

我国人口出生率的分析与预测

孙鑫鑫<sup>1)</sup> 高 燕<sup>1)</sup> 袁 汐<sup>1)</sup> 王增相<sup>2)</sup>

( 1) 山东师范大学数学与统计学院,250358,济南; 2) 阿萨巴斯卡大学商学院,T9S3A3,艾伯塔,加拿大 )

**摘要** 我国的人口出生率经历了很多阶段,尤其是近几年一直保持低人口出生率的态势.为了对未来我国人口出生率进行合理的预测,采用时间序列建模的方法,通过数据预处理、模型识别、参数估计以及模型检验与优化等步骤,建立了 ARIMA(1,2,2)模型,并利用该模型对未来5年我国的人口出生率进行预测,得到的结果分别为12.16‰、12.23‰、12.25‰、12.19‰、12.27‰.根据预测结果得出未来5年我国仍然保持低人口出生率这一结论.基于研究结论,为改善未来我国低人口出生率这一状况提出了合理化建议.

**关键词** 人口出生率; 时间序列; ARIMA 模型; Eviews; SAS  
**中图分类号** O 213.2 **文献标识码** A **doi**: 10.3969/j.issn.1001-4748.2020.03.002

1 引言

我国是一个人口大国,人口问题一直是关乎国计民生的热门话题.根据人口出生率的统计数据,可以发现全国人口出生率近年来总体呈现下降的趋势,尤其是近十年来的人口出生率一直维持在12‰左右,如果人口出生率一直保持在较低的水平势必会给社会带来许多问题.2016年二胎政策全面放开,文献[1]曾提出全面实施二胎政策可能会缓解人口出生率下降的趋势.事实上2016年和2017年这两年的全国人口出生率较之前确实有所上升,分别达到了12.95‰和12.43‰,但这可能只是短期效应,未来会怎样发展还需通过严谨科学的研究进行预测,只有这样才能更客观清晰地了解未来我国人口出生率的趋势.

本文对我国从1949年到2017年的人口出生率进行分析,建立时间序列模型预测未来5年我国的人口出生率是继续呈下降趋势还是会达到一个平稳的状态.由于影响出生率的因素非常多,因此在建模过程中可能会出现很多诸如平稳性检验、白噪声检验不通过或者模型拟合情况不好等问题,本文旨在解决这些问题,最终使残差变成白噪声序列,并且建立一个合理的模型对未来五年我国的人口出生率进行预测.

2 数据来源及分析

**2.1 数据来源** 从《中国统计年鉴》中选取自1949至2017年全国的人口出生率作为原始数据,并选择最合适的预测方法预测未来5年的全国人口出生率(表1).

**2.2 数据特征分析** 将1949-2017年的全国人口出生率绘制成折线图,如图1所示.

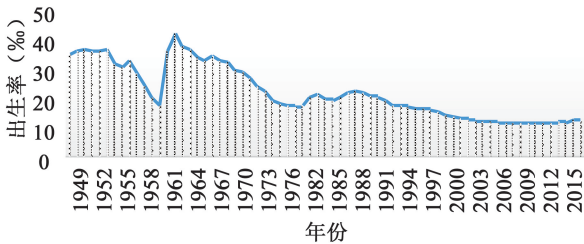


图1 全国人口出生率折线图

收稿日期:2019-11-09  
基金项目:国家自然科学基金青年基金资助项目(11601287).  
通讯作者:王增相,男,博士,副教授,硕士生导师.

表 1 1949–2017 全国人口出生率统计

年份	出生率(‰)	年份	出生率(‰)	年份	出生率(‰)
1949	36.00	1972	29.77	1995	17.12
1950	37.00	1973	27.93	1996	16.98
1951	37.80	1974	24.82	1997	16.57
1952	37.00	1975	23.01	1998	15.64
1953	37.00	1976	19.91	1999	14.64
1954	37.97	1977	18.93	2000	14.03
1955	32.60	1978	18.25	2001	13.38
1956	31.90	1979	17.82	2002	12.86
1957	34.03	1980	18.21	2003	12.41
1958	29.22	1981	20.91	2004	12.29
1959	24.78	1982	22.28	2005	12.40
1960	20.86	1983	20.19	2006	12.09
1961	18.02	1984	19.90	2007	12.10
1962	37.01	1985	21.04	2008	12.14
1963	43.37	1986	22.43	2009	11.95
1964	39.14	1987	23.33	2010	11.90
1965	37.88	1988	22.37	2011	11.93
1966	35.05	1989	21.58	2012	12.10
1967	33.96	1990	21.06	2013	12.08
1968	35.59	1991	19.68	2014	12.37
1969	34.11	1992	18.24	2015	12.07
1970	33.43	1993	18.09	2016	12.95
1971	30.65	1994	17.70	2017	12.43

从图 1 可以看出,我国的人口出生率大致可以分为五个阶段:早期的高出生率阶段,20 世纪 70 年代出生率大幅下降阶段,20 世纪 80 年代出生率有所反弹阶段,20 世纪 90 年代出生率持续下降阶段和近年来低出生率阶段.根据国家统计局人口普查数据,可以得到如下结论.

1) 1949–1970 年的高出生率阶段.这个阶段的人口出生率平均值在 30‰左右,新中国成立后,社会安定,经济发展,人们的生活水平和医疗条件都不断改善,因此在这个阶段总体上保持着高出生率.2) 1971–1979 年出生率大幅度下降阶段.这个阶段人口出生率大幅度下降至 17.82‰,这是因为人口高速增长带来了巨大的压力,人们认识到控制人口已迫在眉睫,在全国发出了实行计划生育的号召,并制定和完善了计划生育政策,导致这个阶段出生率呈现出大幅下降趋势.3) 1980–1987 年出生率小幅度的反弹阶段.这个阶段人口出生率回升至 23.33‰,这是因为在 20 世纪 60 年代初“第二次人口生育高峰”中出生的人口陆续进入生育年龄,使得人口出生率出现回升.4) 1988–2004 年持续的下降阶段.这个阶段人口出生率下降到了 12.29‰,主要是受到了计划生育政策不断加强和完善的影响.5) 2005–2018 年人口低出生率阶段.这个阶段的人口出生率平均值在 12.10‰左右,这个阶段仍然受计划生育政策的影响,加之养育孩子的成本增大,女性的生育观念发生转变等因素,使得出生率保持在较低水平.

### 3 人口出生率的时间序列模型

3.1 时间序列分析基本步骤 时间序列分析是常用的定量预测方法之一.应用时间序列分析建模可以寻找出序列值之间相关关系的统计规律,并拟合出适当的数学模型来描述这种规律,进而利用这个拟合模型预测序列未来的走势.

建立时间序列分析模型首先要对数据进行平稳性检验以及纯随机性检验,然后计算样本自相关系数和偏自相关系数.依据计算出来的自相关系数和偏自相关系数,按照 ARMA 模型定阶的基本原则对模型进行定阶.时间序列分析的基本模型如下<sup>[2–4]</sup>:

1) 自回归模型(AR 模型).一般的  $p$  阶自回归模型记为  $AR(p)$ ,它有如下结构:

$$\begin{cases} x_t = \varphi_0 + \varphi_1 x_{t-1} + \varphi_2 x_{t-2} + \cdots + \varphi_p x_{t-p} + \varepsilon_t, \\ \varphi_p \neq 0, \\ E(\varepsilon_t) = 0, Var(\varepsilon_t) = \sigma_\varepsilon^2, E(\varepsilon_t \varepsilon_s) = 0, s \neq t, \\ E(x_s \varepsilon_t) = 0, \forall s < t. \end{cases} \tag{1}$$

该模型的基本假设如下:

H<sub>1</sub> 模型最高阶数为  $p$  阶;

H<sub>2</sub> 随机干扰项序列  $\{\varepsilon_t\}$  为零均值白噪声序列;

H<sub>3</sub> 现在的随机干扰与过去的序列值不相关.

2) 移动平均模型(MA 模型).一般的  $q$  阶移动平均模型记为  $MA(q)$ , 它有如下结构:

$$\begin{cases} x_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q}, \\ \theta_q \neq 0, \\ E(\varepsilon_t) = 0, Var(\varepsilon_t) = \sigma_\varepsilon^2, E(\varepsilon_t \varepsilon_s) = 0, s \neq t. \end{cases} \tag{2}$$

该模型的基本假设如下:

H<sub>4</sub> 模型最高阶数为  $q$  阶.

H<sub>5</sub> 随机干扰项序列  $\{\varepsilon_t\}$  为零均值白噪声序列.

3) 自回归移动平均模型(ARMA 模型).一般的自回归移动平均模型记为  $ARMA(p, q)$ , 它有如下结构:

$$\begin{cases} x_t = \varphi_0 + \varphi_1 x_{t-1} + \cdots + \varphi_p x_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \cdots - \theta_q \varepsilon_{t-q}, \\ \varphi_p \neq 0, \theta_q \neq 0, \\ E(\varepsilon_t) = 0, Var(\varepsilon_t) = \sigma_\varepsilon^2, E(\varepsilon_t \varepsilon_s) = 0, s \neq t, \\ E(x_s \varepsilon_t) = 0, \forall s < t. \end{cases} \tag{3}$$

该模型基本假设是  $AR(p)$  模型基本假设和  $MA(q)$  模型基本假设的结合.

4) 求和自回归移动平均模型(ARIMA 模型).一般的求和自回归移动平均模型记为  $ARIMA(p, d, q)$ , 它有如下结构:

$$\begin{cases} \Phi(B) \nabla^d x_t = \Theta(B) \varepsilon_t, \\ E(\varepsilon_t) = 0, Var(\varepsilon_t) = \sigma_\varepsilon^2, E(\varepsilon_t \varepsilon_s) = 0, s \neq t, \\ E(x_s \varepsilon_t) = 0, \forall s < t, \end{cases} \tag{4}$$

其中  $\nabla^d = (1 - B)^d$ ,  $\Phi(B) = 1 - \varphi_1 B - \cdots - \varphi_p B^p$ ,  $\Theta(B) = 1 - \theta_1 B - \cdots - \theta_q B^q$ .

当  $d=0$  时,  $ARIMA(p, d, q)$  模型实质上就是  $ARMA(p, q)$  模型, 因此, 差分运算结合  $ARMA$  模型其实就是  $ARIMA$  模型.

3.2 数据预处理 建立时间序列模型之前, 首先要对人口出生率序列进行平稳性和纯随机性检验.

1) 平稳性检验及平稳化处理. 首先画出原始数据的时序图, 如图 2 所示.

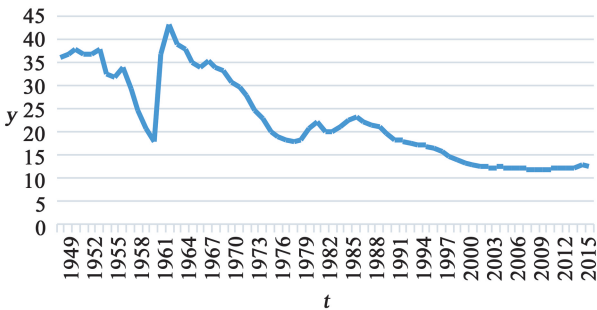


图 2 1949-2017 全国人口出生率时序图

由图 2 可以明显看出原始数据序列为非平稳序列,但是为了保证结果的客观性,还需对全国人口出生率采用 Eviews 软件进行单位根检验,判断其是否平稳,结果显示单位根统计量  $ADF = -1.660\ 242$ ,大于 Eviews 软件给出的显著性水平(1%~10%)对应的 ADF 临界值,这就说明该序列是非平稳的.为了可以对序列进行分析,要采取措施使其平稳化.让序列变平稳的方法有两种,分别是对数法和差分法.应用对数法可以绘制出 1949 年到 2017 年全国人口出生率自然对数的时序图(图 3).

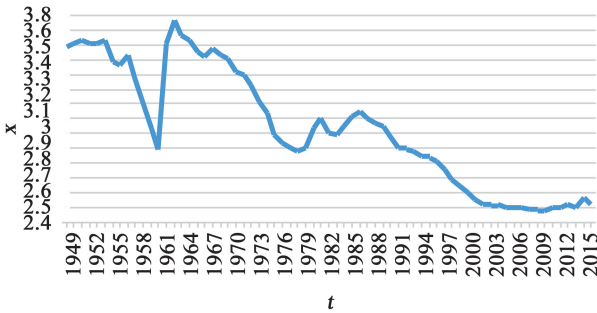


图 3 1949-2017 全国人口出生率取对数时序图

图 3 显示取对数之后的序列保持了原序列的变化趋势,运用 Eviews 软件对取对数之后的人口出生率序列进行单位根检验,结果显示单位根统计量  $ADF = -1.147\ 253$ ,大于 Eviews 软件给出的显著性水平(1%~10%)对应的 ADF 临界值,这说明该序列仍是非平稳的.上述全国人口出生率对数序列含有曲线趋势,通常二阶差分就可以提取出曲线趋势的特征值,因此接下来对原始数据取对数后的序列进行二阶差分,检验其是否平稳(图 4).

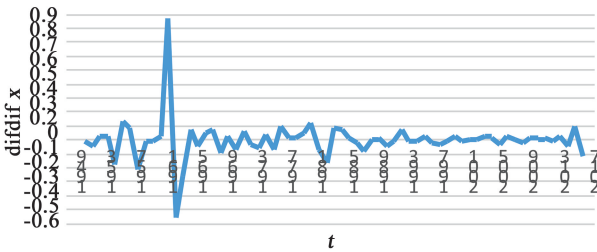


图 4 1949-2017 全国人口出生率取对数后二阶差分时序图

由图 4 可以看出,该序列始终在 0 点左右随机波动,并且波动的范围有界,因此能够确定原始数据取对数再进行二阶差分之后的序列平稳.但是这样判断数据的平稳性有主观色彩,为了保证客观,在此采用 Eviews 软件对取对数后二阶差分的数据进行单位根检验,判断其是否平稳.结果显示单位根统计量  $ADF = -7.921\ 190$ ,小于 Eviews 软件给出的显著性水平(1%~10%)对应的 ADF 临界值,这就说明该序列是平稳的.因此,应该建立 ARIMA 模型,且差分的阶数  $d$  等于 2.

2) 纯随机性检验.在将数据平稳化之后,为了确定平稳序列是否值得继续分析,还需要对得到的平稳序列进行纯随机性检验.纯随机性检验也叫白噪声检验,它是用来检验平稳序列是否为纯随机序列的一种方法,若得到的平稳序列为非白噪声序列,则可以进行模型拟合.

- 纯随机性检验的原假设:延迟期数小于或等于  $m$  期的序列值之间相互独立.
- 纯随机性检验的备择假设:延迟期数小于或等于  $m$  期的序列值之间有相关性.
- 纯随机性检验的检验统计量为

$$LB = n(n + 2) \sum_{k=1}^m \left( \frac{\hat{\rho}_k^2}{n - k} \right) \sim \chi^2(m).$$

(5)

当  $p > \alpha$  时,不能拒绝原假设,需要停止对该序列的统计分析.

$P < \alpha$  时,拒绝原假设,可以继续进行建模.

取置信水平  $\alpha = 0.05$ ,采用 SAS 软件进行白噪声检验<sup>[5,6]</sup>,其结果如表 2 所示.

表 2 白噪声的自相关检验									
延迟阶数	卡方检验	DF	P 值		自相关				
6	25.39	6	0.000 3	0.475	-0.163	-0.269	-0.167	0.048	0.053
12	29.82	12	0.003 0	-0.123	-0.126	-0.082	-0.082	-0.063	-0.085

由表 2 显示的纯随机性检验的结果可以看出,LB 统计量的  $P$  值小于置信水平  $\alpha = 0.05$ ,因此可以判定取对数再二阶差分后的序列是非白噪声序列,可以利用该序列值进行模型拟合.

3.3 模型的识别 通过观察平稳序列的自相关结果和偏自相关结果来选择阶数适当的模型进行拟合,用 SAS 软件可以得到序列的自相关结果和偏自相关结果(表 3、表 4).

表 3 自相关																								
延迟	协方差	相关系数	-1	9	8	7	6	5	4	3	2	1	0	1	2	3	4	5	6	7	8	9	1	标准差
0	0.028 129	1.000 00												*	*	*	*	*	*	*	*	*	*	0
1	0.013 374	0.475 44												*	*	*	*	*	*					0.122 169
2	-0.004 590	-0.163 27									.	*	*			.								0.147 217
3	-0.007 570	-0.269 05								.	*	*	*				.							0.149 896
4	-0.004 700	-0.167 01								.		*	*				.							0.156 938
5	0.001 350	0.047 98								.				*			.							0.159 569
6	0.001 487	0.052 84								.				*			.							0.159 784
7	-0.003 460	-0.123 03								.		*	*				.							0.160 045
8	-0.003 550	-0.126 23								.		*	*				.							0.161 450
9	-0.002 310	-0.082 30								.		*	*				.							0.162 917
10	-0.002 310	-0.082 25								.		*	*				.							0.163 536

表 4 偏自相关																						
延迟	相关系数	-1	9	8	7	6	5	4	3	2	1	0	1	2	3	4	5	6	7	8	9	1
1	0.475 44								.			*	*	*	*	*	*					
2	-0.503 02						*	*	*	*	*	*				.						
3	0.150 10								.				*	*	*	.						
4	-0.221 97								.	*	*	*				.						
5	0.267 16								.				*	*	*	.						
6	-0.373 98						*	*	*	*	*	*				.						
7	0.153 71								.				*	*	*	.						
8	-0.195 65								.	*	*	*				.						
9	0.022 60								.							.						
10	-0.203 31								.	*	*	*	*			.						

从表 3 中可以看出,除了延迟 0 阶和 1 阶的自相关系数在 2 倍标准差范围之外,其余阶数的自相关系数都在 2 倍标准差范围内来回波动.依据自相关系数的这个特点可以判断该序列有短期相关性,从而进一步说明该序列是平稳的.同时,可以认为该序列自相关系数为 2 阶截尾.表 4 中偏自相关系数呈现出拖尾的性质.

结合自相关系数和偏自相关系数的性质,初步为拟合模型定阶为  $MA(2)$ ,建立的模型为  $ARIMA(0,2,2)$ ,模型公式如下:

$$\begin{cases} \Phi(B) \nabla^2 x_t = \Theta(B) \varepsilon_t, \\ E(\varepsilon_t) = 0, Var(\varepsilon_t) = \varepsilon_\varepsilon^2, E(\varepsilon_t \varepsilon_s) = 0, s \neq t, \\ E(x_s \varepsilon_t) = 0, \forall s < t, \end{cases} \quad (6)$$

其中,  $\nabla^2 = (1 - B)^2, \Phi(B) = 1, \Theta(B) = 1 - \theta_1 B - \theta_2 B^2$ .

3.4 参数估计 确定好模型的阶数之后,应该对拟合的模型进行参数估计,本文使用条件最小二乘估计方法对初步拟合的 ARIMA(0,2,2)模型进行参数估计(表 5).

表 5 条件最小二乘估计					
参数	估计	误差	<i>t</i> 值	<i>P</i> 值	延迟
MU	-0.005 599	0.027 75	-0.20	0.840 7	0
MA(1,1)	-1.223 130	0.122 30	-10.00	<0.000 1	1
MA(1,2)	-0.245 250	0.117 88	-2.08	0.041 5	2

由表 5 可以看出 MU 不显著,而其他参数都是显著的.接下来要去掉常数项再次估计未知参数,其结果如表 6 所示,由表 6 可以看出两个未知参数均显著.

表 6 条件最小二乘估计					
参数	估计	标准误差	<i>t</i> 值	<i>P</i> 值	延迟
MA(1,1)	-1.231 29	0.119 97	-10.26	<0.000 1	1
MA(1,2)	-0.250 16	0.116 64	-2.14	0.035 7	2

3.5 模型检验 模型的显著性检验主要检验模型整体的有效性.一个好的拟合模型可以提取观察值序列中几乎所有与样本相关的信息,也就是说残差序列应该为白噪声序列.因此模型的显著性检验也被称为残差序列的白噪声检验.

模型检验原假设: $\rho_1 = \rho_2 = \cdots = \rho_m = 0, \forall m \geq 1$ .  
模型检验备择假设:至少存在某个  $\rho_k \neq 0, \forall m \geq 1, k \leq m$ .  
模型检验统计量为

$$LB = n(n + 2) \sum_{k=1}^m \left( \frac{\hat{\rho}_k^2}{n - k} \right) \sim \chi^2(m).$$

(7)

假如拒绝原假设,这说明残差序列中还留有相关信息,拟合的模型 ARIMA(0,2,2)不显著,模型的显著性检验不通过;假如不能拒绝原假设,就认为该拟合模型是显著有效的,模型的显著性检验通过.模型的显著性检验结果如表 7 所示.

表 7 残差白噪声检验									
延迟阶数	卡方	DF	<i>P</i>	自相关					
6	6.57	4	0.160 5	-0.046	-0.156	-0.013	-0.178	0.124	0.128
12	9.56	10	0.480 1	-0.185	0.040	-0.030	-0.036	0.004	-0.022
18	11.10	16	0.803 4	-0.032	-0.057	0.014	0.058	-0.007	0.096
24	15.49	22	0.840 3	0.135	0.038	-0.115	0.015	0.047	0.092

从表 7 可以看出,延迟 6 阶、12 阶、18 阶、24 阶的 LB 统计量的 *P* 值均大于  $\alpha = 0.05$ ,可知残差序列通过了白噪声检验,也就是说残差序列为白噪声序列.该拟合模型 ARIMA(0,2,2)显著成立,可以表示为

$$\nabla^2 \ln x_t = (1 + 1.231\ 29B + 0.250\ 16B^2) \varepsilon_t.$$

(8)

3.6 模型的优化 当拟合模型为 ARIMA(0,2,2)时,AIC 的值为-100.27,对模型重新进行定阶,尝试是否有其他合适的模型 AIC 的值小于-100.27.经过尝试后,发现当拟合模型为 ARIMA(1,2,2)时,AIC 的值更小,为-100.793,根据 AIC 准则,更优的模型为 ARIMA(1,2,2).当拟合模型为 ARIMA(1,2,2)时,其参数估计的结果如表 8 所示.由表 8 中的 *P* 值可以看出参数估计均显著,其模型检验结果如表 9 所示.

表 8 优化后条件最小二乘估计					
参数	估计	误差	<i>t</i> 值	<i>P</i> 值	延迟
MA(1,1)	-1.692 47	0.136 21	-12.43	<0.000 1	1
MA(1,2)	-0.708 18	0.125 35	-5.65	<0.000 1	2
AR(1,1)	-0.457 84	0.197 90	-2.31	0.023 9	1



表 9 优化后残差白噪声检验

延迟阶数	卡方	DF	P	自相关					
6	4.10	3	0.251 0	-0.031	-0.077	-0.084	-0.136	0.100	0.114
12	6.40	9	0.699 7	-0.155	0.018	-0.029	-0.045	0.013	-0.045
18	8.07	15	0.921 0	-0.016	-0.075	0.030	0.038	0.025	0.097
24	12.39	21	0.928 6	0.143	0.037	-0.098	0.009	0.054	0.094

由表 9 可以看出残差序列为白噪声序列,最终拟合模型 ARIMA(1,2,2) 显著成立.该模型得到的拟合结果的输出形式可以表示为

$$(1 + 0.457\ 84B)\nabla^2\ln x_t = (1 + 1.692\ 47B + 0.708\ 18B^2)\varepsilon_t.$$

(9)

4 预测结果

采用建立的 ARIMA(1,2,2)模型对未来 5 年的全国人口出生率进行预测<sup>[7-11]</sup>,结果如表 10 所示.

表 10 预测结果

	预测值	标准差	95%的置信区间	
70	2.498 1	0.111 6	2.279 4	2.716 8
71	2.504 2	0.177 3	2.156 7	2.851 7
72	2.505 3	0.218 4	2.077 3	2.933 4
73	2.500 9	0.254 4	2.002 3	2.999 5
74	2.506 9	0.286 1	1.946 2	3.067 6

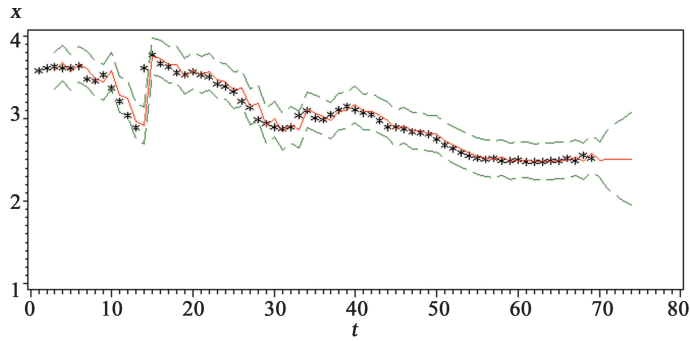


图 5 拟合与预测图

图 5 中黑色星号表示全国人口出生率对数序列观察值,红色连续曲线表示拟合序列曲线,绿色曲线表示拟合序列的 95%上下置信限.根据图 5 可以初步判断模型的拟合效果较好,接下来利用 2018 年的真实数据与预测数据进行比较,进一步验证模型的有效性.

本文以 1949–2017 年的数据为基础建立时间序列模型,预测出了 2018–2022 年的全国人口出生率,其中预测的 2018 年人口出生率对数为 2.498 1,它的 95%置信区间为(2.279 4,2.716 8),而 2018 年人口出生率的真实值为 10.94‰,取对数后为 2.39,真实值在置信区间范围内,证明模型拟合效果良好,该模型得到的数据可信.

由于建立 ARIMA(1,2,2)模型时采用的序列是原始数据取对数之后的数据,因此得到的预测结果也是取对数之后的形式,还需要对得到的结果进行变换,才能最终得到未来五年的全国人口出生率(表 11).

表 11 未来五年全国人口出生率预测值

时间	全国人口出生率对数	全国人口出生率(‰)
2018	2.498 1	12.159 369 20
2019	2.504 2	12.233 768 04
2020	2.505 3	12.247 232 58
2021	2.500 9	12.193 463 14
2022	2.506 9	12.266 843 84

5 结论与建议

通过对 1949–2017 年全国人口出生率的原始数据进行时间序列分析,建立了 ARIMA(1,2,2)模型,对未来五年的全国人口出生率进行预测,得到未来五年全国人口出生率分别为 12.16‰、12.23‰、12.25‰、12.19‰、12.27‰.从数据可以看出 2016 年和 2017 年人口出生率确实是有所提高,但是预测结果显示未来 5 年仍然保持低人口出生率,这个结果与引言中二胎政策放开使得 2016 年和 2017 年出生率短暂提高是一致的.

- 针对人口出生率维持较低水平这一现状,提出以下建议:
- 1) 完善二胎政策.全面实施二胎政策的计划刚刚起步,仍存在许多没有考虑到的细节.因此要积极完善二胎政策,既不能造成大众都不响应该政策,又不能导致出生率突然增大影响国家其他方面的发展.
  - 2) 建立健全社会保障制度.比如为老人开办养老保险;为生育二胎的家庭提供补助;加强对公共托幼服务的职业培训和政府监督等.
  - 3) 为女性提供更多就业支持.女性担心生育之后找工作难度加大,因此为女性提供更多就业支持有助于提高人口出生率,就业环境越宽松,女性才能越放心地去生育.
  - 4) 人口政策应该致力于改变趋势性的斜率.本文所建立的二阶差分模型表示人口出生率带有与时间相关联的趋势性,二阶差分之后所获得的时间序列实际上描述了出生率趋势之斜率随时间的变动情况.具体来说,如果希望人口增长提速(如用来弥补劳动力短缺),应该维持趋势性的斜率为正来促使人口加速增长.

6 参考文献

[ 1 ] 刘雨婷,刘艳琪,邓智年,等.“全面二孩”政策下的人口数量与结构预测[J]. 湖南文理学院学报:自然科学版, 2018,30(4):17–21.

[ 2 ] 王 燕.应用时间序列分析(第四版)[M].北京:中国人民大学出版社,2016:110–121.

[ 3 ] 刘敬敏,丁维勇,贾会涛.全面放开二胎政策对人口出生率的影响研究——以唐山市为例[J].当代经济,2016,(13):100–101.

[ 4 ] 特伦斯·C·米尔斯.金融时间序列的经济计量学模型(中译本)[M]. 北京:经济科学出版社,2002:31–67.

[ 5 ] 郭 敏.我国人口出生率预测——基于 ARIMA 模型分析[D].昆明:云南财经大学,2018.

[ 6 ] 高惠璇.SAS 系统 SAS/ETS 软件使用手册[M]. 北京:中国统计出版社,1998:65–116.

[ 7 ] 万兆泉.关于国内生产总值的时间序列分析及预测[D].南昌:江西财经大学,2016.

[ 8 ] 徐雅静,汪远征.ARIMA 模型在河南省 GDP 预测中的应用及 SAS 实现[J].中国科技信息,2006,(10):216–219.

[ 9 ] 周正柱,王俊龙.长江经济带区域生态环境质量综合评价与预测研究[J].山东师范大学学报:自然科学版,2018,33(4):474–480.

[ 10 ] 陈小玲.基于 ARIMA 模型与神经网络模型的股价预测[J]. 经济数学,2017,34(4):30–34.

[ 11 ] 郝 琳.我国人口老龄化问题研究及对策[D]. 曲阜:曲阜师范大学,2014.



ANALYSIS AND PREDICTION OF BIRTH RATE IN CHINA

Sun Xinxin<sup>1)</sup>      Gao Yan<sup>1)</sup>      Yuan Xi<sup>1)</sup>      Wang Zengxiang<sup>2)</sup>

( 1 ) School of Mathematics and Statistics, Shandong Normal University, 250358, Jinan, China;  
2) Athabasca University Faculty of Business, T9S3A3, Alberta, Canada )

**Abstract** The birth rate in our country has gone through many stages, especially in recent years, China has kept a low birth rate trend. In order to make a reasonable prediction of the future birth rate of China, after steps of the data preprocessing, model identification, parameter estimation and model validation, and optimization, we first developed an ARIMA ( 1, 2, 2 ) model for the birth rate best fitted for the years of 1949 to 2017, then applied the model to predict the birth rates in China for the five years of 2018 to 2022. The prediction results suggest that the birth rate will remain low for the next five years. We concluded with identified policy variables for birth rate management.

**Key words** birth rate; time series; ARIMA model; Eviews; statistical analysis system

( 上接第 255 页 )

COMPACT DIFFERENCE SCHEME FOR  
ONE-DIMENSIONAL FRACTIONAL CATTANEO EQUATION

Wang Xiaohan      Jiang Ziwen

( School of Mathematics and Statistics, Shandong Normal University, 250358, Jinan, China )

**Abstract** The Fick law is adopted as basic model for the description of heat and mass transfer. But the validity of using this law has been questioned extensively. The main reason is that this law implies an infinite velocity of propagation. In order to consider the finite velocity of propagation, several new constitutive models have been introduced, for example, the widely used Cattaneo model. Recently, Compte and Metzler generalized the Cattaneo model to the time fractional models, which was supported by three different scheme: continuous time random walks, non-local transport theory, and delayed flux-force relation. In this paper, a compact finite difference method is proposed for the one-dimensional generalized fractional Cattaneo equation, which adopts L1 approximation for the time fractional derivative and compact approximation for the space integer derivative. A numerical example shows that the convergence order of the scheme is  $o(\tau^{2-\alpha} + h^4)$ .

**Key words** fractional Cattaneo equation; L1 approximation; compact difference scheme