

运用 ARIMA 模型对无锡市人口发展趋势预测

黄梦婷

(江苏信息职业技术学院商学院 江苏 无锡 214000)

【摘要】无锡市人口发展“十四五”规划和2035年远景目标纲要提出持续优化人口结构，顺应人口发展规律和趋势，推动人口结构优化、社会融合、素质提升，保障各类群体合法权益，促进人口长期均衡发展与家人和谐幸福。本文构建ARIMA模型对无锡市未来十年人口进行预测，为涉及到人口问题的相关决策提供参考。

【关键词】人口预测；ARIMA；建模

The ARIMA model is used to predict the population development trend of Wuxi City Mengting Huang

(Business School, Jiangsu Polytechnic of Information Technology, Wuxi, Jiangsu, 214000)

[Abstract] The “14th Five-Year Plan” for population development and the outline of the 2035 vision goal of Wuxi City propose to continuously optimize the population structure, conform to the law and trend of population development, promote the optimization of population structure, social integration, quality improvement, safeguard the legitimate rights and interests of various groups, and promote the long-term balanced development of population and family harmony and happiness. In this paper, the ARIMA model is constructed to forecast the population of Wuxi City in the next ten years, so as to provide reference for the relevant decision-making related to population issues.

[Key words] Population forecast; ARIMA; Modeling

1 人口预测的重要性

深入有效的挖掘历次人口普查数据，预测无锡市中长期人口发展趋势，将为科学编制“十四五”规划、推动高质量发展、科学制定人口发展战略提供重要信息支撑，为更好地把人口因素融入经济社会政策提供重要的决策依据，无锡市“十四五”规划和2035年远景目标纲要提出要坚持以人民为中心，创造高品质生活，中长期人口发展趋势预测在大力促进高质量就业、完善多层次社会保障体系、加快教育现代化建设、提高全民健康水平、积极应对人口老龄化方面发挥着举足轻重的作用。

2 ARIMA 模型简述及数据说明

2.1 ARIMA 模型简述

ARIMA 模型，也称为求和自回归移动平均模型，实际上是差分运算与 ARMA 模型的结合。ARIMA(p, d, q) 模型记为：

$$\begin{cases} \Phi(B)\nabla^d x_t = \Theta(B)\varepsilon_t \\ E(\varepsilon_t) = 0, \text{Var}(\varepsilon_t) = \sigma_t^2, E(\varepsilon_t \varepsilon_s) = 0, s \neq t \\ E(x_s \varepsilon_t) = 0, \forall s < t \end{cases} \quad (1)$$

(1) 式中， $\nabla^d = (1-B)^d$ ；

$\Phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$ ，为平稳可逆 ARMA(p, q) 模型的自回归系数多项式；

$\Theta(B) = 1 - \theta_1 B - \dots - \theta_q B^q$ ，为平稳可逆 ARMA(p, q) 模型的移动平滑系数多项式；模型中的 d 为差分的阶数； $\{\varepsilon_t\}$ 为零均值白噪声序列。

ARIMA 建模分析需要对序列进行平稳性检验、白噪声检验、模型定阶、残差检验以及模型预测五个步骤，在各环节中依据检验结果进行适当的步骤调整，具体步骤如图 1 所示：

2.2 数据说明

在无锡市统计年鉴获取到 1975-2022 年无锡市年末户籍总人口数，共 48 期数据，整体呈现上升趋势，截止到 2022 年无锡市年末户籍人口数达 519 万人。

3 ARIMA 人口预测

3.1 平稳性检验和差分处理

首先对短期、单变量的人口时间序列数据做 ADF 平稳性检验 (Augmented Dickey-Fuller test)，也称为单位根检验，如果序列是平稳的可进行接下来的建模分析过程，否则需要对序列进行差分处理，直到差分后的序列平稳为止。ADF 检验的原假设为：序列是平稳的，通过计算不同显著性水平下的统计量及 p 值，来判断序列是否平稳。运用无锡市 48 个年份的户籍人口数序列做 ADF 平稳性检验结果如表 1 所示：

本数据中，adf 结果为 -1.9，大于 1%、5% 和 10% level 的统计值，不能拒绝原假设，说明数据不是平稳的，需要对序列进行差分处理。经过一阶差分处

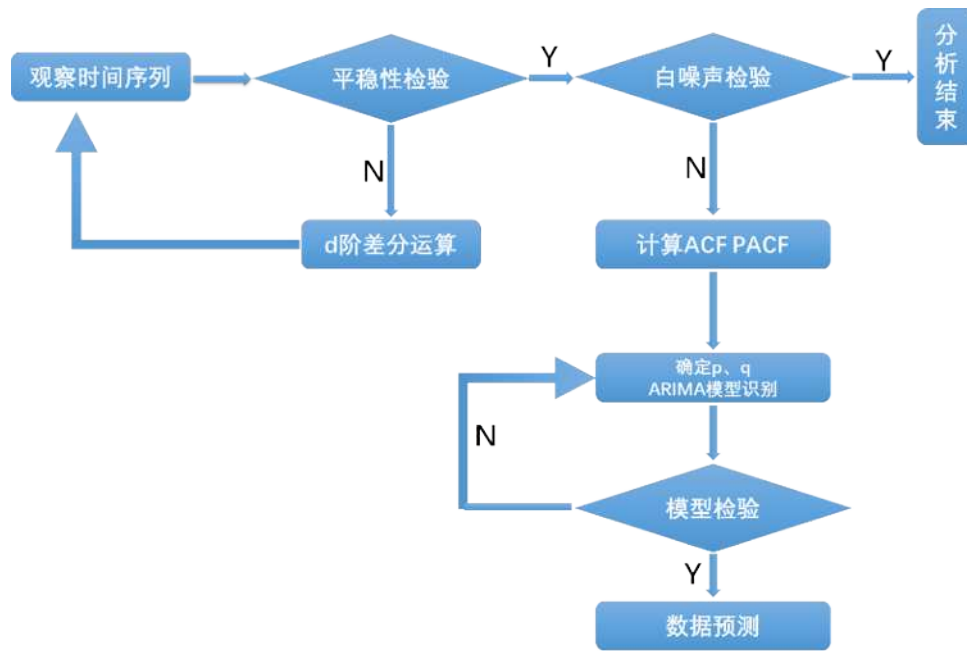


图1 ARIMA 建模流程

表1 不同程度拒绝原假设的统计值

显著性水平	1%	5%	10%
统计值	-4.331573	-3.23295	-2.7487

理后，序列表现平稳，因此 ARIMA (p, d, q) 模型中的 d 等于 1。

3.2 白噪声检验

通过 ADF 检验证明序列是平稳的，接下来需要对平稳的时间序列做白噪声检验，判断该序列是否符合随机序列的特点，如果为白噪声说明序列不适合建模分析。对 1975-2022 年无锡市户籍人口数序列进行白噪声检验，判断人口数是否为白噪声序列。常用的方法有 3 种，自相关图、Box-Pierce 检验和 Ljung-box 检验。本文采用 Ljung-box 做白噪声检验，结果如表 2 所示，表中数据 p 值表明一阶差分后的序列不是白噪声序列，可以进行下一步的建模步骤。

表2 Ljung-box 做白噪声检验

lb_stat	lb_pvalue	bp_stat	bp_pvalue
6	46.030591	2.91932E-08	42.114051751
12	53.9879553	2.74623E-07	48.426574842
			2.63409E-06

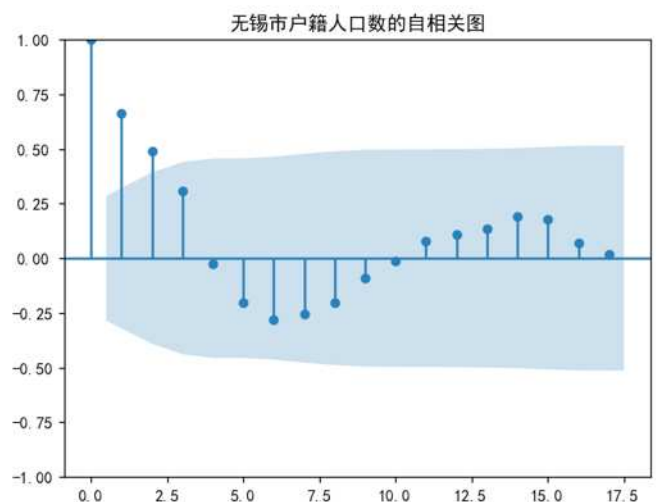
3.3 模型定阶及估计

模型定阶即确定 ARIMA (p, d, q) 模型中的 p 和 q 的值，通常有两种方法，第一种方法是计算自相关系数 ACF 和偏自相关系数 PACF，根据它们表现出来的“拖尾”性和“截尾”性选择适当的 p 和 q 值拟合模型，ARMA 模型定阶的基本原则如表 3 所示；第二种方

法是根据 AIC 和 BIC 信息准则判别，通过寻找最小的 AIC、BIC 确定 p 和 q 的值。本文采用第一种方法，绘制 1975-2022 年无锡户籍人口序列的自相关序列图和偏自相关序列图，如图 2 所示，一阶差分后的人口数据自相关图拖尾，在 3 阶位置就开始基本落在 2 倍标准差范围，偏自相关图 1 阶之后呈现拖尾。因此设定模型为 ARIMA (1, 1, 3) 来拟合无锡市历年户籍人口数的变动趋势。

表3 ARMA 模型定阶的基本原则

ACF	PACF	模型定阶
拖尾	P 阶截尾	AR(P) 模型
q 阶截尾	拖尾	MA(q) 模型
拖尾	拖尾	ARMA(p, q) 模型



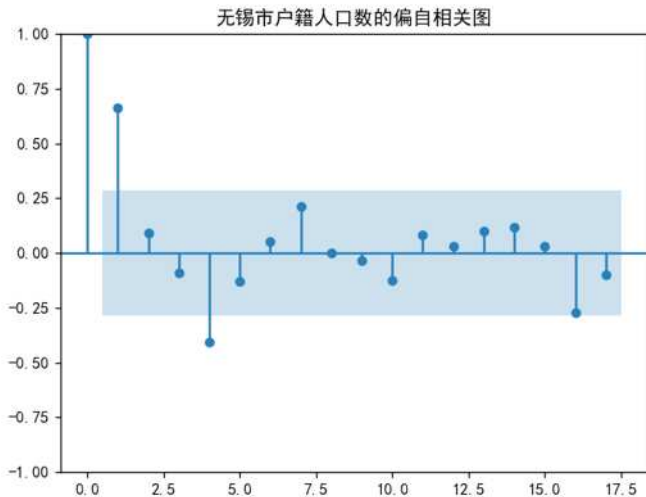


图2 自相关图和偏自相关图

3.4 残差检验

运用ARIMA(1, 1, 3)模型拟合无锡市户籍人口数变动效果好坏可用残差检验进行度量。通过计算原始序列值与预测值的差值得到残差序列，对残差序列做白噪声检验，即为残差检验。如果残差序列为白噪声，说明模拟拟合效果好；若未通过白噪声检验，说明残差序列中还存在有用信息未拟合到模型中，需要对构建的ARIMA(1, 1, 3)模型进一步的调参改进。如图3所示，该模型残差项符合正态分布特征，即通过白噪声检验，建模结束。

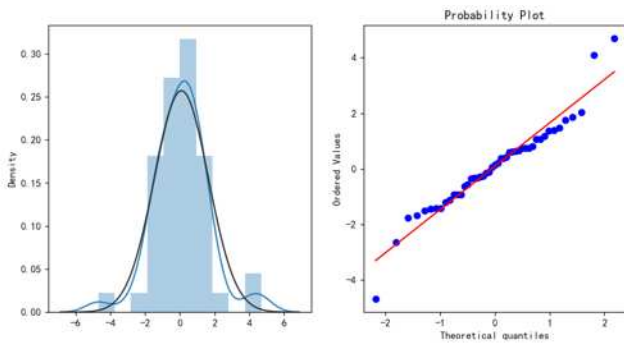


图3 残差正态性检验

3.5 无锡市人口预测

通过平稳性检验、白噪声检验、模拟定阶和残差检验，基于2005-2022年无锡市户籍人口时间序列数据，构建了无锡市户籍人口预测模型ARIMA(1, 1, 3)，

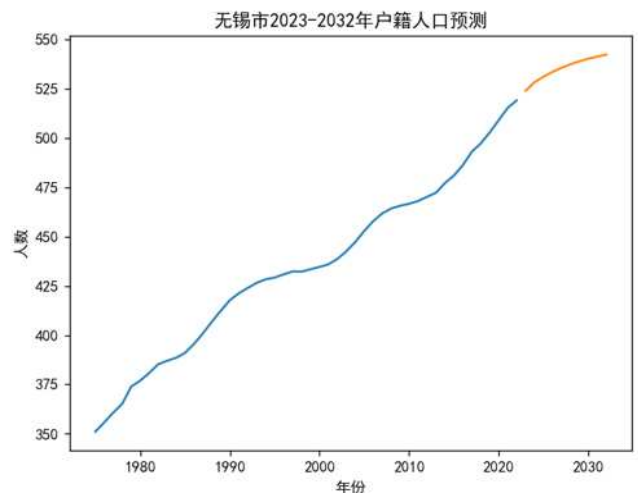


图4 无锡市 2023-2032 年户籍人口数预测

下面运用该模型对无锡市未来10年户籍人口数量预测，如图4所示可见在未来10年无锡市户籍人口数仍呈上升趋势，但增长势头逐渐减缓，到2032年无锡市户籍人口数达542.2万人。

4 结论与建议

本文通过构建ARIMA模型，对无锡市户籍人口进行预测分析得到未来十年人口预测数据，其中关键年份2023、2025、2030、2032年人口数分别为：523.9、531.0、540.1、542.2万人。ARIMA模型通过历史数据，拟合户籍人口数的变动规律，没有考虑到外界因素如国家政策、宏观经济、自然因素等对人口数的影响，因此该仅适用于短期预测，无锡市户籍人口的长期预测还需进一步的研究。

参考文献：

- [1] 刘晓峰. 河南省未来人口发展趋势[J]. 河南教育学院学报(哲学社会科学版), 2007, 26(4): 43-48. DOI: 10.3969/j.issn.1006-2920.2007.04.010.
- [2] 赵跃进. 陕西未来人口发展趋势分析[J]. 西北人口, 2008, 29(5): 41-43, 48. DOI: 10.3969/j.issn.1007-0672.2008.05.026.

作者简介：

黄梦婷(1996.07.21-)，河南信阳人，硕士，承担数据分析、数据挖掘、金融等课程教学。