

基于 ARIMA 模型对浙江省农村常住居民人均生活消费支出的分析

文 / 蔡永

本文对浙江省 1980 年至 2021 年的农村常住居民人均生活消费支出的数据进行分析，建立 ARIMA 模型，并利用该模型对未来 3 年的农村常住居民人均生活消费支出进行预测。结合实验结果，有助于深入了解浙江省农村居民的消费行为和经济发展趋势，为相关政策的制定提供参考。

关键字：农村居民家庭人均生活消费支出；ARIMA 模型；预测

一、引言

农村居民家庭生活消费支出指农村常住居民家庭用于日常生活的全部开支，是反映和研究农民家庭实际生活消费水平高低的重要指标。在我国，带动农村的发展对促进全国经济的整体发展有重要意义^[1]。近年来，随着我国社会经济的快速发展，农村地区的经济和社会发展也取得了显著成效。然而，在经济和社会发展的同时，农村地区也面临着许多问题。例如，研究表明，财政支农支出对浙江省农村居民消费水平有正向推动作用，财政支农支出对浙江省农村居民消费水平的赋能效应正在逐步减弱^[2]。浙江省城镇居民与农村居民消费项目的差异食品消费支出在农村居民和城镇居民之间占的比重最大，农村居民收入及消费均落后于浙江省经济发展，消费结构呈现城乡二元结构^[3]。人均生活消费支出是反映农村居民经济水平的重要指标，也是反映农村居民生活质量的重要指标。因此，对于农村常住居民人均生活消费支出的研究具有重要的实际意义。

二、ARIMA(p,d,q) 模型的结构

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

这样的结构模型也被称为求和自回归移动平均模型，其中 $\nabla^d = (1-B)^d$, $\Phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$ ，属于平稳可逆 ARMA(p,q) 模型的自回归系数多项式。 $\Theta(B) = 1 - \theta_1 B - \dots - \theta_q B^q$ ，属于平稳可逆 ARMA(p,q) 模型的移动平均系数多项式。模型中 $\nabla^d = (1-B)^d x_t = \sum_{i=0}^d (-1)^i C_d^i x_{t-i}$ ，d 表示 d 阶差分。

该模型具有以下优点：（1）不需要借助事物发展的因果关系去分析这个事物过去的和未来的联系；（2）具有很少的信息量，对趋势性、随机性和相关性的数据都能有很好的预测结果；（3）该模型的结构比较简单，预测原理也很容易理解，只需要内生变量而不需要借助一些其他类型的外生变量。基于此，本文利用了 ARMA 模型预测了未来 3 年浙江省农村常住居民人均生活消费支出进行预测，并结合研究结果，给予相关政策的制定和实施提供建议。



蔡永，男，汉族，江西景德镇人，硕士研究生，讲师。单位：景德镇学院信息工程学院（江西景德镇，333000）。研究方向：应用数学。

三、浙江省农村常住居民人均生活消费支出的预测

本文所使用的数据来源于浙江省统计年鉴官方网站，选取了 1980—2022 年浙江省农村常住居民人均生活消费支出改变值序列。

(一) 数据的预处理

对于序列的稳定性，有两种检验方法，一种是基于时序图进行判定的图示检查；二是建立检验统计量，并进行假设检验。根据 Carmer 分解定理，d 阶差分运算可以将序列中的有效确定性信息充分提取。对原始序列、一阶差分后序列和二阶差分序列作 ADF 检验，如表 1 所示，可以看出，原始序列和一阶差分序列都是非平稳的，二阶差分序列是平稳的。

表 1 序列 ADF 检验结果的 P 值

类型	滞后阶数	原始序列	一阶差分序列	二阶差分序列
No drift no trend	1	0.99	0.66	0.01
	2	0.99	0.89	0.01
	3	0.99	0.80	0.01
No drift no trend	1	0.99	0.73	0.01
	2	0.99	0.94	0.01
	3	0.99	0.89	0.01
No drift no trend	1	0.99	0.09	0.01
	2	0.99	0.54	0.03
	3	0.99	0.32	0.04

验证二阶差分序列中有用的信息已经被提取完毕，剩下的全是随机扰动，是无法预测和使用的。下面，对序列进行白噪声检验，检验所得结果如表 2 所示，可以看出在滞后阶数为 6 和 12 时，检验统计量的 P 值分别为 0.02514 和 0.001876，均小于 0.05，说明二阶差分序列为非白噪声序列。

表 2 二阶差分序列的白噪声检验

滞后阶数	检验统计量	P 值
6	11.13	0.02514
12	24.517	0.001876

(二) 建立模型

为了选择合适的模型拟合原始序列，下面给出原始数据的二阶差分序列的自相关图与偏自相关图，结果如图 1 所示。

图 1 (a) 为原始序列的自相关图，具有二阶截尾性，根据平稳时间序列模型的性质，可选用 ARIMA(0, 2, 2) 模型拟合原始序列。图 1(b) 为原始序列的偏自相关图，具有二阶截尾性，根据平稳时间序列模型的性质，可选用 ARIMA(2, 2, 0) 模型拟合原始序列。两个模型的参数估计值、检验 P 值、AIC 值和 BIC 值，结果如表 3 所示。

表 3 模型参数估计值、检验 P 值、AIC 值和 BIC 值

模型	参数	参数估计值	参数检验 P 值	AIC 值	BIC 值
ARIMA(0,2,2)	ma1	-0.9568192	9.274496e-07	631.7695	636.8361
	ma2	0.2646802	2.735276e-02		
ARIMA(2,2,0)	ar1	5.716441	5.916093e-07	627.9939	633.0605
	ar2	2.995902	2.340551e-03		

从表 3 可以看出，两个模型中，各参数估计值的检验 P 值均小于 0.05，故参数显著非零。根据 AIC 值和 BIC 值，本文所拟合的模型为 ARIMA(2, 2, 0)。

(三) 模型检验

为了验证本文拟合模型的有效性，给出模型 ARIMA(2, 2, 0) 的残差序列的自相关检验图、偏自相关检验图、

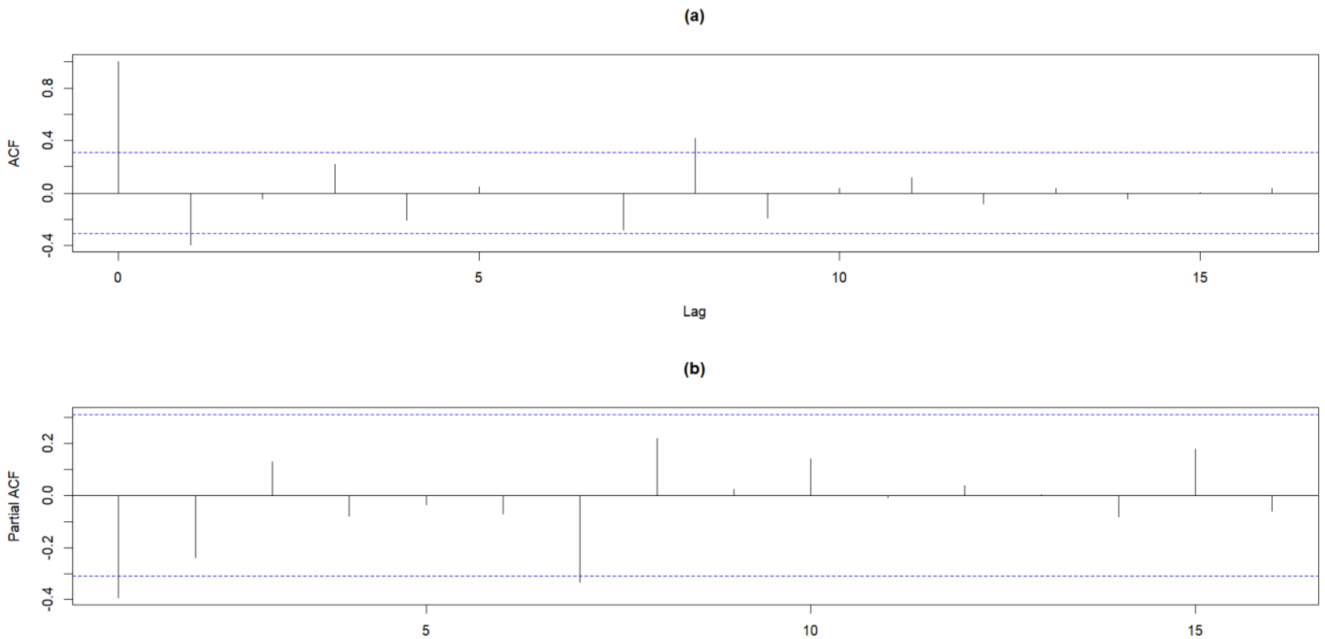


图 1 自相关图与偏自相关图

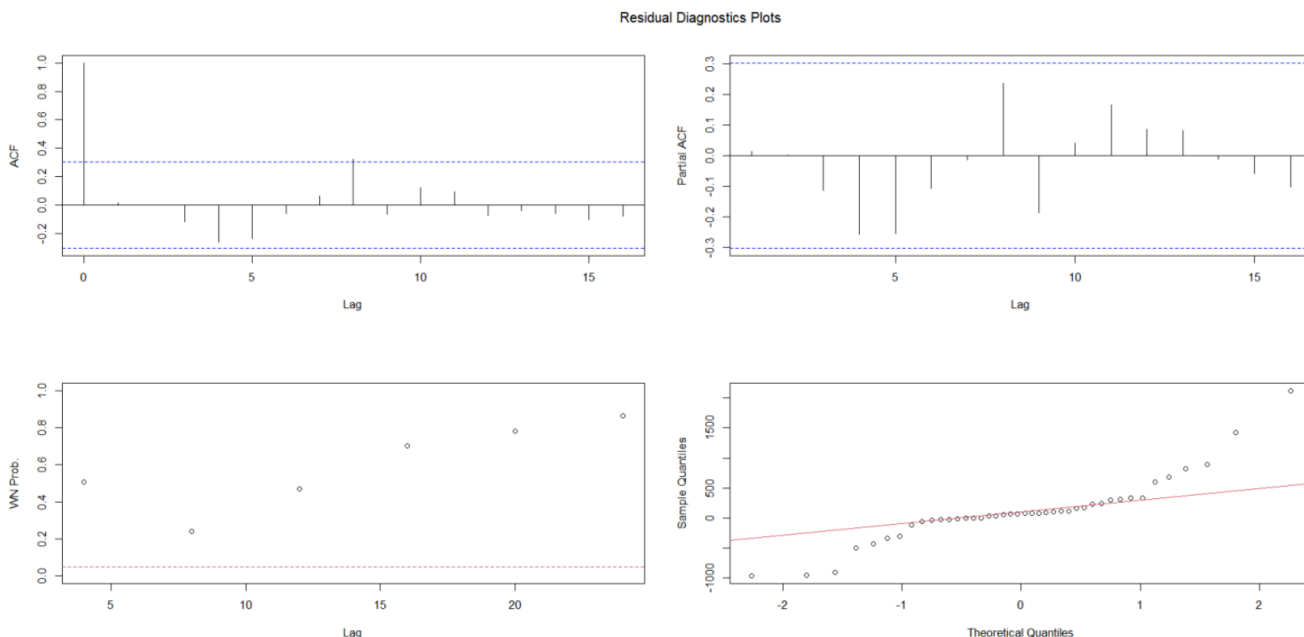


图2 ARIMA(2,2,0)模型的检验图

表4 预测值及95%预测区间

时间	预测值	Lo.80	Hi.80	Lo.95	Hi.95
2023	26162.90484	25437.88905	26887.92064	25054.08887	27271.72082
2024	28166.38664	27179.75339	29153.01989	26657.4613	29675.31198
2025	30601.45354	29168.81684	32034.09025	28410.42478	32792.48231

白噪声检验图和QQ检验图，如图2所示。可以看出，ARIMA(2,2,0)模型效果拟合良好，即可以用该模型进行预测。

因此，得到浙江省农村居民生活消费支出改变值序列的ARIMA(2,2,0)模型为：

$$(1+1.077B)(1+0.5732B)\nabla^2 x_t = \varepsilon_t, \varepsilon_t \sim N(0, 320054)$$

(四) 模型预测

使用该模型预测未来3年浙江省农村常住居民生活消费支出的改变值及其80%和95%预测区间，如表4所示。

预测图如图3所示，其中，虚线代表1980—2022年浙江省农村常住居民生活消费支出改变值序列的历史观测值，黑色实线表示该序列未来3年的预测值，浅色阴影区域和深色阴影区域分别表示预测值的80%和95%预测区间。

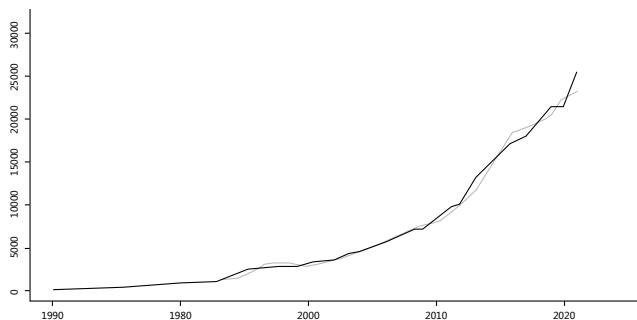


图3 浙江省农村居民生活消费支出的改变值的预测图

四、结论

浙江省农村常住居民生活消费支出的变化是由许多复杂的因素造成的。通过将1980—2022年浙江省农村常住居民生活消费支出的改变量作为样本，本文基于ARIMA模型预测出了2023—2025年浙江省农村常住居民生活消费支出改变量，由实验结果可知，近3年的浙江省农村常住居民生活消费支出逐步递增。但改变量不是很大，依然有一定的提升空间，实验结果也为能够制定合理的增加农村常住居民生活消费支出的政策，提供一定的经济参考依据。

参考文献：

- [1] 齐心怡，高青松．使用人均收入作为衡量人民生活水平指标之反思[J]．湖南工业大学学报（社会科学版），2021，26（05）：29-37.
- [2] 徐舒琪，王国梁．财政支农支出能有效提升农村居民消费水平吗？——来自浙江省的经验证据[J]．中国集体经济，2023（31）：22-24.
- [3] 纪岩峰．城镇、农村居民消费结构现状分析[J]．商，2015（09）：60.