

# 基于 SARIMA 模型的美国 Core CPI 指数实证分析

姓名：杜冰      学号：1807403036

## 摘要

本文依据 1957 年至 2020 年的美国居民核心消费价格指数（Core CPI）数据，利用非平稳时间序列分析（SARIMA）构建 Core CPI 预测模型，并对结果进行实证分析。结果显示，该模型拟合效果比较理想，将 2021 年 1 月至 5 月数据的预测结果与真实值进行比较，发现绝对误差很小。最后，使用此模型对美国 2021 年 6 月至 2022 年 12 月的 Core CPI 数据进行了预测，从而为美国政府宏观经济调控提供参考。

关键词：SARIMA 模型; 核心居民消费价格指数; 宏观调控

## 1 研究背景和意义

新型冠状病毒肺炎疫情在全世界范围的大面积传播，对世界经济的发展产生重大冲击，以至各行各业大多进入停产停工阶段。为了更加直观快速地了解这次疫情对人民日常生活的影响，本文拟通过时间序列模型来分析居民核心消费价格指数（Core CPI）的变化趋势。

CPI 指数，全称为居民消费价格指数，是一个反映居民家庭一般所购买的消费商品和服务价格水平变动情况的宏观经济指标<sup>[1]</sup>。由于食品和能源价格会受到一些异常因素的影响而产生大幅波动，并模糊消费价格的真实变动趋势，因此，一般将剔除了食品和能源消费价格之后的居民消费价格指数作为核心价格指数（Core CPI），并认为其是衡量通货膨胀的最佳指标<sup>[2]</sup>。Core CPI 的高低直接影响着国家宏观经济调控措施的出台与力度，同时，Core CPI 指数的高低也间接影响资本市场的变化。因此，研究 Core CPI 指数是否稳定及其变化趋势具有重要意义。

本文将美国的 Core CPI 指数月度数据建立 SARIMA 模型，并验证模型的可靠性，进而对 2021 年 6 月至 2022 年 12 月的 Core CPI 指数进行预测与分析，分析美国 Core CPI 指数的变化情况。

## 2 理论介绍

基本的时间序列模型有自回归（AR）模型、移动平均（MA）模型、自回归滑动平均（ARMA）模型以及求和自回归滑动平均模型（ARIMA）等<sup>[3]</sup>。

ARIMA( $p, d, q$ ) 模型的原理是把非平稳时间序列转化为差分运算和平稳时间序列的结合，从而能够通过  $d$  阶差分转化为 ARMA 模型。因此，在使用此模型分析前，要对数据进行平稳性分析。而 ARMA 模型分为 AR（自回归模型）、MA（滑动平均模型）、ARMA（自回归滑动平均模型）3 类。

AR( $p$ ) 模型的表达式为  $X_t = a_0 + a_1X_{t-1} + a_2X_{t-2} + \cdots + a_pX_{t-p} + \varepsilon_t$ ，其中  $\varepsilon_t \sim WN(0, \sigma^2)$ ；MA( $q$ ) 模型的表达式为  $X_t = \mu + \varepsilon_t - b_1\varepsilon_{t-1} - b_2\varepsilon_{t-2} - \cdots - b_q\varepsilon_{t-q}$ ；ARMA( $p, q$ ) 模型的表达式为  $X_t = a_0 + a_1X_{t-1} + a_2X_{t-2} + \cdots + a_pX_{t-p} + \varepsilon_t - b_1\varepsilon_{t-1} - b_2\varepsilon_{t-2} - \cdots - b_q\varepsilon_{t-q}$ ，其中  $\varepsilon_t \sim WN(0, \sigma^2)$ 。

SARIMA 模型是在 ARIMA 模型中引入季节性的项，记为 ARIMA( $p, d, q$ )( $P, D, Q$ )<sup>S</sup>，其中  $p, d, q$  表示模型中非季节部分的模型参数， $P, D, Q$  表示模型中季节性部分的模型参数，S 为季节性周期。SARIMA 模型又根据季节  $D$  阶差分后序列是否平稳被细分为季节 ARIMA( $P, D, Q$ )<sup>S</sup> 模型和乘法季节 ARIMA( $p, d, q$ )( $P, D, Q$ )<sup>S</sup> 模型<sup>[4]</sup>。

ARIMA( $p, d, q$ )( $P, D, Q$ )<sup>S</sup> 模型的表达式为：

$$\phi(B)\Phi(B)\nabla^d\nabla_S^D X_t = \theta(B)\Theta(B)\varepsilon_t$$

其中  $\varepsilon_t \sim WN(0, \sigma^2)$ ,  $\nabla^d = (1 - B)^d$ ,  $\nabla_S^D = (1 - B^S)^D$ ,  $\phi(B) = 1 - \phi_1B - \cdots - \phi_pB^p$ ,  $\Phi(B) = 1 - \Phi_1B^S - \cdots - \Phi_PB^{SP}$ ,  $\theta(B) = 1 - \theta_1B - \cdots - \theta_qB^q$ ,  $\Theta(B) = 1 - \Theta_1B^S - \cdots - \Theta_QB^{SQ}$ 。

AR 或者 MA 模型中的季节性特征其实可以从偏自相关图和自相关图中的季节性延迟中被观察到<sup>[5]</sup>。比如，一个 ARIMA(0,0,0)(0,0,1)<sup>12</sup> 模型会显示出：自相关图中延迟期数为 12 上有一个明显突起，但是没有其他明显的突起；偏自相关图中季节性延迟（即当延迟为 12,24,36,⋯）上出现指数性的衰减。同样的，一个 ARIMA(0,0,0)(1,0,0)<sup>12</sup>：自相关图中季节性延迟上出现指数型衰减；偏自相关图中延迟为 12 上的一个明显突起。

目前 SARIMA 模型广泛应用于能源经济、气候环境和医药科学等领域，在农林和工业产业经济领域的预测精度也比较高，尤其针对如降雨量、流感病患数量等包含季节变化性质的研究有较好的表现，在国际上受到很多学者推崇。

### 3 SARIMA 模型在美国 Core CPI 指数预测中的应用

#### 3.1 数据来源

通过使用美联储经济数据库（FRED）平台<sup>[6]</sup>，我们获得了 1957 年 1 月至 2021 年 5 月美国 CPI 指数月度统计数据，部分数据如表1所示。其中 1957 年 1 月至 2020 年 12 月的数据作为训练集，2021 年 1 月至 2021 年 5 月的数据作为测试集。

表 1: 美国 Core CPI 指数月度统计数据 (部分)

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1957	28.5	28.6	28.7	28.8	28.8	28.9	29	29	29.1	29.2	29.3	29.3
1958	29.3	29.4	29.5	29.5	29.5	29.6	29.6	29.6	29.7	29.7	29.8	29.9
1959	29.9	29.9	30	30	30.1	30.2	30.2	30.2	30.3	30.4	30.4	30.5
1960	30.5	30.6	30.6	30.6	30.6	30.7	30.6	30.6	30.6	30.8	30.8	30.7
1961	30.8	30.8	30.9	30.9	30.9	31	31	31.1	31.1	31.1	31.2	31.2
1962	31.2	31.2	31.3	31.3	31.4	31.4	31.4	31.5	31.5	31.5	31.5	31.6
1963	31.5	31.6	31.7	31.7	31.7	31.8	31.8	31.9	31.9	32	32	32.1
1964	32.2	32.2	32.2	32.2	32.2	32.3	32.3	32.3	32.3	32.4	32.5	32.5
1965	32.6	32.6	32.6	32.7	32.7	32.7	32.7	32.7	32.8	32.8	32.9	33
1966	33	33.1	33.1	33.3	33.4	33.5	33.6	33.7	33.8	34	34	34.1
...												
2020	266.306	266.876	266.82	265.836	265.637	266.282	267.722	268.654	269.155	269.35	269.819	269.94
2021	270.025	270.299	271.214	273.7	275.718							

#### 3.2 平稳性检验

首先，作出美国 1947 年 1 月至 2020 年 12 月美国 Core CPI 指数时序图，如图1所示。

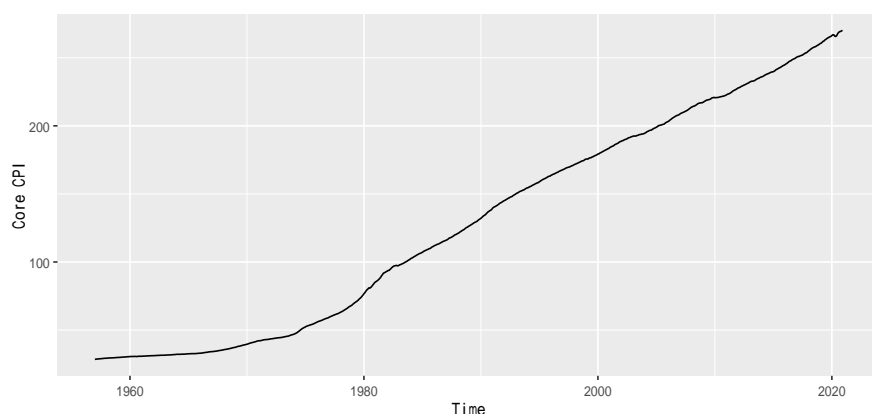


图 1: 美国 1947 年 1 月至 2021 年 5 月美国 Core CPI 指数时序图

由图1可以看出，美国自 1947 年以来 Core CPI 指数总体呈现出明显的上升趋势，因此该序列为非平稳序列。下面对该数据作不含趋势与漂移的 ADF 单位根检验，结果如表2所示。检验对应的  $p$  值为 0.07776，在取  $\alpha = 0.05$  的置信水平时，认为该序列非平稳。因此需要尝试对数据集进行差分处理，以实现该序列的平稳化。

表 2: Core CPI 序列 ADF 单位根检验

	Dickey-Fuller	Lag order	p-value
ADF Test	-3.259	9	0.07776

首先对 Core CPI 序列进行差分运算，将其转化为平稳的时间序列。绘制 1 阶差分以及 1 阶差分的曲线如图2所示。

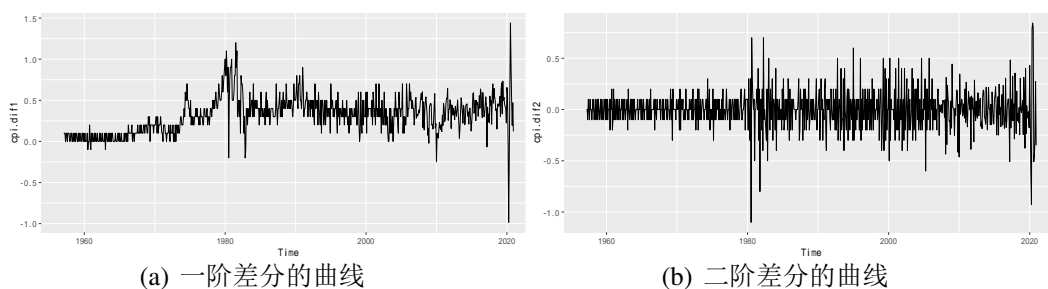


图 2: 1 阶差分及 2 阶差分后的曲线

对差分后的序列进行平稳化检验，如表3所示，可以看出 1 阶差分后序列仍然不平稳，而 2 阶差分后序列平稳。

表 3: 差分后 Core CPI 序列 ADF 单位根检验

	Dickey-Fuller	Lag order	p-value
1 阶差分	-3.3255	9	0.06629
2 阶差分	-13.109	9	< 0.01

### 3.3 模型识别与定阶

SARIMA 模型的识别与定阶是模型构建至关重要的一步。由平稳性检验可知，Core CPI 序列 1 阶差分后不平稳而 2 阶差分后平稳，因此， $ARIMA(p, d, q)$  ( $P, D, Q$ )<sup>s</sup> 中阶数  $d$  为 2。Core CPI 序列 2 阶差分后的 Core DCPI 序列的自相关及偏自相关图如图3所示，接下来通过该图来找到合适的 SARIMA 模型。

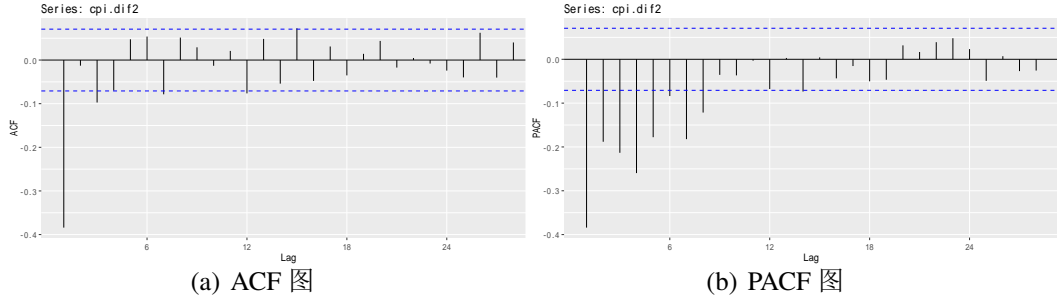


图 3: Core DCPI 序列的自相关图及偏自相关图

从图中可以看出，PACF 图拖尾，而 ACF 图中在延迟为 7 时的突起显示出数据包含非季节性 MA(7) 的成分，而延迟为 12 的突起则显示出季节性 MA(1) 的成分。因此，我们先从  $ARIMA(0, 2, 7)(0, 0, 1)^{12}$  模型开始，这表明数据经过了 2 阶差分，包含非季节性的 MA(7) 和季节性 MA(1) 成分。我们使用 R 语言 forecast 包中的自动定阶，获得的模型为  $ARIMA(0, 2, 5)(0, 0, 2)^{12}$ 。

表 4: Core CPI 序列 SARIMA 模型参数评价

	AIC	AICc	BIC
$ARIMA(0, 2, 7)(0, 0, 1)^{12}$	-630.38	-630.14	-588.61
$ARIMA(0, 2, 5)(0, 0, 2)^{12}$	-636.26	-636.07	-599.13

由表4中 SARIMA 模型的信息准则比较可知，两个模型差别不大，但使用 R 语言 forecast 包中的自动定阶获得的模型  $ARIMA(0, 2, 5)(0, 0, 2)^{12}$  相对更好，因此我们采用此模型，此模型的参数估计如表5所示。

表 5: Core CPI 序列  $ARIMA(0, 2, 5)(0, 0, 2)^{12}$  的参数估计

	ma1	ma2	ma3	ma4	ma5	sma1	sma2	$\sigma^2$
coefficients	-0.6983	-0.0881	-0.1557	-0.026	0.1421	-0.1179	-0.0968	0.02516
s.e.	0.0363	0.0440	0.0430	0.0409	0.0367	0.0405	0.0394	

### 3.4 模型诊断

#### 3.4.1 模型参数显著性检验

$ARIMA(0, 2, 7)(0, 0, 1)^{12}$  模型参数显著性检验结果如表6所示，除 MA(4) 对应系数以外，所有 p 值均小于 0.05。故将 MA(4) 剔除后再进行拟合，模型拟合结果如表7所示。

表 6: 参数显著性检验结果

参数	回归系数	t-统计量	p-value
MA(1)	-0.6983	-19.2382	8.13E-68
MA(2)	-0.0881	-2.00315	0.022758
MA(3)	-0.1557	-3.62148	0.000156
MA(4)	-0.0261	-0.63784	0.261884
MA(5)	0.1421	3.87155	5.87E-05
SMA(1)	-0.1179	-2.91036	0.001858
SMA(2)	-0.0968	-2.45762	0.007104

表 7: Core CPI 序列 ARIMA(0,2,5)(0,0,2)<sup>12</sup> 剔除 MA(4) 后的参数估计

	ma1	ma2	ma3	ma5	sma1	sma2	$\sigma^2$
Coefficients	-0.7014	-0.0894	-0.1641	0.1293	-0.1181	-0.0980	0.02514
s.e.	0.0359	0.0454	0.0413	0.0307	0.0404	0.0394	

对剔除 MA(4) 后的 ARIMA(0,2,7)(0,0,1)<sup>12</sup> 模型进行参数显著性检验, 结果如表8所示, 所有 p 值均小于 0.05, 由此认为所有参数均通过显著性检验。

表 8: 参数显著性检验结果

参数	回归系数	t-统计量	p-value
MA(1)	-0.7014	-19.5363	1.68E-69
MA(2)	-0.0894	-1.96853	0.024685
MA(3)	-0.1641	-3.97449	3.86E-05
MA(5)	0.1293	4.212958	1.41E-05
SMA(1)	-0.1181	-2.92338	0.001783
SMA(2)	-0.0980	-2.48655	0.006556

### 3.4.2 残差白噪声检验

剔除 MA(4) 后的 ARIMA(0,2,7)(0,0,1)<sup>12</sup> 模型残差图如图4所示。图中在延迟 25 阶时有一个显著峰值, 其余所有残差都在显著性临界值内, 因此认为该残差类似于白噪声。

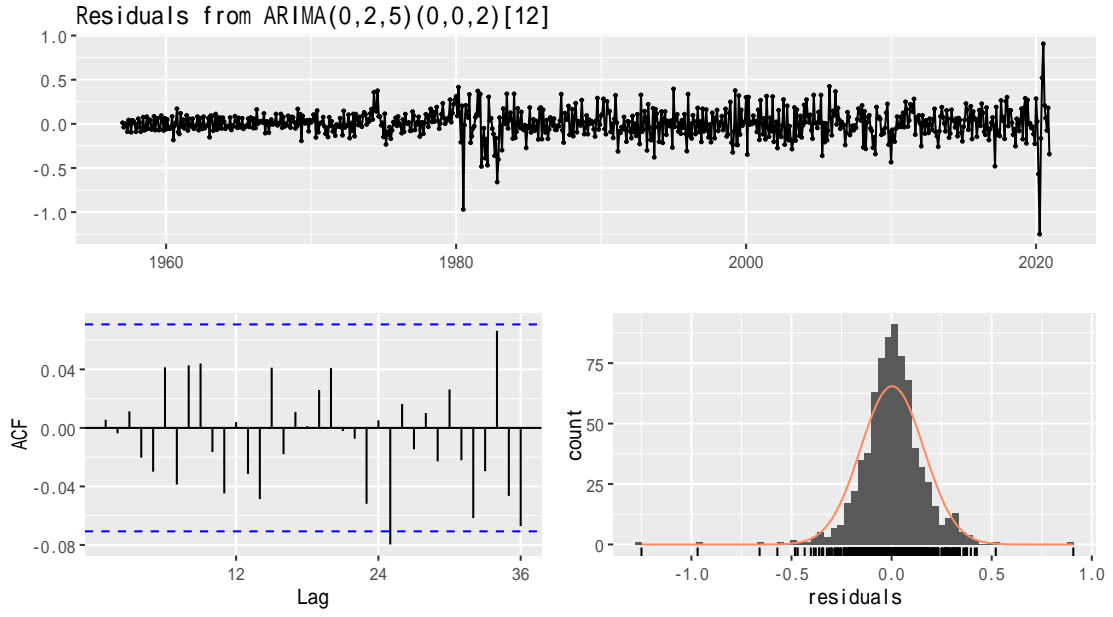


图 4: 模型  $ARIMA(0,2,5)(0,0,2)^{12}$  模型诊断

对残差序列进行延迟不同阶数的 Ljung-Box 检验，检验结果如表9所示，表中  $p$  值均大于 0.05，在取  $\alpha = 0.05$  的置信水平时可以认为该残差序列为白噪声序列。

表 9: Ljung-Box 检验结果

Lag	Q 统计量	df	p-value
6	2.4758	6	0.8712
12	8.3536	12	0.7569
18	12.664	18	0.8111
24	16.718	24	0.8604

### 3.5 模型预测与分析

由此构建的  $ARIMA(0,2,5)(0,0,2)^{12}$  模型为：

$$\nabla^2 CPI_t = \theta(B)\Theta(B)\varepsilon_t$$

其中  $\varepsilon_t \sim WN(0, 0.02516)$ ,  $\theta(B) = 1 + 0.7014B + 0.0894B^2 + 0.1641B^3 - 0.1293B^5$ ,  $\Theta(B) = 1 + 0.1181B^{12} + 0.0980B^{24}$ 。

### 3.5.1 模型预测效果分析

以下对样本进行预测，如图5所示，预测区间为 1978 年到 2017 年，对比真实值与预测值发现误差较小，预测效果非常好。

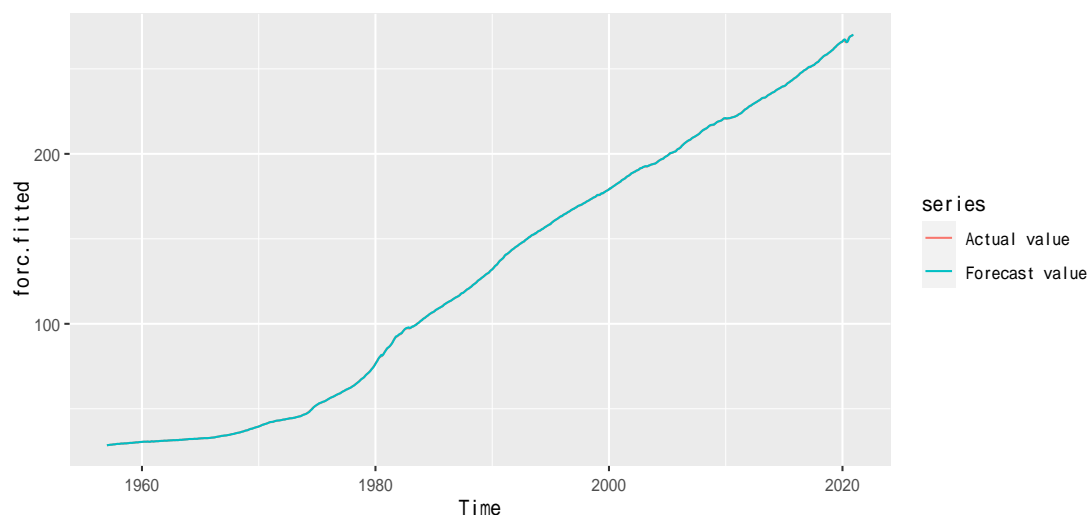


图 5: 对比 Core CPI 预测值与实际值

再对 2021 1 月至 2022 年 5 月的 Core CPI 进行预测，结果如表10所示, 由该表可知预测值与真实值之间的相对误差均小于 2%，预测值与真实值之间的均方误差较小，表明模型的预测精确。

表 10: 2021 年 1-5 月 Core CPI 预测结果

Time	Forecast	Actual	Fractional error	MSE
Jan-21	270.2585	270.025	0.09%	3.407061
Feb-21	270.6298	270.299	0.12%	
Mar-21	271.0891	271.214	0.05%	
Apr-21	271.6658	273.700	0.74%	
May-21	272.1518	275.718	1.29%	

### 3.5.2 模型预测

在模型预测验证精确有效的基础上，进一步对我国 2021 年 7 月至 2022 年 12 月的 Core CPI 进行预测，结果见图6，可见，美国 Core CPI 的基本走势趋于上升。



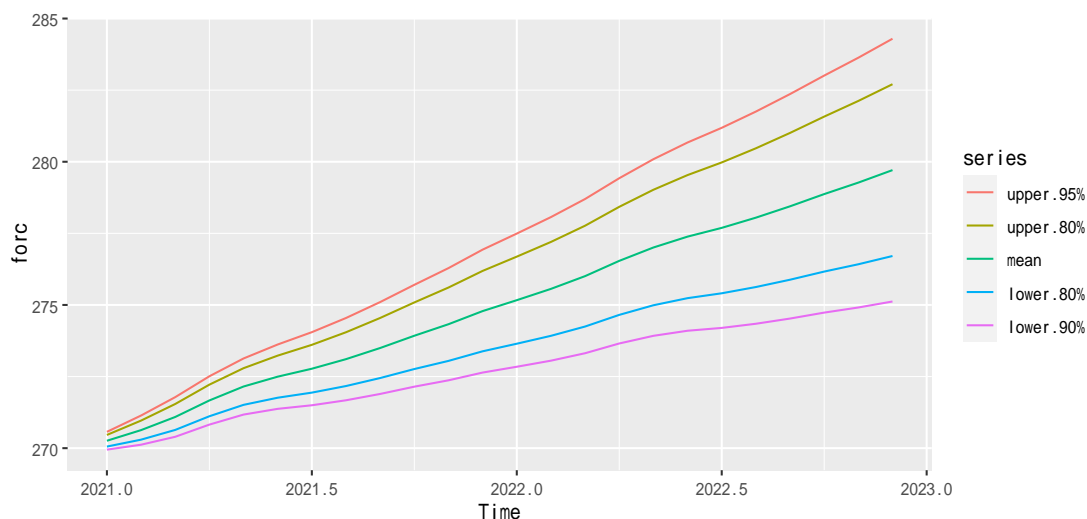


图 6: 2021 年至 2022 年的 Core CPI 预测图

## 4 研究结论与展望

### 4.1 研究结论

本文利用对 1957 年至 2020 年的美国 Core CPI 数据建立了具有较高预测精度的 SARIMA 模型，并对 2021 年 1 月至 6 月的 Core CPI 数据进行了预测与分析，模型预测的相对误差均小于 2%。经过检验，该预测模型产生的误差在合理范围内，则用该模型预测 2021 年的 Core CPI 指数具有一定的准确性和参考价值。但是 SARIMA 模型的短期预测是比较理想的，要从长期预测中获得良好的结果还需要进一步研究。

针对模型预测的结果，美国政府应该给予重视并积极采取如下措施：

1. 发行债券，加大财政补贴力度；
2. 适当地进行超前的基础设施建设，例如交通运输、教育、医疗等行业；
3. 做好合理的引导，加大国家的调控力度，减少市场的流动性，使物价上涨压力减缓；
4. 在防止通货膨胀的同时也要密切关注是否产生通货紧缩等，以此刺激经济的发展，与企业、民众共克时艰。

## 4.2 学习展望

通过这次课程大作业，使我明白了自己原有知识还有所欠缺，需要学习的内容还有很多。学习是一个长期积累的过程，在以后的学习、生活中都应该不断提高自己，扩充自己的知识面，培养自己的综合素质。在此次大作业我也学到了很多知识，更培养了我查找数据和文献，以及独立思考的能力。这次大作业也让我体验到了数据分析过程中的艰难以及成功的喜悦。虽然该上文中建立的模型还不是很完善，但是在数据分析过程中所学到的东西是我这次大作业最大的收获和财富，是我终身受益。

经过分析验证，模型预测的相对误差均小于 2%，说明了模型的准确性。模型还可以进行进一步的研究和改善，这就需要不断的学习专业知识和提高实践能力，为以后的学习生活打下坚实的基础。本文还可以做的进一步改善是：

1. 虽然线性指数平滑模型其实都是 ARIMA 模型的特例，但是非线性的指数平滑模型在 ARIMA 模型中并没有对应的部分。因此该数据也可以使用 ETS 模型进行预测及分析，观察是否有更好的预测效果；
2. 该模型中的点预测是预测分布的均值，对于该模型可以进一步采用回归分析中的分位数回归以及众数回归，观察是否能得到更好的预测效果；
3. 本次研究使用的一元时间序列分析，未来可以对多维时间序列进行分析，以提供更加准确的 Core CPI 预测模型。

## 参考文献

- [1] 维基百科. 居民消费价格指数 [DB/OL]. <https://zh.m.wikipedia.org/wiki/%E6%B6%88%E8%B2%BB%E8%80%85%E7%89%A9%E5%83%B9%E6%8C%87%E6%95%B8>, Accessed on 6 June 2021.
- [2] Thomas Herold, Wesley Crowder. Financial Terms Dictionary - Principles of Economics Explained [M]. Independently published, 2017.
- [3] 周永道, 王会琦, 吕王勇. 时间序列分析及应用 [M]. 北京: 高等教育出版社, 2015.
- [4] 薛毅, 陈立萍. 时间序列分析与 R 软件 [M]. 北京: 清华大学出版社, 2020.
- [5] Hyndman, R.J., & Athanasopoulos, G. (2018) Forecasting: principles and practice, 2nd edition[M/OL], OTexts: Melbourne, Australia. OTexts.com/fpp2. Accessed on 6 June 202.
- [6] U.S. Bureau of Labor Statistics, Consumer Price Index for All Urban Consumers: All Items Less Food and Energy in U.S. City Average [DB/OL], retrieved from FRED, Federal Reserve Bank of St. Louis; <https://fred.stlouisfed.org/series/CPILFESL>, June 15, 2021.

## 代码及运行结果

```
#清空环境
rm(list=ls())

#导入时间序列数据
cpi <- read.csv("D:/CPILFESL.csv", header = T)
cpi <- ts(cpi$CPILFESL, start = c(1957,1), end = c(2020,12),
         frequency = 12)
head(cpi)
## [1] 28.5 28.6 28.7 28.8 28.8 28.9

#时序图
library(forecast)
## Registered S3 method overwritten by 'quantmod':
##   method      from
##   as.zoo.data.frame zoo
pdf("fg1.pdf", family = "GB1", width = 8, height = 4)
autoplot(cpi, xlab = "Time", ylab = "Core_CPI")
dev.off()
## png
## 2

#平稳性检验
library(tseries)
adf.test(cpi)
##
## Augmented Dickey-Fuller Test
##
## data: cpi
## Dickey-Fuller = -3.259, Lag order = 9, p-value = 0.07776
## alternative hypothesis: stationary

#序列差分
cpi.dif1 <- diff(cpi)
pdf("fg2.pdf", family = "GB1", width = 8, height = 4)
```

```

autoplot(cpi.dif1)
dev.off()
## png
## 2
cpi.dif2 <- diff(cpi.dif1)
pdf("fg3.pdf", family = "GB1", width = 8, height = 4)
autoplot(cpi.dif2)
dev.off()
## png
## 2

#差分后序列平稳性检验
adf.test(cpi.dif1)
##
## Augmented Dickey-Fuller Test
##
## data: cpi.dif1
## Dickey-Fuller = -3.3255, Lag order = 9, p-value = 0.06629
## alternative hypothesis: stationary
adf.test(cpi.dif2)
##
## Augmented Dickey-Fuller Test
##
## data: cpi.dif2
## Dickey-Fuller = -13.109, Lag order = 9, p-value = 0.01
## alternative hypothesis: stationary

#模式识别与定阶
pdf("fg4.pdf", family = "GB1", width = 8, height = 4)
ggAcf(cpi.dif2)
dev.off()
## png
## 2
pdf("fg5.pdf", family = "GB1", width = 8, height = 4)
ggPacf(cpi.dif2)

```

```

dev.off()
## png
## 2

#模型拟合
cpi.fit <- Arima(cpi,order=c(0,2,7),seasonal=c(0,0,1))
cpi.fit
## Series: cpi
## ARIMA(0,2,7)(0,0,1) [12]
##
## Coefficients:
##          ma1          ma2          ma3          ma4          ma5          ma6          ma7
##          sma1
##      -0.6765 -0.1143 -0.1572 -0.0227  0.1313  0.0578 -0.0436
##      -0.1099
## s.e.   0.0365   0.0444   0.0446   0.0416  0.0462  0.0429   0.0391
##      0.0466
##
## sigma^2 estimated as 0.02533: log likelihood=324.19
## AIC=-630.38 AICc=-630.14 BIC=-588.61
cpi.fit.auto <- auto.arima(cpi)
cpi.fit.auto
## Series: cpi
## ARIMA(0,2,5)(0,0,2) [12]
##
## Coefficients:
##          ma1          ma2          ma3          ma4          ma5          sma1
##          sma2
##      -0.6983 -0.0881 -0.1557 -0.0261  0.1421 -0.1179
##      -0.0968
## s.e.   0.0363   0.0440   0.0430   0.0409  0.0367   0.0405
##      0.0394
##
## sigma^2 estimated as 0.02516: log likelihood=326.13
## AIC=-636.26 AICc=-636.07 BIC=-599.13

```

```

#模型参数显著性检验
len <- length(cpi) - 7
se <- c(0.0363,0.0440,0.0430,0.0151,0.0367,0.0405,0.0394)
for(i in 1:7) print(cpi.fit.auto$coef[i]/se[i])
##          ma1
## -19.23815
##          ma2
## -2.003145
##          ma3
## -3.621476
##          ma4
## -1.727668
##          ma5
##  3.87155
##          sma1
## -2.910361
##          sma2
## -2.457616
for(i in 1:7){
  if(cpi.fit.auto$coef[i]>0)
    print(pt(cpi.fit.auto$coef[i]/se[i], df = len, lower.tail = FALSE))
  else
    print(pt(cpi.fit.auto$coef[i]/se[i], df = len, lower.tail = TRUE))
}
##          ma1
## 8.134269e-68
##          ma2
## 0.02275796
##          ma3
## 0.0001561761
##          ma4
## 0.0422267
##          ma5
## 5.871672e-05

```

```

##          sma1
## 0.001857742
##          sma2
## 0.007104123

#剔除变量MA(4)后拟合新模型
cpi.fit.new <- Arima(cpi,order=c(0,2,5),seasonal=c(0,0,2),
  transform.pars=F,fixed=c(NA,NA,NA,0,NA,NA,NA))
cpi.fit.new
## Series: cpi
## ARIMA(0,2,5)(0,0,2)[12]
##
## Coefficients:
##          ma1          ma2          ma3 ma4          ma5          sma1          sma2
##      -0.7014 -0.0894 -0.1641   0  0.1293 -0.1181 -0.0980
## s.e.   0.0359   0.0454   0.0413   0  0.0307   0.0404   0.0394
##
## sigma^2 estimated as 0.02514: log likelihood=325.93
## AIC=-637.85 AICc=-637.7 BIC=-605.36

#模型参数显著性检验
len.new <- length(cpi) - 6
se.new <- c(0.0359,0.0454,0.0413,0,0.0307,0.0404,0.0394)
for(i in 1:7) print(cpi.fit.new$coef[i]/se.new[i])
##          ma1
## -19.53633
##          ma2
## -1.968525
##          ma3
## -3.974491
## ma4
## NaN
##          ma5
## 4.212958
##          sma1
## -2.923383

```



```

##      sma2
## -2.48655
for(i in 1:7){
  if(cpi.fit.new$coef[i]>0)
    print(pt(cpi.fit.new$coef[i]/se.new[i], df = len, lower
      .tail = FALSE))
  else
    print(pt(cpi.fit.new$coef[i]/se.new[i], df = len, lower
      .tail = TRUE))
}
##          ma1
## 1.680034e-69
##          ma2
## 0.02468511
##          ma3
## 3.862139e-05
## ma4
## NaN
##          ma5
## 1.411304e-05
##          sma1
## 0.001782614
##          sma2
## 0.006556102

#残差白噪声检验
pdf("fg6.pdf", family = "GB1", width = 8, height = 4.5)
checkresiduals(cpi.fit.new)
##
##  Ljung-Box test
##
## data: Residuals from ARIMA(0,2,5)(0,0,2)[12]
## Q* = 16.718, df = 17, p-value = 0.4736
##
## Model df: 7. Total lags used: 24

```

```

dev.off()
## png
## 2
for(i in 1:4) print(Box.test(cpi.fit.new$residuals, type = "
    Ljung-Box", lag = 6 * i))
##
## Box-Ljung test
##
## data: cpi.fit.new$residuals
## X-squared = 2.4758, df = 6, p-value = 0.8712
##
##
## Box-Ljung test
##
## data: cpi.fit.new$residuals
## X-squared = 8.3536, df = 12, p-value = 0.7569
##
##
## Box-Ljung test
##
## data: cpi.fit.new$residuals
## X-squared = 12.664, df = 18, p-value = 0.8111
##
##
## Box-Ljung test
##
## data: cpi.fit.new$residuals
## X-squared = 16.718, df = 24, p-value = 0.8604

#模型预测效果
cpi.forc <- forecast(cpi.fit.new)
forc.fitted <- cbind(cpi, cpi.forc$fitted)
colnames(forc.fitted) <- c("Actual_value", "Forecast_value")
pdf("fg7.pdf", family = "GB1", width = 8, height = 4)
autoplot(forc.fitted)
dev.off()

```

```

## png
## 2
cpi.forc$mean[1:5]
## [1] 270.2585 270.6298 271.0891 271.6658 272.1518

#模型预测
forc <- cbind(cpi.forc$lower,cpi.forc$mean,cpi.forc$upper)
colnames(forc) <- c("lower.80%", "lower.90%", "mean", "upper.80%",
  , "upper.95%")
new_cols <- c("upper.95%", "upper.80%", "mean", "lower.80%", "
  lower.90%")
forc <- forc[, new_cols]
pdf("fg8.pdf", family = "GB1", width = 8, height = 4)
autoplot(forc)
dev.off()
## png
## 2

```