

Statistical analysis and prediction of the construction demand of smart elderly care

Junyi Zhu ^{1,#}, Jingyi Sun ^{1,#,*}, Zhiyuan Chen²

¹School of Economics, Jiangsu University of Technology, Changzhou, 213001, China

²School of Mathematics and Science, Jiangsu University of Technology, Changzhou, 213001, China

*Corresponding author: 15190342359@163.com

#These authors contributed equally.

Abstract: China is facing a severe problem of population aging, and the number of elderly population is large and continues to grow. Under the current social background, the acceptance of new technology products among the elderly shows a steady increasing trend. From the perspective of empirical research, the popularization of new technology products has become one of the important indicators to measure the modern living level of the elderly. In addition, with the general improvement of education level and the diversification of information access channels, the cognitive ability and learning willingness of the elderly for new technology products have been significantly enhanced.

Keywords: Intelligent pension, ARIMA model, predictive analysis.

1. Introduction

With the increasingly prominent problem of population aging, the intelligent pension model, supported by big data and supported by artificial intelligence precision services, which has become an important means for countries in the world to alleviate the pension contradiction and cope with the deep aging of the population. In 2019, the National Medium- and Long-term Plan to Actively Address the Aging of the Population pointed out that the innovation-driven development strategy should be further implemented and the development of the smart health and elderly care industry should be continuously promoted. By 2020, the total number of elderly people aged 60 and above in the mainland will be 264 million, accounting for 18.7% of the total population. The annual net increment of the elderly population has also increased sharply from the lowest value in 2021 to the highest value in 2023. [1] The aging level of China's population has been pressed the acceleration key, and China's aging level has entered the "fast track" of growth. At this stage, the traditional home care service system has been difficult to meet the diversified needs of the elderly. In 2022, The State Council issued the 14th Five-Year Plan for the Development of undertakings for the Elderly and the Elderly Care Service System, which clarified the combination of Internet information technology and elderly care services.[2] In 2023, the National Standards Commission issued the Guidelines for National Standards Project Approval in 2023, which listed the standards for living services such as smart elderly care as the focus of the project approval. Continuous policy support and top-level design play a key role in the process of promoting the smart pension industry from market format to national strategic planning, and also provide a strong guarantee for the prosperity and development of smart pension related research.[3][4]

2. The basic fundamental of ARIMA Model construction

2.1 ARIMA Description of the model parameters

ARIMA The model is determined by 3 important parameters (p, d, q):The p is the autoregressive coefficient (Auto-Regressive), which means that the sequence value lags the p order; The d corresponds to the "I" in the ARIMA model, indicating that the temporal data needs to become a



stationary sequence; The q is the sliding mean coefficient (Moving Average), indicating that the error term lags the order q.

2.2 ARIMA Mathematical representation of the model:

autoregressive model (AR)

The formula of the p-order autoregressive model is defined as:

$$y_t = \mu + \sum_{i=1}^p \gamma_i y_{t-i} + \varepsilon_t \quad (1)$$

Here, y_t is the current value, μ is the constant term, γ_i is the autocorrelation coefficient, and ε_t is the error.

Sliding mean model (MA)

The formula of the q-order autoregressive model is defined as:

$$y_t = \mu + \sum_{i=1}^q \theta_i \varepsilon_{t-i} + \varepsilon_t \quad (2)$$

Here, θ_i is the sliding mean coefficient.

Autoregressive-Sliding Average Model (ARMA)

The formula for the self-regression-sliding mean model is defined as:

$$y_t = \mu + \sum_{i=1}^p \gamma_i y_{t-i} + \varepsilon_t + \sum_{i=1}^q \theta_i \varepsilon_{t-i} \quad (3)$$

2.3 ARIMA Confirmation of the model parameters

In the time-series analysis, the ARIMA model is used to model and predict the non-stationary time-series data.

Autocorrelation function (ACF)

The autocorrelation function reflects the correlation of adjacent observation points in the temporal data, which is formulated as follows:

$$\rho_k = \frac{cov(y_t, y_{t+k})}{\sqrt{var(y_t)}\sqrt{var(y_{t+k})}} \quad (4)$$

The autocorrelation function is a tool to measure the correlation of the time series at different time points. It shows the degree to which the time series is correlated with itself at different lag times. In the ARIMA model, ACF plots can help us identify the order (q) of the model in MA. If the ACF graph rapidly decays down to zero, after some lag order, this suggests that the order of the MA part may be just before this point.

Partial autocorrelation function (PACF)

The partial autocorrelation function is the correlation between and measured after a random variable removes the effect of an intermediate k-1 value by the formula:

$$\varphi_{kk} = \begin{cases} \rho_1, & k = 1 \\ \frac{\rho_k - \sum_{j=1}^{k-1} \varphi_{k-1, j} \varphi_{k-j}}{1 - \sum_{j=1}^{k-1} \varphi_{k-1, j} \varphi_{k-j}}, & k > 1 \end{cases} \quad (5)$$

The partial autocorrelation function is an extension of the autocorrelation function that considers the intermediate variables of the time series. The PACF diagram can help us to identify the order (p) of the AR part in the model. If the PACF plot intercintercudes after some lag order, this suggests that the order of the AR part may be just before this point.

2.4 Differential order number (d)

Differential order is how many differential operations on the original time series to make it smooth. This is usually determined by observing the ACF and PACF plots and by performing a unit root test (e. g., the ADF test). The purpose of the difference is to eliminate the trend and seasonal components in the time series and thus satisfy the stationarity requirement of the ARIMA model.

After determining the difference order d , we can preliminarily determine the order p and q of the AR and MA parts by observing the ACF and PACF plots. Then, information criteria such as the Akaike Information criterion (AIC) or the Bayesian Information criterion (BIC) can be used to compare the goodness of fit of different models and choose the optimal model. Finally, a residual analysis was used to test whether the assumptions of the model were met. This includes checking for the independence of the residuals, homoscedasticity, and normality. If the model diagnosis shows problems with the model, it may need to reconsider the order of the model or try other models.

3. Results

The data selected for this analysis is the situation of the elderly aged 65 and above in China in the past decade (2012-2023), with a table, including the year and the number of elderly people aged 65 and above (10,000). Considering the practical significance, the number of elderly population aged 65 and above (10,000) is taken as the observed data. We plotted the time series and performed the taking log analysis. Taking the log data fluctuated in a smaller range and meeting the criteria for further analysis. Next, we conducted the unit root test, we tested the original unit root, found that the p value of the original data was greater than 0.05, failed the unit root test of the first order difference, $p=0.0474 < 0.05$ for the first order difference, and the unit root test passed. Then, we carried out stationarity test and pure random sequence test, and found that both autocorrelation and partial autocorrelation maps have parts beyond the standard deviation of two columns, believing that there is a short-term sequence correlation, that is, this time series data is stable. If the p -value of the Q-test is all less than 0.05, then the original data is not a pure random sequence. To investigate the characteristics of the autocorrelation coefficient: the end of order 1 and the end of order 1, we tried to use multiple models to fit the development of the sequence, and investigate the fitting optimization problem of the sequence model. We obtained the results of the ARMA model and found that the P value was not less than 0.05 and the model test failed. Subsequently, we explored the model, and made P -values greater than 0.05, indicating that the residual was white noise, and P -values were less than 0.05. The model test was passed, and the equation was obtained:

$$X(t) - 0.056496 = -0.56496 \times (X(t-1) - 0.056496) \varepsilon(t) - 0.917761 \times \varepsilon(t-1) \quad (6)$$

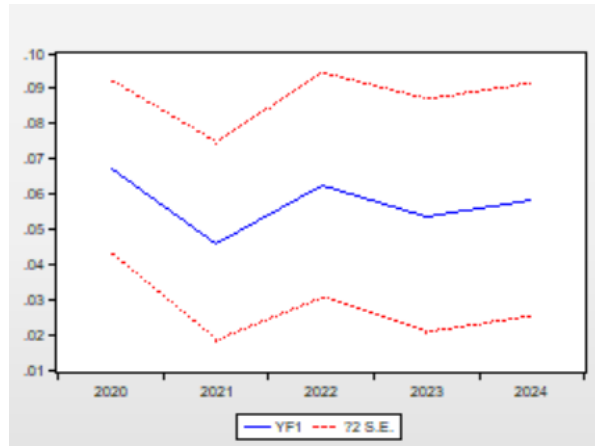


Figure 1. Shows the prediction results

Finally, the author used the best model to predict the number of people aged 65 and above at the end of 2024. According to the above forecast chart, the number of elderly people aged 65 and above will reach 223.99 million in 2024.

4. Prospect

This paper has made a detailed prediction of the demand of national smart pension construction and the changing trend of the elderly population. The author believes that the author can further analyze and predict the demand of smart pension construction in each region, so as to achieve accurate positioning, precise and precise policies of local construction. In addition, the author believes that the research on the theme of smart pension needs to learn from international experience, promote the localization of the theoretical system, pay attention to the smart pension in rural areas and towns, and strengthen the research on key technology products.[5]

References

- [1] The "Fast track" to the National Bureau of Statistics. Population Opportunities and Challenges for high-quality Development in the New Era — Interpretation of the seventh National Population Census Bulletin [EB / OL].2021-05-12.
- [2] The Chinese government network. The State Council issued the 14th Five-Year Plan for the Development of the Elderly and the Service System for the Elderly [EB / OL].2022-02-21.
- [3] Fa if ice, Wang Changqing. The development status of community smart pension in Jiangsu province from the perspective of dynamic ability [J]. Health Service Management in China, 2023,40 (08): 578-584.
- [4] Zhang Jingtang, Rui Guoqiang. Big data-driven intelligent health and pension: realistic representation, internal elements and optimization path [J]. Social Science in Hunan Province, 2023, (05): 132-141.
- [5] Xia Siyang, Zhu Xuefang. Research on smart health Information service system for the elderly — Based on multi-source data fusion technology [J / OL]. Intelligence Science, 1-14 [2024-04-28].