

基于 ARIMA 模型对江西省 CPI 的预测研究

杨芷熠

江西财经大学国际学院, 中国·上海 330000

【摘要】 CPI (居民消费价格指数) 一方面可表明消费者的购买能力, 另一方面也是观察通货膨胀水平的重要指标。对 CPI 的科学准确预测可以帮助有关部门更好地把握宏观经济形势。本文通过国家统计局收集江西省 2016 年 1 月至 2020 年 2 月的 CPI 月度数据, 借助 R 语言, 对江西省 CPI 月度数据建立 ARIMA (12, 1, 12) 模型, 从而对江西省未来四期 CPI 进行预测。结果表明, 江西省未来四期 CPI 将会持续上涨, 相关部门应尽快采取措施进行宏观调控。

【关键词】 CPI 预测; ARIMA 模型; R 语言

引言

居民消费价格指数 (Consumer Price Index, 简称 CPI), 是反映居民家庭所购买的生活消费商品的价格和服务的价格水平, 随时间变动趋势和变动幅度的相对数^[1]。对 CPI 合理科学和准确地预测可以帮助有关部门和机构更好地把握宏观经济的发展趋势, 从而制定相应的政策和采取精准的调控措施。而 CPI 这一经济指标的影响因素较多, 且已有较多的研究来探讨其影响因素, 但是总有一些影响因素尚且存在争议或者可能仍有潜在因素也会影响 CPI, 加之我们要对其进行预测, 因此, 我们可以建立 ARIMA 模型, 一方面 ARIMA 模型涵盖了 CPI 的影响因素, 包括已有影响因素和潜在影响因素; 另一方面, 我们可以通过 ARIMA 模型对 CPI 进行预测。

1 文献综述

不少文献做过 CPI 影响因素的研究。昆鹏在《对于影响 CPI 因素的研究》中分析了食品类消费价格指数和工业生产出厂价格指数的变动对于居民消费价格指数波动的影响程度^[2]。杜一平在《PI 的影响因素分析》中提出广义货币供应量和外汇储备是 CPI 的影响因素^[3]。

邹劲松在《基于 ARIMA 模型的经济形势预测分析——以重庆市 GDP 为例》中用 ARIMA 模型对我国 2000 年至 2015 年共 16 年的 GDP 序列数据进行预测分析^[4]。杨志在《居民消费价格指数的时间序列分析》中指出对于季节性、月度性的密集变化时间序列, 用 ARIMA 模型的预测分析更加具有准确性^[5]。对已有研究总结, 笔者发现时间序列预测模型中较为经典的模型是 ARIMA 模型。

2 数据来源

本文通过国家统计局采集数据样本, 收集了江西省 2016 年 1 月至 2020 年 2 月的 CPI 同比价格月度数据。

3 模型介绍

ARIMA (p, d, q) 是求和自回归移动平均 (Auto Regressive Integrated Moving Average) 模型的简写形式, 用于平稳序列或通过差分而平稳的序列分析。显然, 江西省月度 CPI 是时间序列数据, 但是我们不能确定其是否为平稳时间序列, 于是我们可采用 ARIMA (p, d, q) 模型构建预测模型。

ARIMA 模型:

$$\mu_t = c\phi_1\mu_{t-1} + \dots + \phi_p\mu_{t-p} + \theta_1\varepsilon_{t-1} + \dots + \theta_q\varepsilon_{t-q}$$

其中, c 是常数, ϕ 是自回归模型系数, θ 是移动平均模型系数, ε_t 是白噪声序列。p 为自回归模型阶数, q 为移动平均模型阶数^[6]。

构建 ARIMA (p, d, q) 模型以及预测总共分如下五步。

第一步, 确定适当的 d 值, 即差分的次数, 以使得原序列转变为平稳性时间序列。这也是 ARIMA (p, d, q) 模型与 ARMA (p, q) 模型的区别, 即对一般时间序列而言, 往往不满足平稳性, 而差分后的序列可以是平稳的, 从而再过度到 ARMA (p, q) 模型。

第二步, 平稳性检验。检验方法主要有如下三种方法。

(1) 时序图检验。平稳时间序列的性质之一是其均值和方差都是常数, 因此对平稳时间序列而言, 其序列值应该始终在某一常数附近随机波动, 并且序列值波动的范围是有界的。如若某时间序列时序图具有明显的波动趋势或具有周期性, 其往往是非平稳时间序列。

(2) 自相关图检验。平稳时间序列另一性质是一般具有短期相关性, 因此对平稳时间序列而言, 相邻几期的序列值相关性通常较大, 而间隔时期较长的序列值之间相关性通常较小。另外, 平稳时间序列的自相关系数会随着延迟期数的增加逐步衰减趋向于零, 且趋于零的速度较快, 最后保持在零上下随机波动。对非平稳时间序列而言, 其自相关系数也可能逐渐衰减, 但其趋于零的速度通常很慢。还有需要说明的是, 平稳时间序列在延迟期数较小时, 其自相关系数可以很大, 但是随着延迟期数不断增加, 其自相关系数应该迅速衰减并趋于零。

(3) 单位根检验。据平稳性时间序列不存在单位根的性质, 可以检验某一时间序列是否存在单位根, 若存在, 则为非平稳时间序列。用的最多的检验方法就是扩展 Dickey-Fuller (ADF) 检验, 其创建的回归模型如下。

$$\nabla Y_t = \phi Y_t + \beta_1 Y_{t-1} + \dots + \beta_{k-1} Y_{t-(k-1)} + \varepsilon_t$$

其中, k 代表模型中能容许的最大时间延迟数量, 第一项来自于 AR 模型, 最后一项来自于 MA 模型。

ADF 原假设 (H0): 时间序列非平稳——模型回归系数 ϕ 约为零。

ADF 备择假设 (H1): 时间序列是平稳的——模型回归系数 ϕ 小于 0。

第三步, 将原序列转变为平稳序列后, 建立模型拟合样本数据。

第四步, 对模型进行评价。对模型合理性进行评价也即模型诊断, 即检验模型是否充分提取序列的信息进行回归, 这可通过检验残差序列是否为白噪声序列来说明。

第五步, 模型预测。通过 ARMA 模型的拟合结果给出的参数估计, 得到回归方程, 进一步, 我们利用回归方程对未来数据进行

拟合从而得到预测结果，同时给出预测结果的置信区间。

4 模型构建

4.1 平稳性检验

该时间序列具有波动性，且2019年至2010年CPI呈上升趋势。为探究其是否为平稳序列，需要进行单位根检验进一步确定。

原序列P值远大于0.05，应该接受原假设，即认为该序列具有单位根，是非平稳时间序列。因此，我们需要对其差分处理。

4.2 一阶差分处理

一阶差分序列P值小于0.05，不能接受原假设，即不认为该序列具有单位根，是平稳时间序列。从而，一阶差分序列为平稳时间序列。

4.3 模型定阶

12阶之后函数值呈拖尾性，且趋于0，因此可将q取为12。由图3可知，12阶之后函数值呈拖尾性趋于0，因此可将p取为12。从而建立ARIMA(12, 1, 12)模型。

5 模型评价

残差序列在零附近波动，且方差大致不变。直方图显示残差大致是正态分布且均值接近0。因此，残差可看成是均值为零方差为常数的正态分布。这也说明序列样本中的信息基本被提取出来了，因此可认为ARIMA(12, 1, 12)模型是合理的。

6 模型预测

江西省未来四期(2020年3月至2020年6月)CPI的预测值依次约为105.16, 105.24, 105.53, 106.10。同时，我们给出预测值的置信区间。

对2020年3月CPI预测值105.16，其置信度为80%的置信区间为[104.77, 105.55]，置信度为95%的置信区间为[104.56, 105.76]。

对2020年4月CPI预测值105.24，其置信度为80%的置信区间为[104.63, 105.85]，置信度为95%的置信区间为[104.31, 106.18]。

对2020年5月CPI预测值105.53，其置信度为80%的置信区间为[104.76, 106.30]，置信度为95%的置信区间为[104.35, 106.71]。

对2020年6月CPI预测值106.10，其置信度为80%的置信区间为[105.21, 107.00]，置信度为95%的置信区间为[104.74, 107.47]。

同时我们给出江西省未来四期(2020年3月至2020年6月)CPI预测值的可视化结果。

Forecasts from ARIMA(12,1,12)

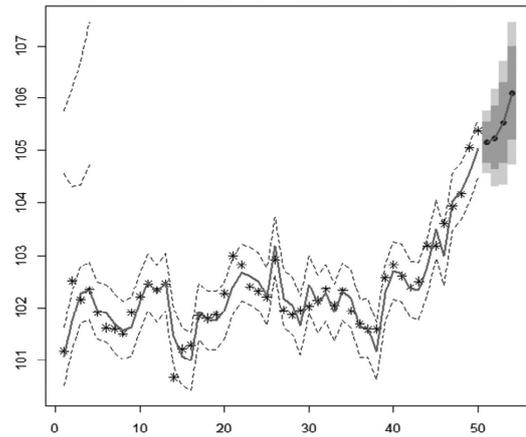


图1 江西省未来四期CPI的预测图

由图1可知，江西省未来四期(2020年3月至2020年6月)CPI有上涨趋势，这意味着通货膨胀的到来，日常购买的物品价格上涨，或者说居民的购买力下降，即相同的东西，在未来几期要花更多的钱才能买到。

7 结语

CPI(居民消费价格指数)是宏观经济中的重要指标之一，反映经济的景气状况，一方面它表明消费者的购买能力，另一方面它是观察通货膨胀水平的重要指标。因此，对其进行准确的预测可以帮助调控宏观经济，具有现实意义。本文运用ARIMA模型对江西省CPI进行预测，结果显示未来四期(2020年3月至2020年6月)的居民消费价格指数会持续上升，为应对江西省CPI上涨，有关部门应尽快制定相应的政策和措施进行宏观调控，以保证江西省居民生活水平和质量免于通货膨胀的影响。

参考文献:

- [1] 栾惠德. 居民消费价格指数的实时监测——基于季节调整的方法[J]. 经济科学, 2007, 29(2): 59-67.
- [2] 对于影响CPI因素的研究[J]. 魅力中国, 2014, (4).
- [3] CPI的影响因素分析[J]. 财讯, 2018, (30).
- [4] 邹劲松. 基于ARIMA模型的经济形势预测分析——以重庆市GDP为例[J]. 商场现代化, 2013(34): 185-187.
- [5] 杨志. 居民消费价格指数的时间序列分析[J]. 经济研究导刊, 2013(35): 8-11.
- [6] 吴连恕. 基于ARIMA模型的浙江省CPI预测研究[J]. 商情, 2019(27).