

基于 ARIMA 和 Holt 指数平滑模型的 山东省药品集中采购价格指数研究

蔡文鼎, 魏新江

(鲁东大学 数学与统计科学学院, 山东 烟台 264039)

摘要: 本文实现了对“十三五”期间山东省药品集中采购价格的动态分析和预测。利用采购平台提供的数据编制链式拉氏价格指数、帕氏价格指数, 将价格指数作为对药品集中采购价格预测分析的重要指标依据, 以此进行量化评价。结合 ARIMA 和 Holt 指数平滑模型, 对已有的药品价格指数数据进行预测, 两个模型的平均绝对误差分别为 0.361 1 和 0.372 3, 表明模型能够有效拟合序列的整体趋势, 以及提取价格指数序列动态及持续特征。通过对价格指数进行趋势预测以评估政府及相关部门稳控药价等政策的效果, 为政府和平台管理者提供科学的决策依据。

关键词: 山东省药品集中采购; 价格指数; ARIMA; Holt 双参数指数平滑

中图分类号: O213 **文献标志码:** A **文章编号:** 1673-8020(2022)03-0218-07

“十三五”期间, 山东省药品集中采购平台努力营造安全有序的采购环境, 积极落实《山东省药品集中采购实施办法(试行)》等政策。政府健全药品供应保障措施, 完善药品挂网方式, 实行价格联动机制, 实现政府和市场双向维护基本药物集中采购流程, 确保质量合格和价格合理双过关。政府也通过量价挂钩的采购方式, 进一步稳定药品价格, 促使药品采购价格在平稳中实现下降, 以减轻人民的用药负担, 让人民放心用药、用得起药。因此, 如何准确获取药品采购价格波动等信息, 并对其进行动态监测, 使管理者能及时作出应对措施以稳控药价, 这无疑是一项重要工作。

价格指数拥有反应价格水平的特性, 对价格指数的编制和统计模型的预测分析, 对采购价格的监测将产生重大作用。目前, 国内许多学者对药品价格指数编制进行相关研究。李悦等^[1]通过比较各类指数, 最终选择了马氏价格指数作为价格管理的指标依据, 但没有考虑消费结构的变化, 且权重固定; 马芳芳等^[2]从药品定义、价格以及数量计量单位等方面, 对不同价格指数的优缺点进行对比分析, 为理解各类价格指数的作用提

供重要依据; 刘华等^[3]研究了链式 CPI 在美国财政政策中的应用, 对比传统 CPI 的优点, 链式 CPI 更能反应民众生活成本的变化。此外, 王淑玲等^[4]对药品价格综合指数进行研究; 张崖冰^[5]探讨了药品价格指数指标在提供价格管理信息方面的重要作用; 王淑庆^[6]立足药品交易市场, 对比分析各类价格指数的意义。文献[4—6]主要研究价格指数的编制意义, 但对药品价格指数的预测缺少深入探讨。

在有关药品价格分析和价格指数预测方法上, 国内外学者也进行很多相关研究。Shi 等^[7]通过建立广义回归模型, 研究了 2010—2017 年中国省级收入水平与药品价格之间的关联; 文献[8—12]利用 ARIMA 模型对居民消费价格指数进行研究预测, 指出该模型预测精度高、序列信息提取充分; 文献[13—15]结合具体案例, 利用时间序列和指数平滑模型进行预测研究, 分析得出相对于传统指数平滑模型, Holt 双参数指数平滑模型预测精度更高; 文献[16]用多元回归模型对不同类型中药材价格进行预测分析, 实现了对某一药品价格的预测研究。

收稿日期: 2021-08-01; 修回日期: 2022-02-21

基金项目: 国家自然科学基金(61973149); 山东省自然科学基金重点项目(ZR2020KF029); 2021 年度省社科规划研究项目(21CSDJ20)

第一作者简介: 蔡文鼎(1997—), 男, 山东滨州人, 硕士研究生, 研究方向为经济与社会统计。E-mail: cwd1789426288@163.com

通信作者简介: 魏新江(1977—), 男, 山东东营人, 教授, 硕士研究生导师, 博士, 研究方向为非线性系统控制、鲁棒控制等。E-mail: weixinjiang@163.com

以上研究针对药品价格指数的预测分析较少,缺乏对药品价格的整体趋势预测。本文从山东省药品集中采购平台提供的若干订单数据中提取可用信息,采用链式拉氏与帕氏价格指数,编制药品价格指数;比较药品价格总指数与分类指数的变化趋势,利用ARIMA和Holt指数平滑模型对每月挂网出售药品价格指数进行预测研究;最后,综合分析平台采购的药品价格变动情况,评估“十三五”期间政府及相关采购部门落实稳控药价政策情况。

1 药品采购价格指数统计分析

1.1 价格指数编制

价格指数可以反映特定商品或服务价格在某一时期的变动幅度,是从相对水平的角度去衡量一组价格的变化程度及趋势。由于价格存在量纲且蕴含商品价值信息的特点,难以利用价格指标直接衡量一组价格在某一时期的波动,因此,为了准确衡量研究组的价格水平变动情况,一般通过编制价格指数对价格形成动态监测。

价格指数分为以拉氏、帕氏为代表的固定权重的价格指数和以链式、费氏为代表的非固定权重的价格指数^[1-2]。利用不同方法编制的价格指数,由于不同权重代表的意义不同,所以要根据分析目的选取合适的模型。本文选取链式拉氏价格指数和帕氏价格指数模型,编制每月均在平台挂网出售的药品价格总指数及中药、西药等分类价格指数。

链式拉氏价格指数能够考虑所有商品的结构变化,即不仅考虑商品价格的变化,还考虑数量的变化;能够及时更新权重,采用先计算环比指数、再合成定基指数的方法,减少因结构变化带来的计算误差,比拉氏价格指数更能准确反映商品价格的变化趋势。计算公式如下:

$$L_t = \left[\sum_{i=1}^n W_{i-1} \frac{P_i}{P_{i-1}} \right] \times L_{t-1}, \quad (1)$$

其中: L_t, L_{t-1} 分别为现期和前一期的采购药品价格指数, $t = 1, 2, \dots, n$ 是基于基期的第 t 期; W_{i-1} 表示前一期的权重, P_i 和 P_{i-1} 分别表示现期和前一

期药品代表组的价格。

帕氏价格指数是以现期数量作为与前期对比的权重,可以同时反映价格和数量的变动,能够说明消费结构发生现期变动时引起的支出费用变化不能反映纯价格变动的情况。考虑到民众比较关心用药费用,本文选取帕氏价格指数分析维持现期需求的药品成本和用药费用变化。其计算公式如下:

$$L_t = \frac{\sum_{i=1}^n P_i Q_i}{\sum_{i=1}^n P_0 Q_i}, \quad (2)$$

其中: L_t 表示现期采购药品价格指数, Q_i 表示现期药品销售数量; P_0 表示基期采购药品价格指数,设置为100。

1.2 数据来源

本文所用数据为山东省药品集中采购平台提供的2016年1月—2020年7月共55个月的采购订单。利用这些订单数据进行价格指数编制,步骤简要概括如下:

1) 利用SQL软件对订单大数据进行预处理,即分析订单数据表结构,检查异常值和关联匹配字段间的关系及可行性;

2) 根据每月关联药品ID字段进行内连接匹配,选取55个月均在平台挂网出售的2805种药品,再细分其他字段,提取出每月挂网的中药、化学药、医保药和非医保药分别为681、2029、2452和346种;

3) 根据价格指数模型(1)、(2),以2016年1月作为基期,编制链式拉氏价格指数和帕氏价格指数;

4) 以基于模型(1)编制的月度药品链式拉氏价格指数为预测对象,计算2805种代表品的价格,对药品价格波动进行整体趋势分析。

1.3 价格指数描述分析

利用每月均在平台挂网出售药品的链式拉氏价格指数,以及中药和化学药、医保药和非医保药的分类链式拉氏价格指数,得到每月各类挂网药品价格指数变动趋势,如图1所示。

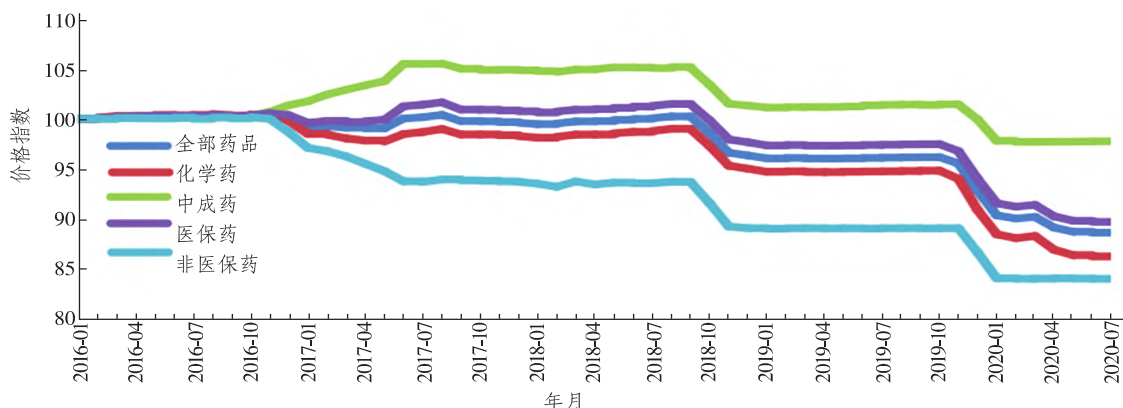


图 1 每月挂网药品的链式拉氏价格指数变动趋势

Fig.1 Monthly trend of chain Laspeyres price index of drugs for hanging network bidding

由图 1 可以看出: 政府采购药品的价格变动整体呈现小幅度波动且有下降趋势; 中药、化学药、医保药以及非医保药的分类价格指数变动趋势与药品价格总指数波动趋势大致相同。这说明山东省政府通过以量换价等政策的实施, 有效监测采购过程, 使药品价格下降; 同时, 自 2019 年底开始, 中成药、化学药等药品价格明显下降, 说明新冠疫情期间政府稳定药品价格的一系列政策取得明显成效。

同样, 根据相关数据得到每月挂网药品的帕氏价格指数变动趋势, 见图 2。由图 1、2 可以得到, 药品价格总体降幅分别为 11.87% 和 20.79%, 说明两指数都有较大幅度的降低。由于两类价格指数能很好地反应价格变动和民众现期支出费用的情况, 能够说明民众在现阶段用药需求下, 药品价格和用药支出费用都明显下降, 这进一步说明山东省药品采购平台对药品采购政策的有效落实。

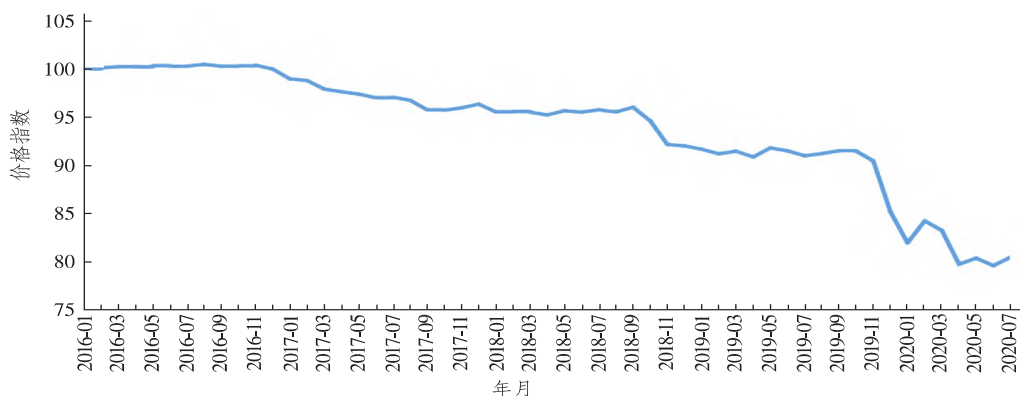


图 2 每月挂网药品的帕氏价格指数变动趋势

Fig.2 Monthly trend of Paasche price index of drugs for hanging network bidding

2 价格指数预测模型

2.1 ARIMA 模型

ARIMA 模型是目前最常用的时间序列建模方法, 可以将非平稳序列经过差分运算得到平稳序列, 再对平稳序列进行建模。该模型可以有效运用序列的趋势信息、动态特征和持续特征, 根据

历史序列和当前序列的相关信息预测未来值, 具有很好的短期预测效果^[13]。

运用 ARIMA 建模时应遵循模型的平稳性检验、初步识别、参数检验和残差序列白噪声检验等步骤。其结构表达式为:

$$\nabla^d x_t = u + \frac{\Theta(B)}{\Phi(B)} \varepsilon_t, \quad (3)$$

其中: $\Phi(B) = 1 - \varphi_1 B - \dots - \varphi_p B^p$, $\Theta(B) = 1 - \theta_1 B - \dots - \theta_q B^q$ 分别为 ARMA(p, q) 的自回归系

数和移动平均系数多项式, B 为延迟算子, $\varphi_i (i = 1, 2, \dots, p)$ 和 $\theta_j (j = 1, 2, \dots, q)$ 为系数权重; u 为差分后序列的均值, ε_t 为零均值白噪声序列; $\nabla^d = (1 - B)^d$ 表示用延迟算子 B 代表序列 d 阶差分, x_t 表示价格指数原始序列。

2.2 Holt 双参数指数平滑模型

Holt 双参数指数平滑模型对有线性趋势的序列有很好的预测作用,能较好地提取长期的变化趋势。该模型有两个待估参数^[15],会对序列的水平值、趋势值进行平滑,用两次平滑结果进行线性外推预测^[16],对时间序列建模有更高的适应性和灵活性。

模型结构表示如下:

$$\begin{cases} \hat{a}(t) = \alpha x_t + (1 - \alpha) (\hat{a}(t-1) + \hat{b}(t-1)), \\ \hat{b}(t) = \beta (\hat{a}(t) - \hat{a}(t-1)) + (1 - \beta) \hat{b}(t-1), \end{cases} \quad (4)$$

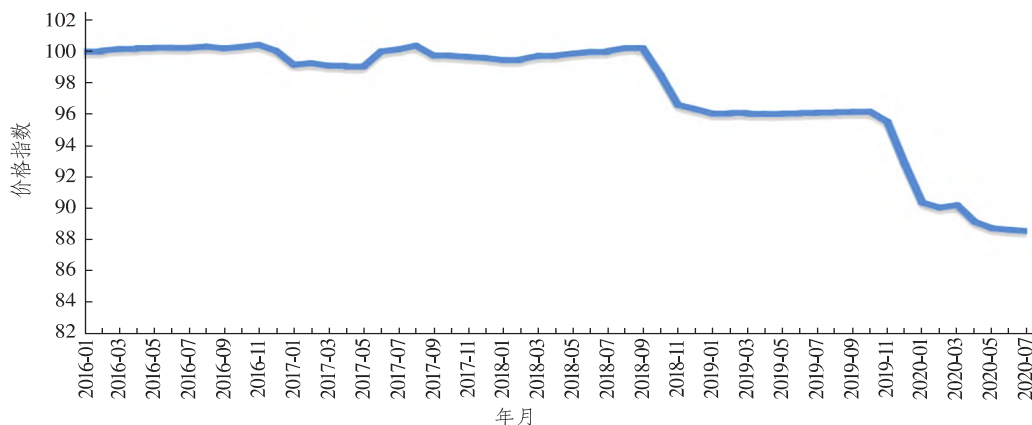


图3 每月挂网药品价格指数趋势

Fig.3 Monthly trend of online drug price index

为判断差分后序列的长期趋势信息是否提取充分,得到图4。由图4可以看出,序列差分后对长期趋势信息的提取比较充分。由 ADF 检验得出,在 0.05 的显著性水平下,可以认为每月挂网药品链式拉氏价格指数序列经过差分后是平稳序列。

3.1.2 模型建立及检验

由于自相关系数代表 x_t 与所有随机变量 ($x_{t-1}, x_{t-2}, \dots, x_{t-k}$) 的相关关系,而偏自相关系数剔除其他随机变量的影响,单纯测度其中一个随机变量 x_{t-k} 对 x_t 的影响。为确立 ARIMA(p, d, q) 模

其中: $\hat{a}(t)$ 和 $\hat{b}(t)$ 为根据序列最新观察值不断修匀的截距项与斜率项; x_t 为序列在 t 时刻最新观察值; α 和 β 为平滑系数,取值范围为 $[0, 1]$ 。进一步由模型(4)得到 m 个时期的预测值 \hat{x}_{t+m} , 即:

$$\hat{x}_{t+m} = \hat{a}(t) + \hat{b}(t) m, \forall m \geq 1. \quad (5)$$

3 实证分析

3.1 ARIMA 模型分析

3.1.1 平稳性检验

为判断原序列是否平稳,得到图3。由图3可以看出,链式拉氏价格指数序列波动趋势较大。对该序列进行 ADF 检验,ADF 检验 t 统计量值为 -1.1138 , $P > 0.05$,说明该序列为非平稳序列。

型阶数,得到图5。由图5可以看出,自相关系数和偏自相关系数均没有明显的截尾特征。所以,可以基于差分序列建立 ARIMA 模型,利用 R 软件基于 BIC 最优原则自动搜索相对最优模型,再结合图5(a)、(b),建立不带漂移项的 ARIMA(1,1,2) 模型。

计算模型参数及标准差,结果见表1。由表1可以得到,参数估计值均大于两倍标准差,说明参数是显著非零的,模型参数有意义。

为进一步说明模型参数通过检验,得到表2。由表2可以看到, P 值远小于 0.05,说明参数显著非零,可以保留。

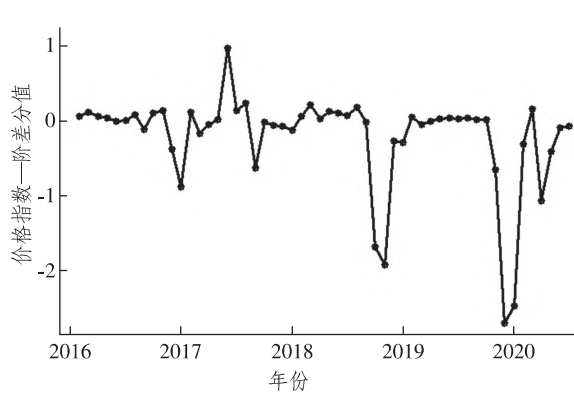
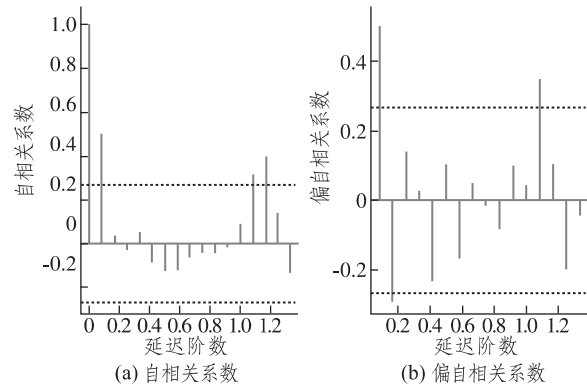


图 4 药品价格指数序列的一阶差分

Fig.4 First-order difference of drug price index series



注: 平行虚线是自相关和偏自相关系数两倍标准差的参考线; 悬垂线的长度表示系数的大小。

图 5 一阶差分序列自相关与偏自相关系数

Fig.5 Autocorrelation and partial autocorrelation coefficients of first-order difference series

表 1 模型参数及标准差

Tab.1 Model parameters and standard deviations

参数	AR(1)	MA(1)	MA(2)
系数	-0.678 6	1.546 6	0.821 1
标准差	0.146 9	0.156 3	0.136 9

表 2 模型参数检验

Tab.2 Test of model parameters

参数	AR(1)	MA(1)	MA(2)
参数检验 t 值	4.620 6	9.893 3	5.999 5
参数 t 检验 P 值	1.317 1E -05	9.285 9E -14	1.027 4E -07

为评估模型的有效性, 检验拟合的模型是否已经充分提取原序列信息, 得到图 6。由图 6(c)

可以看出, P 值均大于 0.05, 说明残差序列为白噪声序列, 证明信息提取充分。

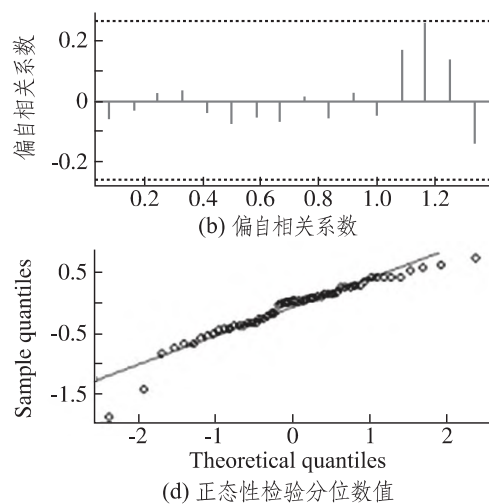
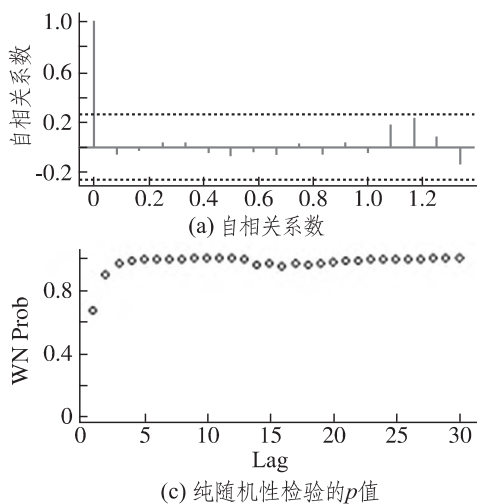


图 6 残差序列的白噪声检验

Fig.6 White noise test for residual sequences

3.1.3 模型预测

建立的 ARIMA(1,1,2) 模型表达式为:

$$(1 - B)x_t = \frac{1 + 1.5466B + 0.8211B^2}{1 + 0.6786B} \varepsilon_t \quad (6)$$

对未来 6 期进行预测, 预测结果见图 7 和表 3。综合预测结果得到, 未来 6 期的药品价格指数波动趋势平稳。

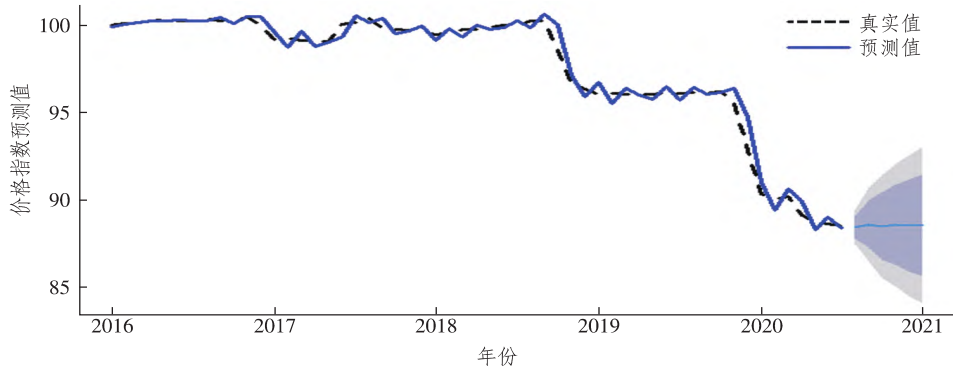


图7 ARIMA 模型拟合及预测

Fig.7 ARIMA model fitting and prediction

表3 ARIMA 模型的未来6期预测值

Tab.3 The predictive values of ARIMA model in the future 6 periods

预测期	2020-08	2020-09	2020-10	2020-11	2020-12	2021-01
预测值	88.481 3	88.637 1	88.531 4	88.603 1	88.554 4	88.587 5

3.2 Holt 双参数指数平滑模型分析

针对每月平台挂网出售药品链式拉氏价格指数数据,分析其线性变动趋势,并利用 Holt 双参数指数平滑模型进行预测。

通过 R 软件,基于最优拟合原则得到模型参数(表4)。由表4可以看出,α和β值均很高,表明不论从水平方向还是趋势方向考虑序列预测,时序数据近期的观测值在当前值的预测分析中都有较高的权重。这主要因为原序列在整个时间段上变化较大,所以当前值的预测更应侧重近期观测值。

表4 模型参数值

Tab.4 Model parameter values

参数	α	β	a(t)	b(t)
估计值	1	0.750 7	88.559 4	-0.110 1

根据表4的参数估计值,得到 Holt 双参数指数平滑模型预测表达式:

$$\hat{x}_{t+m} = 88.559 4 - 0.110 1m, \forall m \geq 1. \quad (7)$$

基于模型(7)预测未来6期的价格指数,结果见图8和表5。综合预测结果表明,未来6期的药品价格指数有缓慢的下降趋势。

3.3 结果分析

基于表3、5的预测结果,通过计算平均绝对误差(MAE)对模型精度进行评价。MAE计算公式为:

$$y_{MAE} = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i|, \quad (8)$$

其中: y_{MAE} 表示对初始价格指数序列 y_i 预测后的平均绝对误差; \hat{y}_i 表示价格指数序列 y_i 的预测值, $i(i = 1, 2, \dots, n)$ 表示基期后的第 i 期。

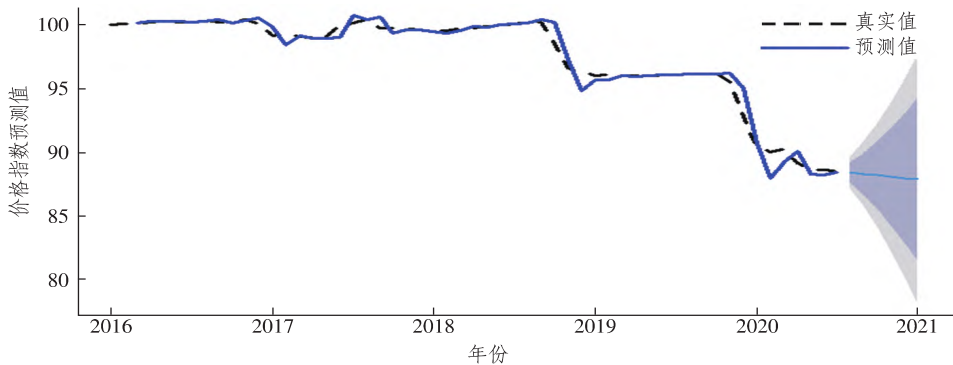


图8 Holt 双参数指数平滑模型拟合及预测

Fig.8 Holt two-parameter exponential smoothing model fitting and prediction

表5 Holt 双参数指数平滑模型的未来6期预测值

Tab.5 The predictive values of Holt two-parameter exponential smoothing model in the future 6 periods

预测期	2020-08	2020-09	2020-10	2020-11	2020-12	2021-01
预测值	88.449 3	88.339 2	88.229 0	88.118 9	88.008 8	87.898 6

由式(8)计算得到,ARIMA和指数平滑模型的MAE值分别为0.361 1和0.372 3。对比文献[8—9],预测数据类型及所用模型一致,但本文ARIMA和Holt双参数指数平滑模型的MAE值更高,说明本文预测效果更好。

此外,图7显示价格指数成波动状态,且向上波动逐渐放缓,向下波动高于向上波动的趋势;由于充分考虑近期观测值的趋势和水平值,图8中的价格指数近期处于缓慢的下降趋势。结合图7、8可以判断,药品集中采购价格将呈现稳中有降的趋势,表明“十三五”期间山东省药品集中采购平台对采购政策的改革是非常有效的。此外,药品价格下降说明采用量价挂钩、以量换价的采购方式以及实行价格联动机制等措施,有利于稳定药物价格,切实减轻人民用药负担。

4 结语

本文利用“十三五”期间山东省药品采购平台的数据,通过编制药价价格指数,基于ARIMA和Holt双参数指数平滑模型对采购药品价格进行动态分析和趋势预测。利用两个预测模型对价格指数序列进行有效拟合,并结合预测模型的特点,综合分析未来药品价格的变动趋势。结果显示,未来药品价格将呈现稳中有降的趋势,说明省采购平台实行药品价格联动、带量采购政策的有效性。

由于样本数据较少,个别值的预测可能会存在偏差,后续将根据平台提供的新月份订单数据划分更为细致的药理分类等新特征数据,重新拟合和优化模型,以期对山东省药品集中采购价格进行更全面的数据分析。

参考文献:

- [1] 李悦,朱凯,俞慧强.常用药品价格指数计算公式的比较[J].中国药房,2012,23(8):690-692.
- [2] 马芳芳,吴晶.药品价格指数的方法学综述[J].中国

卫生政策研究,2015,8(7):61-67.

- [3] 刘华,李亚培.链式CPI在美国财政政策和货币政策中的应用及启示[J].南方金融,2012,425(1):47-49.
- [4] 王淑玲,李飞立,王妍,等.关于我国药品价格综合指数的研究[J].价格理论与实践,2014,359(5):75-77.
- [5] 张崖冰.药品价格指数在药品价格管理中的应用[J].价格理论与实践,2009,295(1):53.
- [6] 王淑庆.药品成交价格指数编制方法研究[D].重庆:重庆大学,2013.
- [7] SHI W, WOUTERS O J, LIU G, et al. Association between provincial income levels and drug prices in China over the period 2010—2017 [J]. Social Science & Medicine, 2020, 263: 113247.
- [8] 杨颖梅.基于ARIMA模型的北京居民消费价格指数预测[J].统计与决策,2015,424(4):76-78.
- [9] 李树,汪飞,王丰效.基于ARIMA模型的青岛市居民消费价格指数实证分析[J].兰州文理学院学报(自然科学版),2020,34(6):11-16.
- [10] 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,2007,158(2):59-67.
- [11] 潘静,张颖,刘璐.基于ARIMA模型与GM(1,1)模型的居民消费价格指数预测对比分析[J].统计与决策,2017,488(20):110-112.
- [12] 夏丽.基于ARIMA模型及回归分析的区域用电量预测方法研究[D].南京:南京理工大学,2013.
- [13] 钱力.农村居民收入区域差异发展趋势预测:基于二次指数平滑法和ARMA模型分析[J].中央财经大学学报,2014,323(7):78-82.
- [14] 李云娥,周云波.中国城乡收入差距未来发展趋势的预测[J].山西财经大学学报,2007,180(10):14-18.
- [15] 黎锁平,刘坤会.平滑系数自适应的二次指数平滑模型及其应用[J].系统工程理论与实践,2004(2):95-99.
- [16] 崔旭盛,丁颖辉,高秀强,等.基于多元线性回归模型的中药材市场价格分类预测实证研究[J].北方园艺,2020,461(14):157-161.

(下转第253页)

tant indicator of atmospheric environmental quality. In order to master the chemical composition characteristics of atmospheric precipitation in Huancui District of Weihai City and to provide theoretical reference for the control of atmospheric pollution in this region, based on the monitoring data of atmospheric precipitation chemical composition in Huancui District of Weihai City from 2015 to 2020, the chemical composition characteristics, variation trend and sources of atmospheric precipitation were analyzed by means of enrichment factor method, Spearman rank correlation coefficient and Pearson correlation coefficient method. The results show that no acid rain happened and the annual average pH of precipitation is 6.13 from 2015 to 2020 in Huancui District of Weihai city. The annual average conductivity was $4.35 \text{ mS} \cdot \text{m}^{-1}$, showing a general downward trend. Cl^- , SO_4^{2-} , Ca^{2+} , Na^+ , NH_4^+ and NO_3^- ions are the main ions in precipitation. The equivalent concentration of total ions in precipitation ranges from $464.5 \mu\text{eq} \cdot \text{L}^{-1}$ to $873.4 \mu\text{eq} \cdot \text{L}^{-1}$. Except for SO_4^{2-} , the equivalent concentrations of other ions show a large decreasing trend, and the change trend of the equivalent concentrations of other ions is not very large, and SO_4^{2-} was mainly affected by human sources, with a contribution rate of 87.90%, while NO_3^- was almost all affected by human anthropogenic activities. Ca^{2+} , Mg^{2+} and K^+ mainly came from soil sources and anthropogenic activities. Cl^- is a marine salt ion, with 85.47% of Cl^- coming from marine sources and 14.26% from anthropogenic activities.

Keywords: atmospheric precipitation; chemical composition; Weihai City

(责任编辑 李维卫)

(上接第224页)

Abstract ID: 1673-8020(2022)03-0218-EA

Centralized Drug Purchase Price Index of Shandong Province Based on ARIMA and Holt Index Smoothing Model

CAI Wending, WEI Xinjiang

(School of Mathematics and Statistics Science, Ludong University, Yantai 264039, China)

Abstract: This paper realizes the dynamic analysis and forecast of the centralized drug purchase prices in Shandong Province during the 13th Five-Year Plan period. By using the data provided by the procurement platform, the chain Laspeyres price index and Paasche price index were compiled, and the price index was used as an important index basis for the prediction and analysis of centralized drug purchase price, so as to carry out quantitative evaluation. Combining the ARIMA with Holt exponential smoothing model to predict the existing drug price index data, the average absolute errors of the two models are 0.361 1 and 0.372 3, respectively, which indicates that the models effectively fit the overall trend of the sequence, and extract well the dynamic and persistent characteristics of price index series. With the trend prediction of the price index, the implementation effect of drug price stabilization policies of the government and relevant departments are evaluated, and scientific decision-making basis is provided for the government and platform managers.

Keywords: centralized drug purchase in Shandong Province; price index; ARIMA; Holt two-parameter exponential smoothing

(责任编辑 顾建忠)