

# 基于 ARIMA 模型的安徽省城乡收入差距的预测

刘兆鹏, 陈敏

(宿州学院数学与统计学院, 安徽 宿州 234000)

**摘要:**根据 1978~2012 年安徽省城乡居民收入的时间序列数据,运用 ARIMA 模型,用 SAS 软件进行拟合,对安徽省城乡居民收入差距情况进行分析,给出了 2013~2015 年安徽省城乡居民绝对收入差距的预测值。结果显示城乡居民收入差距会继续扩大。

**关键词:**ARIMA 模型;城乡收入差距;预测

**中图分类号:**F12

**文献标志码:**A

随着安徽省经济水平的提高,城乡居民的收入都有了较大增长,人民生活水平有了较大的改善。在城乡居民收入增加的同时,收入差距也在不断拉大,从 1978 年的 97 元扩大到了 2012 年的 13824 元,这导致“三农”问题日益显著,城乡发展极不平衡,影响劳动者的积极性,危及社会稳定和谐<sup>[1-2]</sup>。本文采用时间序列分析方法,研究安徽省城乡居民收入差距,建立 ARIMA 预测模型,对安徽省城乡收入差距的发展趋势进行预测和分析。

## 1 ARIMA 模型

具有如下结构的模型称为求和自回归移动平均模型,记为 ARIMA(p,d,q)模型<sup>[3-4]</sup>:

$$\begin{cases} \Phi(B) \nabla^d x_t = \Theta(B) \varepsilon_t \\ E(\varepsilon_t) = 0, \text{Var}(\varepsilon_t) = \sigma_\varepsilon^2, E(\varepsilon_t \varepsilon_s) = 0, s \neq t \\ E\chi_s(\varepsilon_t) = 0, \forall s < t \ln X_t \end{cases}$$

其中,  $\nabla \ln X_t$ ,  $\Phi(B) = 1 - \varphi_1 B - \dots - \varphi_p B^p$  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。

## 2 安徽省城乡收入差距分析与建模

### 2.1 数据分析

选取 1978~2012 年安徽省城镇居民人均可支配

收入和农村居民人均纯收入数据(表 1),从 1978 年到 2012 年,安徽省城镇居民人均可支配收入年均增长 184.38%,农村居民人均纯收入年均增长 81.36%,二者分别增加了 99.63 倍和 63.72 倍,但在安徽省城乡居民收入不断增加的同时,二者的差距也在不断拉大(图 1)。

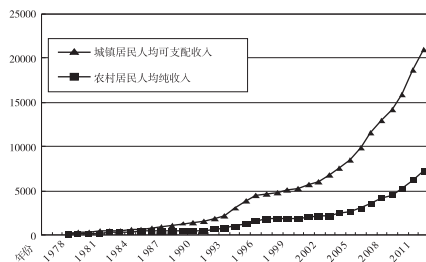


图 1 安徽省城乡居民收入差距

### 2.2 数据的平稳化处理与检验

安徽省城乡居民绝对收入差距序列(表 1)记为  $x_t$ , 其时序图(图 2),由图 2 可知除个别年份外,安徽省城乡居民收入差距总体呈上升趋势,且数据不具有明显的周期变化和季节波动,但呈现出一定的指数增长趋势,是典型的非平稳序列<sup>[5-6]</sup>。

将城乡居民收入差距序列取自然对数,消除异方差性;记为  $\ln x_t$ , 绝对收入差距对数序列蕴含显著的线性趋势,为提取序列  $\ln x_t$  的线性增长趋势,对  $\ln x_t$  进行 1 阶

表1 1978~2012年安徽省城乡居民收入

年份	城镇居民人均可支配收入(元/人)	农村居民人均纯收入(元/人)	城乡收入差距
1978	211.23	113.43	97.8
1979	278.19	170.23	107.96
1980	344.22	164.83	179.39
1981	424.56	246.49	178.07
1982	452.52	259.1	193.42
1983	448.43	304.64	143.79
1984	559.11	322.97	236.14
1985	633.85	369.41	264.44
1986	834.76	396.53	438.23
1987	924.84	429.26	495.58
1988	1074.7	446.05	628.65
1989	1274.7	485.53	789.17
1990	1354.6	515.66	838.94
1991	1485.14	539.16	945.98
1992	1795.54	573.38	1222.16
1993	2233.89	724.5	1509.39
1994	3035.38	973.2	2062.18
1995	3778.86	1302.82	2476.04
1996	4493.92	1607.72	2886.2
1997	4599.27	1808.75	2790.52
1998	4770.47	1863.06	2907.41
1999	5064.6	1900.29	3164.31
2000	5293.53	1934.57	3358.96
2001	5668.8	2020.04	3648.76
2002	6032.4	2117.57	3914.83
2003	6778.03	2127.48	4650.55
2004	7511.43	2499.33	5012.1
2005	8470.68	2640.96	5829.72
2006	9771.05	2969.08	6801.97
2007	11473.58	3556.27	7917.31
2008	12990.35	4202.49	8787.86
2009	14085.74	4504.32	9581.42
2010	15788.2	5275.1	10513.1
2011	18606.13	6232.3	12373.83
2012	21024.21	7200	13824.21

注:数据来源于2013年《安徽省统计年鉴》。

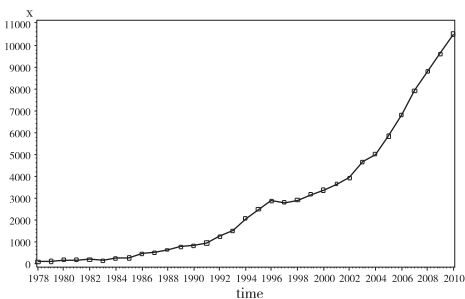


图2 1978~2012年安徽省城乡居民收入差距时序图

差分,记为  $\nabla \ln x_t$ , 得到差分后序列  $\nabla \ln x_t$  时序图(图3)。

由图3可知绝对收入差距对数序列1阶差分后的序列基本平稳,通过观察  $\nabla \ln x_t$  序列的自相关图(图4)和  $\nabla \ln x_t$  序列的 ADF 检验结果和白噪声检验结果(图5、图6),从而确定是否可以采用此  $\nabla \ln x_t$  序列进行模型的建立。

由  $\nabla \ln x_t$  序列1阶差分后的序列自相关图4可知,  $\nabla \ln x_t$  序列的自相关系数一直都比较小,基本控制在2倍的标准差范围以内,具有显著的短期相关性,表明

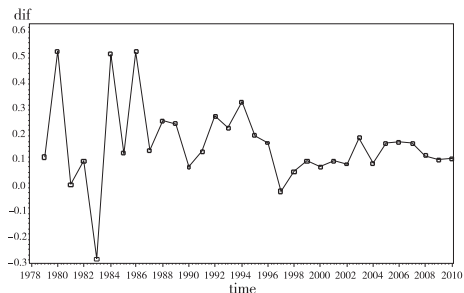


图3 安徽省城乡居民收入差距对数序列1阶差分时序图

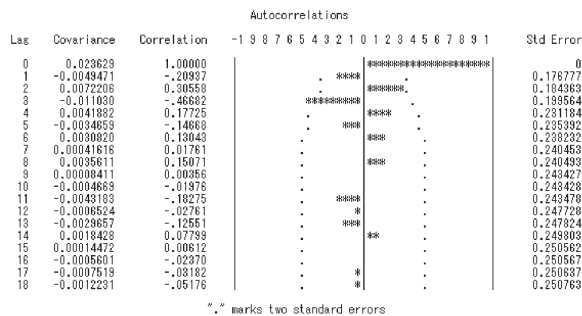


图4 城乡收入差距对数序列1阶差分自相关图

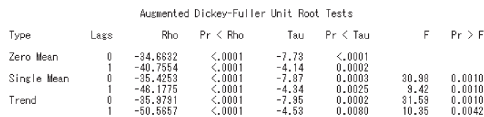


图5 城乡收入差距对数序列1阶差分 ADF 检验结果图

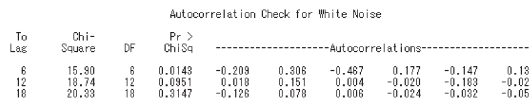


图6 城乡收入差距对数序列1阶差分白噪声检验结果图

$\nabla \ln x_t$  序列为平稳序列;同时图5显示的 ADF 检验结果各个 P 值都小于 0.05,也说明  $\nabla \ln x_t$  序列为平稳序列。图6白噪声检验结果显示  $\nabla \ln x_t$  序列是非白噪声序列,说明该序列还蕴含着值得提取的相关信息。

### 2.3 数据的平稳化处理与检验

对平稳非白噪声  $\nabla \ln x_t$  序列拟合 ARMA(p,q) 模型。自相关图1结果显示  $\nabla \ln x_t$  序列自相关系数3阶截尾,所以可以考虑拟合 MA(3)模型。再考察  $\nabla \ln x_t$  序列偏自相关图(图7),结果显示偏自相关系数3阶截尾,因此也可以拟合 AR(3)模型。

综上,可以考虑选用模型 ARIMA(3,1,0)和 ARIMA(0,1,3)对  $\ln x_t$  序列进行拟合。

通过检验可知两个模型 ARIMA(3,1,0)和 ARIMA(0,1,3)都有效,为了得到最优模型,采用 AIC 和 SBC 准则对模型进行比较。

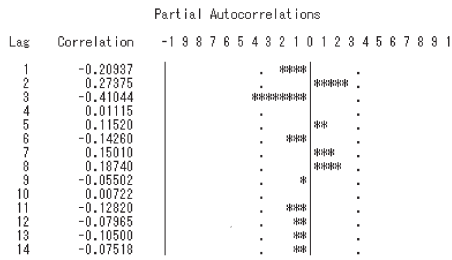


图 7 城乡收入差距对数序列 1 阶差分  
偏自相关图

由表 2, ARIMA(3,1,0) 模型的 AIC 和 SBC 都比 ARIMA(0,1,3) 模型的 AIC 和 SBC 小,因此选用 ARIMA(3,1,0) 模型,综上得最优模型为 ARIMA(3,1,0)。

表 2 AIC 和 SBC 准则对多个 ARIMA 模型比较

项目	ARIMA(3,1,0)	ARIMA(0,1,3)
AIC	-33.0009	-32.7751
SBC	-30.0694	-29.8436

确定模型为 ARIMA(3,1,0) 模型,通过 SAS 对其进行参数估计(图 8),其 ARIMA(3,1,0) 模型为:

$$(1 - B)\ln X_t = 0.14665 + \frac{\varepsilon_t}{1 + 0.47199B^3},$$

$$\varepsilon_t \sim N(0, 0.01965)$$

The ARIMA Procedure

Conditional Least Squares Estimation

Parameter	Estimate	Standard Error	t Value	Approx Pr >  t	Lag
MU	0.14865	0.01728	8.49	<.0001	0
AR1,1	-0.47199	0.16218	-2.91	0.0067	3

Constant Estimate	0.215872
Variance Estimate	0.019651
Std Error Estimate	0.140183
AIC	-33.0009
SBC	-30.0694
Number of Residuals	32

\* AIC and SBC do not include log determinant.

图 8 ARIMA(3,1,0) 模型参数估计结果图

利用此模型对 2012 年安徽省城乡居民收入差距进行预测,最终的预测结果是 14 015.82 元,2012 年实际收入差距为 13 824.21 元,预测相对误差近似为 1.3 % 说明预测值与实际结果是比较接近的,另外将预测值和序列观测值联合作图通过图示也可以直观地看出该模型对序列的拟合效果良好(图 9)。

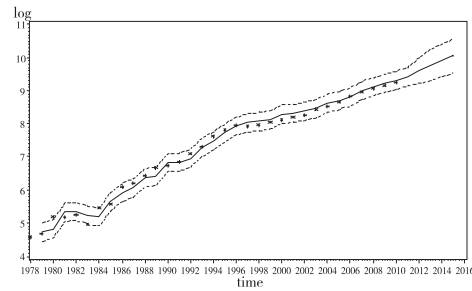


图 9 安徽省城乡居民收入差距序列模型拟合效果图

2013 ~ 2015 年安徽省城乡居民收入差距预测见表 3。

表 3 2013 ~ 2015 年安徽省城乡居民收入  
差距预测值

年份	2013	2014	2015
收入差距	17574.43	20158.88	23031.08

### 3 结束语

本文针对 1978 年至 2012 年安徽省城乡居民收入差距,进行 ARIMA 模型拟合,并且预测了 2013 ~ 2015 年城乡居民收入差距,未来 3 年安徽省城乡居民收入差距仍将继续拉大,城乡发展不平衡,如何保持城乡收入分配均衡,将是重中之重。

### 参考文献:

- [1] 范丽.基于 ARMA 模型的云南省城乡收入差距分析和预测[J].安徽农业科学,2012,40(19):10283-10285.
- [2] 符想花.城市化进程中城乡居民收入差距模拟与预测[J].统计与决策,2011(24):74-76.
- [3] 韩立胜,叶磊.城乡居民收入差距分析与预测—基于 ARMA 模型的实证研究[J].吉林金融研究,2011(7):33-37.
- [4] 白云,冯晓宁.基于 ARMA 模型对十二五期间河北省城乡居民收入及其差距的预测[J].开发研究,2011(5):15-18.
- [5] 涂雄苓.我国城乡居民收入差距预测模型比较研究[J].安徽农业科学,2011,39(24):15057-15060.
- [6] 李娟娟.1978-2007 年中国城乡收入差距的变迁及影响因素分析[J].四川理工学院学报:自然科学版,2010,25(3):50-52.

## Prediction of Urban-Rural Income Gap in Anhui Province Based on ARIMA Model

LIU Zhao-peng, CHEN Min

(School of Mathematics and Statistics, Suzhou University, Suzhou 234000, China)

**Abstract:** According to the time series data of urban-rural income gap of Anhui Province from 1978 to 2012, the income gap between urban and rural residents is analyzed with the ARIMA model and SAS fitting, and the absolute income gap from 2013 to 2015 is forecasted. The results show that the urban-rural income gap will keep enlarging.

**Key words:** ARIMA model; urban-rural income gap; forecast