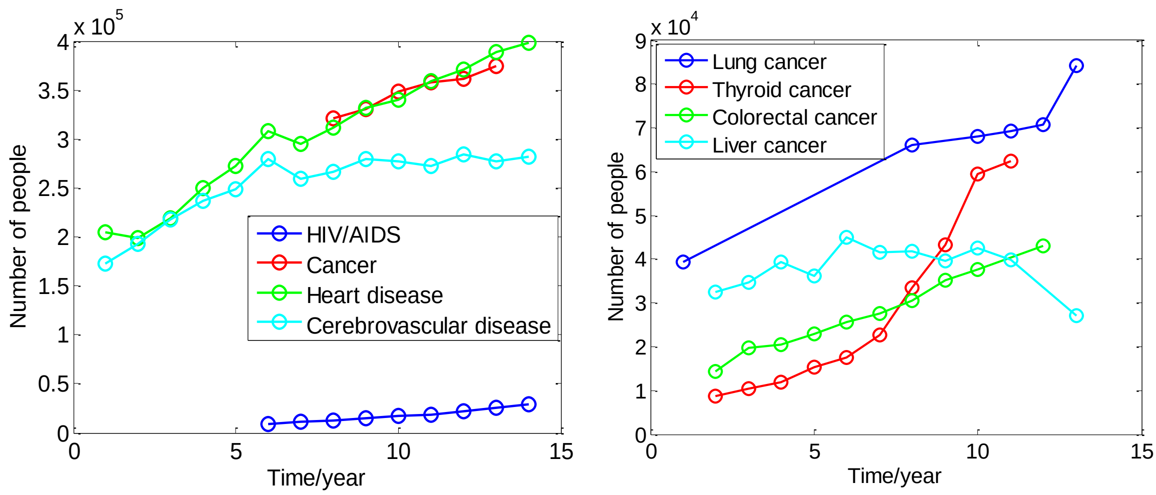


Figure5: Statistics on the number of diseases in Beijing

As shown in the above figure, we found that the proportion of patients with heart disease and cancer is high. Considering the less cancer data, we analyze the specific cancer. Next we use the least squares fitting to Model.

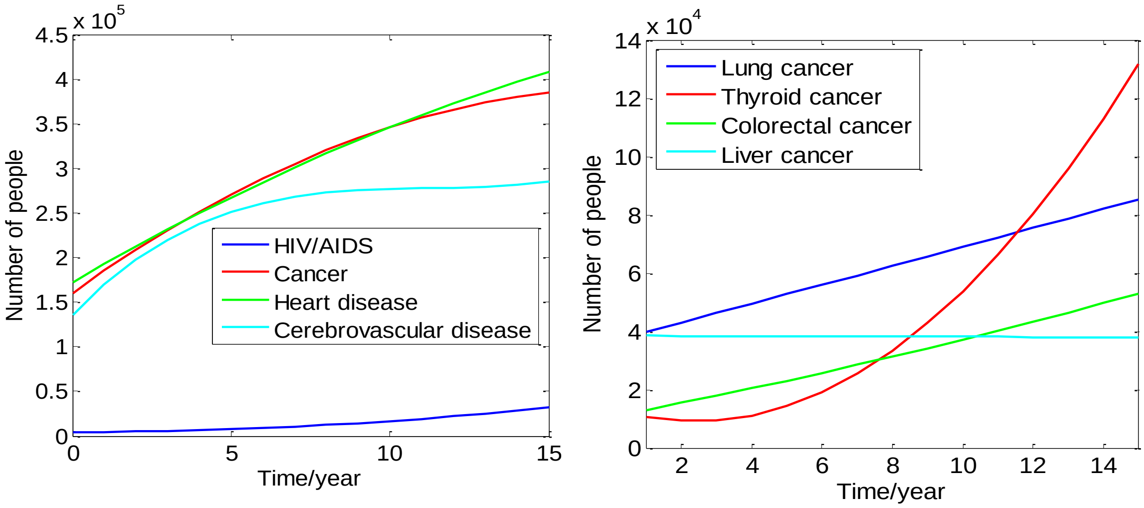


Figure 6：Beijing disease population prediction model

Next we will give a fitting model and prediction of the disease in 2019. **The solution results and test results are shown in Appendix Table 2.1.**

Based on the above calculations, we found that patients with heart disease have increased at a faster rate in recent years. It is expected that the number of patients in Beijing will reach 408,500 in 2019, and the prevalence of cancer is known to be high. We collected four common cancers, such as lung cancer and liver cancer, and performed polynomial fitting in case the cancer data is difficult to obtain. We found that thyroid cancer has grown rapidly in cancer in recent years. It is expected to reach 130,000 in 2019, which is about 34% of the total number of cancer patients. Simultaneously, lung cancer will also reach 85,000, accounting for about 22% of the total number of cancer patients. According to the research results, we propose the following two suggestions:

① Use the Internet to build a medical knowledge science platform, advocate a healthy lifestyle, promote cancer-related knowledge, and emphasize the importance of disease prevention.

②Apply the concept of "Internet + medical health", advocate the integration of medical technology and emerging industry technology, increase investment in drug research and development, cultivate innovative medical talents, and improve the efficiency of treating diseases.

4.3 Markov Queueing Network Model

Different types of patients may need to do different medical checking in the hospital, and different medical checking are distributed in different locations in the hospital. This problem requires us to find an optimal queuing scheme which allows the patient can be given priority to go to the shorter position of the captain to do checks. This scheme is designed to enhance the service efficiency of the entire system and reduce the patient's queuing time.

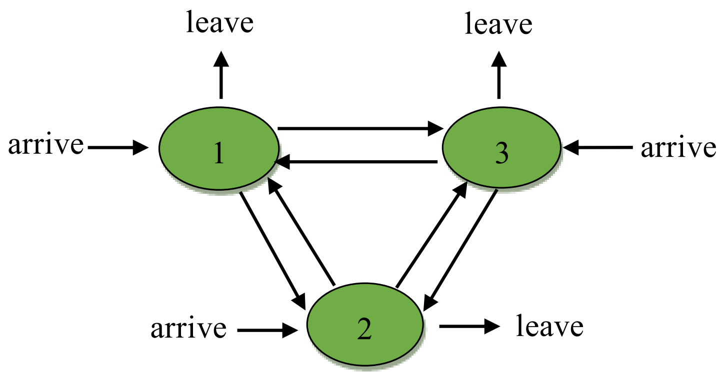


Figure 7: Schematic diagram of hospital patient queuing theory model

As shown on the left of Figure 1, different service stations are connected to each other. The patient output stream of a service station will enter as the input stream of the next service station. In order to obtain multiple services, patients need to queue up for multiple service stations, and there are certain links between different service stations. This problem belongs to the Markov queuing network model. Considering the complexity of solving the problem, let's simplify the problem below.

1.The patient output stream of a service station may enter the next service station, or it may exit the queuing system directly after medical checking. In order to facilitate the solution of the problem, we consider the input stream of each service station to be generated by the Poisson distribution of its own queue, ignoring the impact of the output of the previous service station on the next service station.

2. Different types of patients often have different medical checking. In reality, the checking often includes routine checking, such as blood test, height and weight. Surely, it also includes some special checking, such as B-ultrasound, color ultrasound and so on. Therefore, the arrival rate of patients at different service stations is often different. In this problem, we ignore the impact of different patient arrival rates at different service stations.

3. Different checking have different service rates. Similar to the above, the service rates of different checking are often different.

4.Different types of patients often do several specified medical checking. In this problem, we assume that all patients need to have all the medical checking.

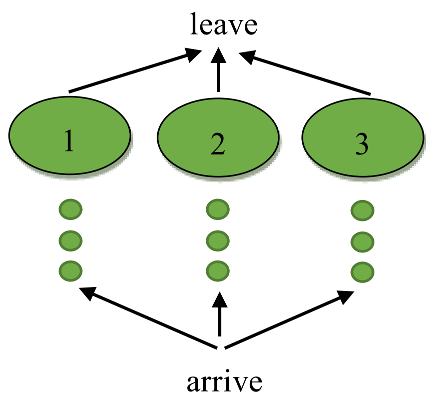


Figure 8: Schematic diagram of simplified model of hospital patient queuing theory

Based on the above assumptions, we can define a complex system as a M/M/S/ ∞ queue theory model. After that, we will mainly do related simulation calculations on the two schemes of patient random queuing and lining up in the shortest queue, and then determine a better queuing scheme.

4.3.1 M/M/S/ ∞ queuing theory model establishment[3][4]

The common queuing system is mainly aimed at single business queuing and multiple service stations handling the same business. The queuing problem here is more complicated, as shown on Figure 7.

Given by the interaction between different services of the Markov queue network model, the solution is more complicated. We assume that all patients need to have all the medical checking, ignoring the impact of the output stream of one service station becoming the input stream of another service station. On this basis, the arrival of each service station patient still meets the Poisson distribution. In this way, the problem is reduced to the multi-service station queuing model M/M/S/ ∞.

After the system is in a stable state, the probability distribution of long N is as follows:

$$\lambda_n = \lambda$$

The system service rate is:

$$\mu = \begin{cases} n\mu & (n=1,2\cdots s) \\ s\mu & (s = s, s+1, s+2\cdots) \end{cases}$$

The system service strength is:

$$\rho_s = \frac{\rho}{s} = \frac{\lambda}{s\mu}$$

When $\rho_s < 1$, the queue length is not getting longer and longer, and the state transition probability of this system is:

$$P_n = \begin{cases} \dfrac{\rho^n}{n!} P_0, & \text{n=1,2,3}\cdots, s \\ \dfrac{\rho^n}{s!s^{n-s}} P_0, & \text{n} \geq s \end{cases}$$

Particularly: $P_0 = \left[\sum_{n=0}^{s-1} \frac{\rho^n}{n!} + \frac{\rho^s}{s!(1-\rho^s)} \right]$, P_0 indicates the probability that the entire system is idle.

When n $\geq s$, the customer needs to wait, and the probability of waiting is:

$$c(s,\rho) = \sum_{n=s}^{\infty} P_n = \frac{\rho^s}{s!(1-\rho^s)} P_0$$

The main system operating indicators we selected were the average queue length and average waiting time.

The average queue length is:

$$L_q = \sum_{n=s+1}^{\infty} (n-s) P_n = \frac{P_0 \rho^s \rho_s}{s!(1-\rho_s)^2} \tag{10}$$

The average waiting time is:

$$W_q = \frac{L_q}{\lambda} = W - \frac{1}{\mu} \tag{11}$$

4.3.2 Solution and Simulation of Queuing Model

For this problem, we made some assumptions to convert complex Markov network queuing into a simple multiple service station queuing. Due to the long distance between different service stations, patients cannot clearly know which queue is shorter. Therefore, patients choosing queue is random. Based on the above queuing theory, we mainly analyze whether the service effect of the whole system is enhanced if the patients choose a shorter queue.

We assume that the hospital has 5 service stations, and the total patient arrival rate is $\lambda = 5 people/\min$ .Besides, we also assume that the service rate per window is $\mu_i = 1.5 people/\min$. From the above assumption, we calculated that the average queue length is: $L_q = 1.75\ people$,and the average waiting time is: $W_q = 2.62\ \min$.

Then, we will simulate and calculate the two modes of random queuing and lining up in the shortest queue, compared the average queue length and the average waiting time of patients under the two modes.

①Random queuing: Each patient randomly choose a queue to line up after arrival.

②Lining up in the shortest queue: Each patient can choose the shortest queue to line up after arrival.

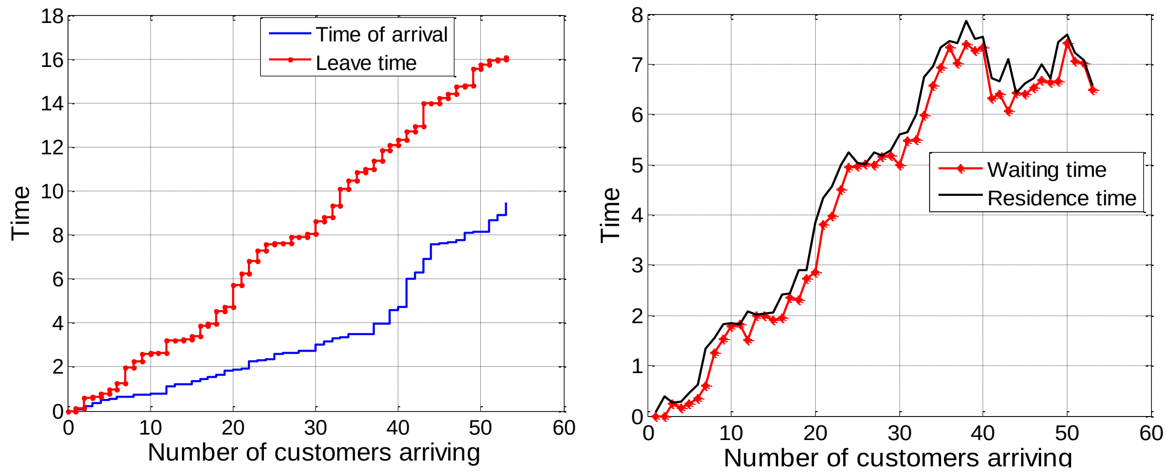


Figure 9: Curve of time of arrival, leave time, waiting time and residence time for each patient under random queuing

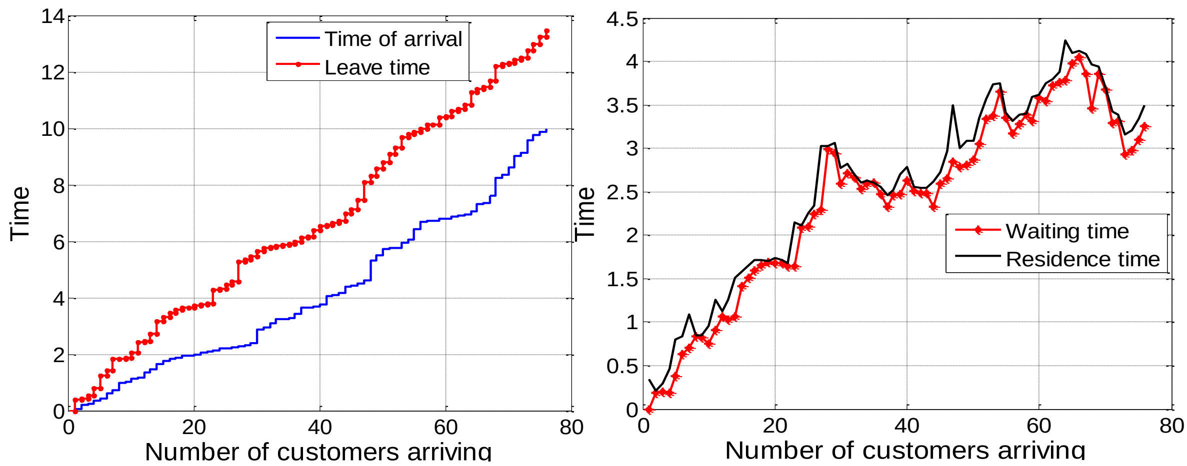


Figure 10: Curve of time of arrival, leave time, waiting time and residence time for each patient under lining up in the shortest queue

Table 5: System performance indicators of different queuing modes

queuing modes	the average waiting time /min	the average residence time /min
Random queuing	4.52	4.74
Lining up in the shortest queue	2.49	2.56

From the above table, we can clearly know that when patient randomly choose a queue to line up after arrival, patients may often be randomly queued to a longer service station, which leads to an increase in patient waiting time. In contrast, if the optimal queuing position is calculated, the average waiting time of the patient can be drastically reduced.

Different checkpoints in hospitals are often distributed in different places. Patients usually randomly select a place to queue, and the remaining checkpoints may be idle, resulting in waste of system resources. In summary, we believe that it is necessary to set up a mechanism to calculate the shortest queue, so that customers can give priority to relevant inspections, so that the reasonable scheduling of the entire system resources can also reduce the patient's time of treatment. In this way, reasonable scheduling of the entire system resources can be achieved, and the patient's medical time can also be reduced.

4.4 Competitive partnership between public hospitals and private hospitals

4.4.1 Competition and cooperation between hospitals

In view of the current problem of patients in large public hospitals, such as "difficult to see a doctor, hospitalization is difficult, surgery is difficult", the cooperation between public hospitals and private hospitals is of great significance. Because private hospitals cannot compete with public hospitals in terms of technology and talent credibility, resulting in small outpatients、 vacant beds and unnecessary waste of medical resources. Public hospitals rely on the government's strong capital investment, insufficient beds, doctors' appointments and other issues often occur[5].

Public hospitals and private hospitals have the following main advantages:

①Public hospitals have better equipment、 strong medical technology and medical teams;

②Private hospitals have more beds that can be controlled and have a small flow of people;

In this question, we consider the distribution of hospitals in Chengdu, China, and cluster small hospitals according to the European distance to form a number of reasonable hospital clusters. Private hospitals rely on public hospitals that are close to each other for cooperation. They can provide beds, surgery and other resources to public hospitals, and learn advanced technology from them to achieve win-win cooperation.

4.4.2 Hospital cluster division based on K-Means algorithm

In this part, we take Chengdu as an example to collect the large and small hospitals in Chengdu, then we take the Chengdu “Wuhou Leather City” as the coordinate origin to establish the coordinate system. As shown in the map on the left side of Figure 3, the red mark identifies a large public hospital, and the red dot indicates small private hospitals, the following will focus on clustering all hospitals.

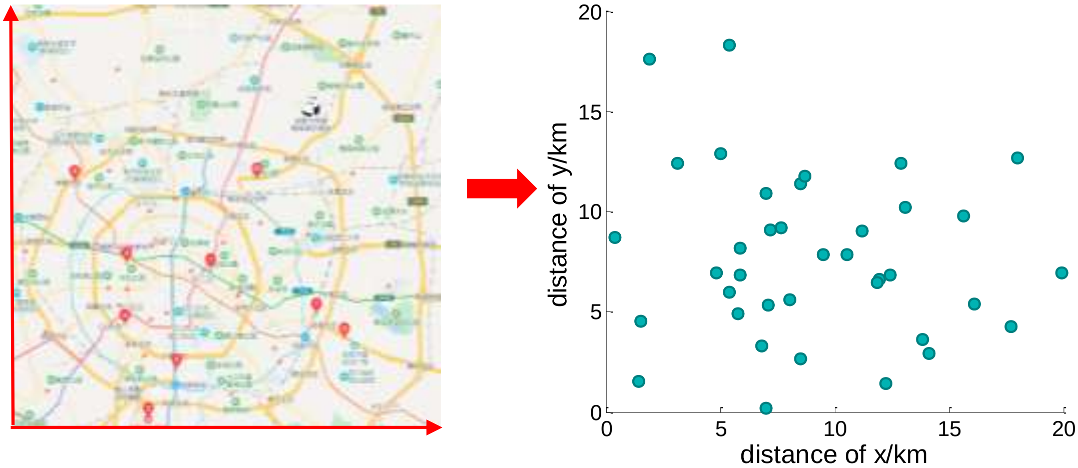


Figure 11: Distribution map of Chengdu hospital

K-Means clustering belongs to the unsupervised machine learning algorithm[6]. All the known data are data without split labels. The goal of clustering is to divide n data into K sets, so that WCSS (within-cluster sum of squares) is the smallest, that is, needs to meet:

$$\arg\min \sum_{i=1}^{k} \sum_{x \in S_i} \| x - \mu_i \|$$

In the above formula, S_i represents a data set belonging to class i, and μ_i represents a cluster center point.

First, we randomly select K cluster centers. In each iteration, we first determine the nearest cluster center for each. This step is called cluster assignment, and then calculates the center of all points in each class. Moving the cluster center of the class to the center point is the step which called move centroid, getting the new position of the k cluster centers, and proceed to the next iteration until the cluster center point is correctly assigned in the center of each class. The specific process is as follows:

Algorithm: K-Means algorithm

Randomly initialize K cluster centers $\mu_1, \mu_2, \cdots, \mu_k$

Iterate through this process until $\mu_1, \mu_2, \cdots, \mu_k$ no longer changes

for $i = 1:k$

Solve the distance between i and each cluster center

Divide i into a center u closest to the distance i in the cluster center

for $k = 1:K$

Average the points that are all divided into cluster centers (update $\mu_1, \mu_2, \cdots, \mu_k$)

In the clustering process, the Euclidean distance is used for calculation. The calculation formula is as follows:

$$dis = \sqrt{(x - x_0)^2 + (y - y_0)^2}$$

x represents the geographic coordinates of each hospital and y represents the coordinates of each cluster center point.

We set up three cluster centers. Finally, we get the distribution of hospital ownership in this city. As shown in Figure 12, three colors represent three clusters of hospital clusters, and triangles mark the cluster center points of each cluster.

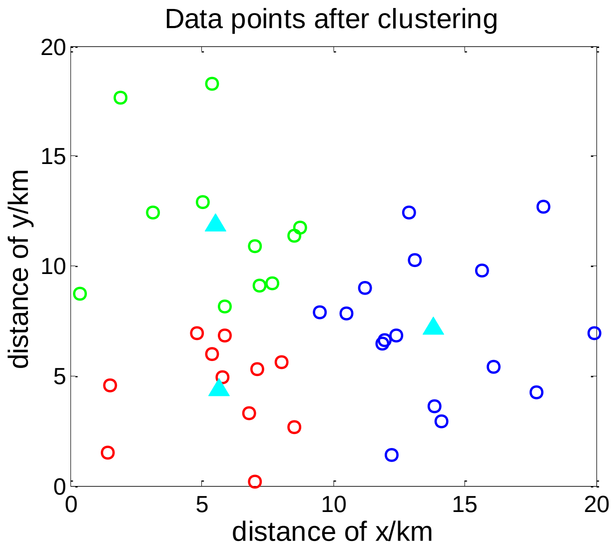


Figure12: Clustering of hospital distribution

Then we statistic the above clustering results, marking the specific number of public hospitals and private hospitals in each category, see Table 6.

Table 6: Statistics on the number of public hospitals and private hospitals

Class ID	The number of public hospital	The number of private hospital	Total Number
1	3	8	11
2	2	9	11
3	4	12	16

In the three categories formed by our clustering, relying on geographical location to form a hospital group based on public hospitals, In the process, private hospitals can provide some resources to support, and at the same time learn advanced management level and medical technology level, gradually enhance credibility, visibility and competitiveness of the private hospitals.

5. Sensitivity Analysis

Some of the variables in our model may be some deviations, which may affect the solution of our model. In view of the above, we implemented a sensitivity analysis to detect the impact of small changes in the variables on the results. For the grey prediction model, the development coefficient and the gray efficiency are the impact indicators. For the polynomial multiple regression model, the model can naturally have better robustness, and no test is performed here. For the queuing theory algorithm, the system's arrival rate and service rate may be great influence to our queuing system, we use them as impact indicators. The results show that our model does not show chaotic behavior and shows good sensitivity. Our sensitivity analysis will be based on a methods of controlling variables in order to understand how the results of the model change as the input parameters change. We will analyze the effect of the following parameter changes on the model:

1. Development coefficient of Grey prediction model
2. Gray efficiency of Grey prediction model
3. Arrival rate and Service rate of queuing system

5.1 Sensitivity analysis of development coefficient

Here we analyze the impact of the development coefficient on the whole model, we only use the 2019 aging degree to predict whether the output is stable.

Table 7: Sensitivity analysis of development coefficient

Development coefficient	-5%	-3%	-1%	0%	+1%	+3%	+5%
Year	0.1183	0.1182	0.1149	0.1178	0.1177	0.1175	0.1174

It can be seen from the above that when the independent variable produces a small range of fluctuations, the predicted value also remains within a certain range, and the model has better robustness. At the same time, we also find that the increase of the development coefficient can reduce the predicted value.

5.2 Sensitivity analysis of Gray efficiency

Similar to previous analysis, we predict the degree of aging in 2019.

Table 8: Gray efficiency of Grey prediction model

Gray efficiency	-5%	-3%	-1%	0%	+1%	+3%	+5%
Year	0.1170	0.1174	0.1176	0.1178	0.1180	0.1183	0.1185

It can be seen that the model is relatively stable and the increase of the development coefficient can increase the predicted value.

5.3 Sensitivity analysis of Arrival rate and Service rate

Here we only modify the value of the arrival rate and observe the impact on the queuing system.

Table 9: Arrival rate of queuing system

Arrival rate	-5%	-3%	-1%	0%	+1%	+3%	+5%
Average waiting time	2.25	2.32	2.4	2. 49	2.5	2.55	2.62
Average residence time	2.38	2.46	2.52	2. 56	2.68	2.74	2.85

It is not difficult to find that the average waiting time and the average residence time have small fluctuations with the change of the arrival rate. We can guess that the service rate has a similar impact, and no relevant analysis is done here.

From these three kinds of sensitivity analysis, we can know that development coefficient, gray efficiency and arrival rate and service rate also impact little on our model. In comparison, arrival rate has a greater impact on the average waiting time. This may be due to the fact that the average waiting time is calculated by software simulation and has certain randomness. It is concluded that the model has good robustness.

6.Strengths and Weakness

6.1 The advantages of the model

1. In the model, we use the gray model to predict the degree of aging of the population. The model itself does not require a large number of samples, and does not require the regular distribution of the sample itself. The accuracy of the prediction is high;

2. In the queuing theory model, we simplify the original complex queuing network model, compare the average waiting time and average queue length of the patients in the two queuing modes, and simulate according to the rules of the two modes, which is convincing.

3. For the competition and cooperation between public hospitals and private hospitals. The hospitals are divided into three hospital groups by clustering the hospitals in Chengdu, which can realize the rational sharing of resources and gradually improve the competitiveness of private hospitals.

6.2 The shortcomings of the model

1. When forecasting future medical needs, we selected five indicators for the four dimensions to solve the model. The factors that may be considered are not comprehensive enough, and some important indicators are ignored.

2. When predicting the main diseases in Beijing in the future, we use simple and efficient polynomial fitting. Due to the difficulty of collecting data, the calculation of the model through a small amount of data may have large calculation errors.

7. Conclusion

After we have established different models for each problem and analyzed the results, we conclude that:

1. In the next few years, the proportion of the elderly population in China will continue to increase steadily, and the growth rate will be stable at around 3%. At the same time, the medical needs of Chinese residents are also on the rise, and the per capita health expenditure of residents may exceed 5,000 yuan. Therefore, optimizing the allocation of medical resources and improving the efficiency of medical services are important topics in the reform of the medical system.

2. The incidence of heart disease and cancer in Beijing is higher than other diseases. It is estimated that in 2019, the number of patients with heart disease in Beijing will reach 408,500. At the same time, we also found that thyroid cancer is growing fastest. The thyroid cancer patients will reach 34% of all cancer patients in 2019, meanwhile lung cancer patients will also reach 22%.

3.We establish the $M/M/S/\infty$ model to simulate the hospital's queuing model. Through computer simulation, we compared the two modes of random queuing and shortest queue queuing. Finally, we found that the shortest queue queuing mode can effectively reduce the average queuing time and average queue length. In the future, we hope the relevant departments can develop an online queuing platform to improve the overall service efficiency of the hospital.

4.We use K-mean clustering to calculate the number of public hospitals and private hospitals in the cluster. We conclude that private hospitals can provide some medical resources support to hospital groups that are mainly public hospitals, and can learn their advanced management level and The level of medical technology, gradually improve the credibility and visibility. Public hospitals can take advantage of resources to support the development of private hospitals and ease the pressure on surgery.

References

[1] Yang hualong, liu jinxia, zheng bin. Improvement and application of grey prediction GM(1,1) model [J]. Practice and understanding of mathematics,2011,41(23):39-46.

[2] Beijing Municipal Healthy Commission Information Center. http://www.phic.org.cn/sy/

[3] Lu Xue. Network queuing simulation based on MATLAB [D]. Dalian Maritime University, 2013.

[4] Srivastava H M, Kashyap B R K. Special functions in queuing theory and related stochastic processes[J]. ACADEMIC PRESS, 111 FIFTH AVE., NEW YORK, NY 10003, 1982.

[5] Wu Zongyong. Research on cooperation mode between public hospitals and private hospitals[D]. Chinese Academy of Social Sciences, 2013.

[6] Wagstaff K, Cardie C, Rogers S, et al. Constrained k-means clustering with background knowledge[C]//Icml. 2001, 1: 577-584.

A letter to the National Health and Wellness Commission of the People's Republic of China

Dear Sir/Madam:

thank you for taking the time to read our letters. The content of this letter is mainly some of the problems we have discovered during the study of the medical problems. We hope this letter can bring some help.

Today there are two main problems in China's current medical system: (1) China's total medical and health resources are seriously insufficient. (2) Medical services are inefficient.

We use the gray prediction method GM (1,1) model to predict the degree of population aging, and found that the growth rate of the elderly population will remain at around 3% in the next five years. The trend of aging is obvious, and the burden on the medical system will be further aggravated. We also selected some indicators to predict per capita health expenditures. We found that the per capita health expenditure may reach 4,500 in the next few years. Therefore, optimizing medical resource allocation is very important.

in addition to this, we forecast the main diseases in Beijing in the future and find that disease patients in Beijing will reach 408,500 in 2019. At the same time, the incidence of thyroid cancer and lung cancer remains high. Relevant departments need to increase drug efforts, meanwhile, effectively use the online platform to publicize health knowledge.

Hospital service efficiency is very important. We carry out related simulations to compare two models, and finally find that the shortest queue method can effectively reduce the average queuing time and queue length. so we suggest that hospitals can use the online queuing system platform to allow patients to choose the shortest queue.

For the allocation of medical resources, we find that there are many beds in private hospitals, but the number of patients is small. The public hospitals are supported by the government and have strong financial strength. If they can take advantage of each other, the hospital's service efficiency will improve greatly. Therefore, it is necessary to establish a cluster model to establish a hospital cluster led by public hospitals to achieve mutual benefit and win-win results.

Of course, optimizing the allocation of medical resources and improving the efficiency of medical services is a long-term process. In addition to the above recommendations, we believe that upgrading medical technology is also an important measure to improve the quality of medical system services.

These are all about our opinions, I will be grateful if you can seriously consider our plan. And we will live up to our common wish, improving the medical level in China.

Yours sincerely

Appendix

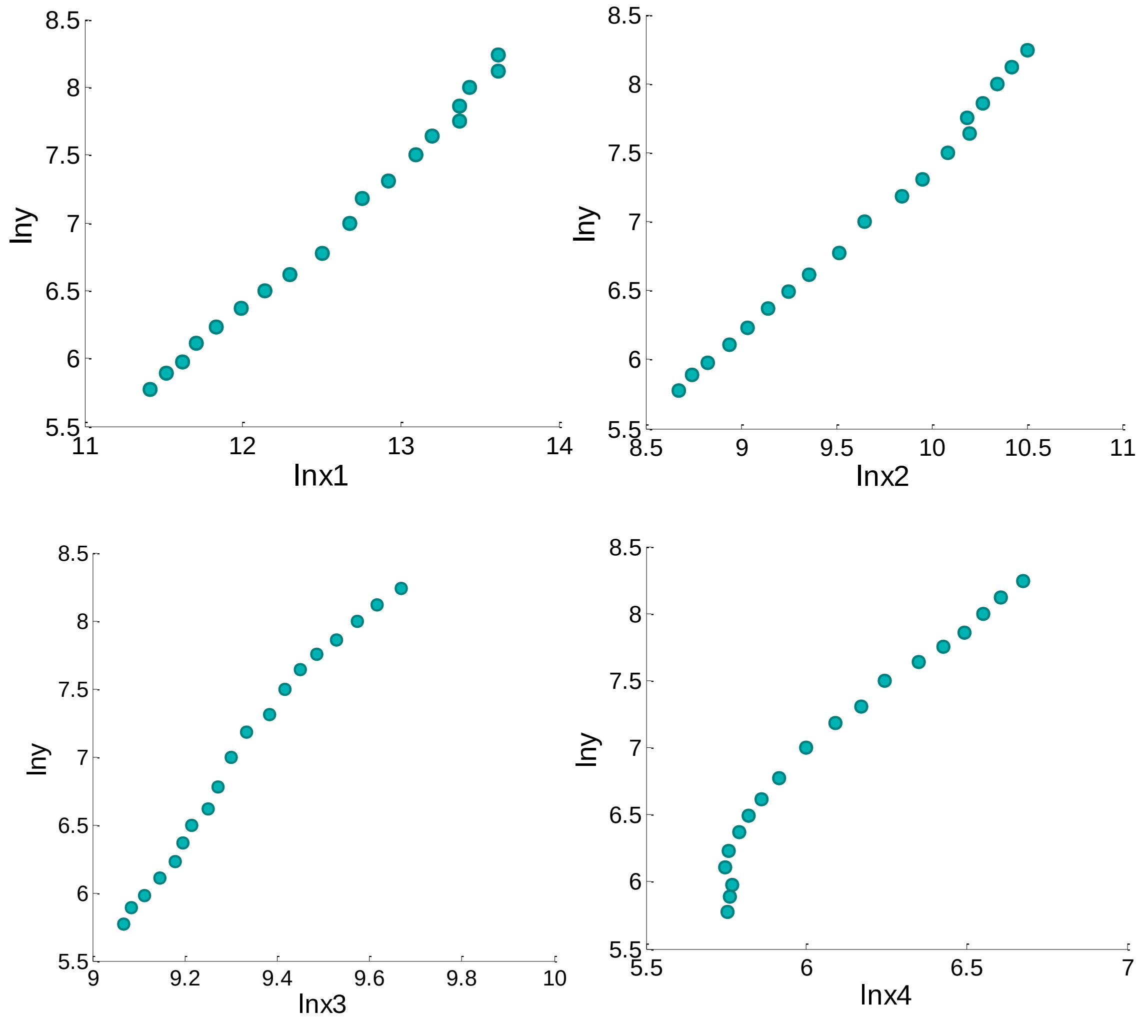
Task 1:

```
clc;
clear;
X0=[0.0671 0.069 0.0696 0.0710 0.0730 0.0750 0.0758 0.0769 0.0793 0.0805 0.0825 0.0847
0.0867 0.0912 0.0939 0.0967 0.1006 0.1047 0.1085 0.1139 0.1194];
% x=1998:2018;
n=length(X0);
for i=2:n %Start modeling feasibility analysis
    Q(i)=X0(i-1)/X0(i);
end
Q(1)=[];
ma=max(Q);
mi=min(Q);
if ma>exp(2/(n+1))
    disp(['The sequence cannot be grey predicted']);
    return
elseif mi<exp(-2/(n+1))
    disp(['The sequence cannot be grey predicted']);
    return
else
    disp(['The sequence can be grey predicted']);
end
clear Q ma mi %End of the test

X1=cumsum(X0);%Accumulating operator vector X1
Z1=ones(n-1,2);
for i=1:(n-1)
    Z1(i,1)=-(X1(i)+X1(i+1))/2;
    Z1(i,2)=1;%Mean generation operator Z1
end
Z1T=Z1';%The transposition Z1T of the mean generator matrix Z1
for j=1:n-1
    Y(j)=X0(j+1);
end
YT=Y';
A=inv(Z1T*Z1)*Z1T*YT;%Least squares estimation calculation parameters a, u
a=A(1);%Z1 parameter a
u=A(2);%System given parameter u
t=u/a;
```

```
t_test=5;
i=1:t_test+n;
X1S(i+1)=(X0(1)-t).*exp(-a.*i)+t;%Calculate the time response sequence and derive the
estimated accumulation vector X1S
X1S(1)=X0(1);
X0S(1)=X0(1);
for j=n+t_test:-1:2
    X0S(j)=X1S(j)-X1S(j-1);%Calculate the inverse accumulation vector X0S of X1S
end
for i=1:n
    Q(i)=X0S(i)-X0(i);%Residual
    E(i)=abs(Q(i))/X0(i);%Relative error
end
AVG=sum(E)/(n-1);%Average relative error
av=0.05;
if AVG>=av;%If the average relative error is greater than av%, enter the residual GM model
    clear cn Q1 CZ1 CZ1T CY CYT CA ca cu ct Q1S
    cn=length(Q);
    Q1=cumsum(Q);%Accumulating operator vector Q1
    CZ1=ones(cn-1,2);
for i=1:(n-1)
    CZ1(i,1)=-(Q1(i)+Q1(i+1))/2;
    CZ1(i,2)=1;%Mean generating operator CZ1
end
CZ1T=CZ1';%Transposition CZ1T of Mean Value Operator Matrices CZ1
for j=1:cn-1
    CY(j)=Q(j+1);
end
CYT=Y';
CA=inv(CZ1T*CZ1)*CZ1T*CYT;%Least squares estimation calculation parameters ca, cu
ca=CA(1);%CZ1 parameter a
cu=CA(2);%System given parameter cu
ct=cu/ca;
i=1:t_test+cn;
Q1S(i+1)=(Q(1)-ct).*exp(-ca.*i)+ct;%Calculate the time response sequence and derive the
estimated accumulation vector Q1S
X1S=X1S+Q1S;%Add the residual fit value to improve accuracy
for j=n+t_test:-1:2
    X0S(j)=X1S(j)-X1S(j-1);%Calculate the inverse accumulation vector X0S of X1S
end
clear av
    for i=1:cn
```

```
        Q(i)=X0S(i)-X0(i);%Residual
        E(i)=abs(Q(i))/X0(i);%Relative error
end
AVG=sum(E)/(n-1);%Average relative error
end
x=1:n;
xs=2:n+t_test;
yn=X0S(2:n+t_test);
plot(x+1998,X0,'^r',xs+1998,yn,'*-b','LineWidth',2); %Drawing
legend({'The original value','Forecast line'},'fontsize',14,'Location','northwest');
set(gca,'fontsize',12)
axis([1998 2025 0.06 0.14]) ;
xlabel('Year','fontsize',18);
ylabel('Aging ratio','fontsize',18);
disp(['The percent relative error is£º',num2str(AVG*100),'%']);
disp(['The fitted value is£º',num2str(X0S(1:n+t_test))]);
```



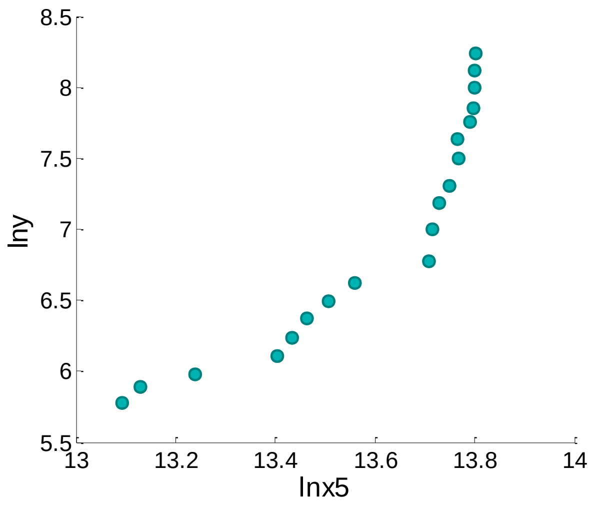


Figure 1.1: Scatter plot corresponding to each indicator

1.2 Stepwise regression process

Dependent Variable: LNY
Method: Least Squares
Date: 11/16/19 Time: 23:35
Sample (adjusted): 1999 2017
Included observations: 19 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LNX1	0.474968	0.013512	35.15200	0.0000
LNT	0.473166	0.074153	6.380987	0.0000

R-squared	0.930729	Mean dependent var	6.991192
Adjusted R-squared	0.926654	S.D. dependent var	0.812088
S.E. of regression	0.219933	Akaike info criterion	-0.091684
Sum squared resid	0.822302	Schwarz criterion	0.007731
Log likelihood	2.870993	Hannan-Quinn criter.	-0.074859
Durbin-Watson stat	0.202026		

Dependent Variable: LNY
Method: Least Squares
Date: 11/16/19 Time: 23:51
Sample (adjusted): 1999 2017
Included observations: 19 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-16.19071	1.082525	-14.95643	0.0000
LNX1	1.613856	0.169750	9.507239	0.0000
LNX2	0.846546	0.059762	14.16540	0.0000
LNT	-0.013099	0.018566	-0.705543	0.4913

R-squared	0.999154	Mean dependent var	6.991192
Adjusted R-squared	0.998985	S.D. dependent var	0.812088
S.E. of regression	0.025872	Akaike info criterion	-4.286671
Sum squared resid	0.010040	Schwarz criterion	-4.087841
Log likelihood	44.72337	Hannan-Quinn criter.	-4.253021
F-statistic	5906.633	Durbin-Watson stat	2.078833
Prob(F-statistic)	0.000000		

Dependent Variable: LNY
Method: Least Squares
Date: 11/16/19 Time: 23:24
Sample (adjusted): 1999 2017
Included observations: 19 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-17.96094	1.961399	-9.157210	0.0000
LNX1	0.291634	0.109131	2.672327	0.0182
LNX2	0.564561	0.121981	4.628285	0.0004
LNX3	1.712754	0.195128	8.777580	0.0000
T	-0.012489	0.010598	-1.178425	0.2583

R-squared	0.999423	Mean dependent var	6.991192
Adjusted R-squared	0.999258	S.D. dependent var	0.812088
S.E. of regression	0.022124	Akaike info criterion	-4.563408
Sum squared resid	0.006852	Schwarz criterion	-4.314872
Log likelihood	48.35238	Hannan-Quinn criter.	-4.521346
F-statistic	6059.762	Durbin-Watson stat	2.031904
Prob(F-statistic)	0.000000		

Dependent Variable: LNY
Method: Least Squares
Date: 11/16/19 Time: 23:25
Sample (adjusted): 1999 2017
Included observations: 19 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-11.02573	1.722555	-6.400802	0.0000
LNX1	0.236807	0.081662	2.899858	0.0124
LNX2	0.488508	0.098397	4.964660	0.0003
LNX3	0.808392	0.257019	3.145261	0.0077
LNX4	0.438846	0.142921	3.070543	0.0089
LNT	0.053648	0.027152	1.975821	0.0698

R-squared	0.999670	Mean dependent var	6.991192
Adjusted R-squared	0.999542	S.D. dependent var	0.812088
S.E. of regression	0.017371	Akaike info criterion	-5.015931
Sum squared resid	0.003923	Schwarz criterion	-4.717687
Log likelihood	53.65134	Hannan-Quinn criter.	-4.965456
F-statistic	7865.229	Durbin-Watson stat	2.309671
Prob(F-statistic)	0.000000		

Task 2:

Table 2.1： Results of different disease calculations

Disease	SSE	RMSR	R^2	model	Prediction
HIV/AIDS	2. 11e6	593. 6	0. 99	$y = 119.8x^2 + 31..86x + 4523$	31960
Cancer	3. 2e7	3261	0. 985	$y = -712.3x^2 + 25670x + 160200$	385000
Heart disease	1. 172e9	10320	0. 98	$y = -329.1x^2 + 20700x + 172000$	408500
Cerebrovascular disease	5. 572e-8	7464	0. 966	$y = 97.9x^3 - 3279x^2 + 37160x + 135200$	28520
Lung cancer	7. 42e7	44307	0. 93	$y = 3256x + 36430$	85270
Thyroid cancer	6. 1e7	2950	0. 98	$y = 770.2x^2 - 3687x + 13570$	131600
Colorectal cancer	7. 6e6	1040	0. 99	$y = 34.4x^2 + 2302x + 10700$	52970
Liver cancer	2. 69e8	5464	0. 9	$y = -36.41x + 38500$	37950

```
function [fitresult, gof] = createFit(x, y)
%CREATEFIT(X,Y)
%   Create a fit.
%   Data for 'untitled fit 1' fit:
%        X Input : x
%        Y Output: y
%   Output:
```

```
%          fitresult : a fit object representing the fit.
%          gof : structure with goodness-of fit info.
%% Fit: 'untitled fit 1'.
[xData, yData] = prepareCurveData( x, y );
% Set up fittype and options.
ft = fittype( 'poly2' );
% Fit model to data.
[fitresult, gof] = fit( xData, yData, ft );
% Plot fit with data.
figure( 'Name', 'untitled fit 1' );
h = plot( fitresult, xData, yData);
% Label axes
grid on
```

Task 3:

```
clc
clear
%    Initialize the patient source
%******************************************
%    Total simulation time
Total_time = 10;
N = 20;
%
lambda = 10;
mu = 5;
%Average arrival time and average service time
arr_mean = 1/lambda;
ser_mean = 1/mu;
%Maximum number of possible patients (round: round to round)
arr_num = round(Total_time*lambda*2);
%patient event table initialization
arr = [];
%According to the negative exponential distribution, each patient reaches the time interval
arr(1,:) = exprnd(arr_mean,1,arr_num);
%The arrival time of each patient is equal to the cumulative sum of time intervals
arr(1,:) = cumsum(arr(1,:));
%The service time of each patient is generated according to the negative exponential
distribution
arr(2,:) = exprnd(ser_mean,1,arr_num);
%Calculate the number of simulation patients, that is, the number of patients at the arrival
time in the simulation time
len_sim = sum(arr(1,:)<=Total_time);
%******************************************
```

```
% Calculate the information of the first patient
%*******************************************
%The first patient enters the system and accepts the service directly without waiting
arr(3,1) = 0;
%Its departure time is equal to its arrival time and service time
arr(4,1) = arr(1,1)+arr(2,1);
%flag=1
%
arr(5,1) = 1;
member = [1];
%*******************************************
% Calculate the information of patient I
%*******************************************
for i = 2:arr_num
%If the arrival time of patient I exceeds the simulation time, the loop is broken
if arr(1,i)>Total_time
break;
%27/5000
%If the arrival time of patient I does not exceed the simulation time, it is included
%The number of people already in the arrival time system
else number = sum(arr(4,member) > arr(1,i));
%
if number >= N+1
arr(5,i) = 0;
else if number == 0
%Its wait time is 0
arr(3,i) = 0;
%its departure time equals the sum of arrival time and service time
arr(4,i) = arr(1,i)+arr(2,i);
%flag= 1
arr(5,i) = 1;
member = [member,i];
%f the system has   patients who are being served
%enters the system
else len_mem = length(member);
%The wait time is equal to the departure time of the previous patient in the queue minus its
arrival time
arr(3,i)=arr(4,member(len_mem))-arr(1,i);
%Its departure time is equal to the departure time of the previous patient in the queue plus its
service time
arr(4,i)=arr(4,member(len_mem))+arr(2,i);
% Identify/flag
```

```
% Represents the number of patients in the system after it enters the system
arr(5,i) = number+1;
member = [member,i];
end
end
end
end
%The total number of patients entering the system at the end of simulation
len_mem = length(member);
%****************************************
% The output
%****************************************
%Plot the arrival and departure times of all patients entering the system during simulation time
%residence time£¨stairs£ºDraw a two-dimensional ladder diagram£©
stairs(0:len_mem,[0 arr(1,member)],'LineWidth',2);
hold on;
stairs(0:len_mem,[0 arr(4,member)],'.-r','LineWidth',2);
legend('Time of arrival ','Leave time ');
set(gca,'FontSize',15);
xlabel('Number of patients arriving','FontSize',18);
ylabel('Time','FontSize',18);
hold off;
grid on;
%During simulation time, the residence time and waiting time of all patients entering the system
% waiting time£¨plot£ºDraw a two-dimensional linear graph£©
figure;
plot(1:len_mem,arr(3,member),'r-*',1:len_mem,arr(2,member)+arr(3,member),'k-','LineWidth',1.8);
legend('Waiting time ','Residence time');
xlabel('Number of patients arriving','fontsize',18);
ylabel('Time','fontsize',18);
set(gca,'FontSize',15);
grid on;
```

Task 4:

```
% KMeans
clc;clear;
X=[7.01713043478261,1.42239130434783,0.379304347826087,1.89652173913043,5.40508695652174,3.12926086956522,1.51721739130435,5.02578260869565,6.82747826086957,8.53434782608696,12.2325652173913,14.1290869565217,13.8446086956522,5.78439130434783,5.40508695652174,4.83613043478261,5.87921739130435,5.87921739130435,7.206782
```

```
60869565,7.01713043478261,7.68091304347826,7.11195652173913,8.06021739130435,8.5
3434782608696,8.72400000000000,9.48260869565217,10.5256956521739,11.18947826086
96,12.8963478260870,13.0860000000000,12.4222173913043,11.9480869565217,11.853260
8695652,18.0169565217391,15.6463043478261,16.1204347826087,17.7324782608696,19.9
134782608696];
Y=[0.189652173913043,1.51721739130435,8.72400000000000,17.6376521739130,18.3014
347826087,12.4222173913043,4.55165217391304,12.8963478260870,3.31891304347826,2.
65513043478261,1.42239130434783,2.93960869565217,3.60339130434783,4.93095652173
913,5.97404347826087,6.92230434782609,6.82747826086957,8.15504347826087,9.103304
34782609,10.9050000000000,9.19813043478261,5.31026086956522,5.59473913043478,11.
3791304347826,11.7584347826087,7.87056521739131,7.84211739130435,9.008478260869
56,12.4222173913043,10.2412173913043,6.82747826086957,6.63782608695652,6.4481739
1304348,12.7066956521739,9.76708695652174,5.40508695652174,4.26717391304348,6.92
230434782609];
N=38;
% plot(X, Y, 'r*');
% xlabel('X');
% ylabel('Y');
% title('data plot');
n = 3;
m = 1;
eps = 1e-7;
u1 = [X(1),Y(1)];
u2 = [X(2),Y(2)];
u3 = [X(3),Y(3)];
U1 = zeros(2,100);
U2 = zeros(2,100);
U3 = zeros(2,100);
U1(:,2) = u1;
U2(:,2) = u2;
U3(:,2) = u3;
ND = zeros(3,N); %init dis
while(abs(U1(1,m) - U1(1,m+1)) > eps || abs(U1(2,m) - U1(2,m+1) > eps || abs(U2(1,m) -
U2(1,m+1)) > eps || abs(U2(2,m) - U2(2,m+1)) > eps))
    m = m +1;
    % dis
for i = 1 : N
    D(1,i) = sqrt((X(i) - U1(1,m))^2 + (Y(i) - U1(2,m))^2);
end
for i = 1 : N
    D(2,i) = sqrt((X(i) - U2(1,m))^2 + (Y(i) - U2(2,m))^2);
end
```

```
for i = 1 : N
        D(3,i) = sqrt((X(i) - U3(1,m))^2 + (Y(i) - U3(2,m))^2);
end
A = zeros(2,N); % A data
B = zeros(2,N); % B data
C = zeros(2,N);
for k = 1: N
        [MIN,index] = min(D(:,k));
        if index == 1    % along to first class
                A(1,k) = X(k);
                A(2,k) = Y(k);
        elseif index == 2
                B(1,k) = X(k);
                B(2,k) = Y(k);
        else
                C(1,k) = X(k);
                C(2,k) = Y(k);
        end
end
indexA = find(A(1,:) ~= 0); % class one center
indexB = find(B(1,:) ~= 0);
indexC = find(C(1,:) ~= 0);
U1(1,m+1) = mean(A(1,indexA));
U1(2,m+1) = mean(A(2,indexA));

U2(1,m+1) = mean(B(1,indexB));
U2(2,m+1) = mean(B(2,indexB)); % update centers

U3(1,m+1) = mean(C(1,indexC));
U3(2,m+1) = mean(C(2,indexC));
end
figure;
size=8;
plot(A(1,indexA) , A(2,indexA), 'ob','MarkerSize',size,'LineWidth',2');
hold on
plot(B(1,indexB) , B(2,indexB), 'or','MarkerSize',size,'LineWidth',2');
hold on
plot(C(1,indexC) , C(2,indexC), 'og','MarkerSize',size,'LineWidth',2');
hold on
centerx = [U1(1,m),U2(1,m),U3(1,m)];
centery = [U1(2,m),U2(2,m),U3(2,m)];
```

```
plot(centerx , centery, '^c','MarkerSize',size+5,'MarkerFaceColor','c');
xlabel('distance of x/km','fontsize',18);
ylabel('distance of y/km','fontsize',18);
set(gca,'FontSize',15);
grid off;
title('Data points after clustering','fontsize',18);
disp(['update time:',num2str(m)]);
```