

基于 ARIMA 和 GM(1,1) 模型的陕西省个人卫生支出预测

于菲^{1*} 耿顺利¹ 高建民² 范小静¹ 董琬月¹ 吕强¹

1. 西安交通大学医学院 陕西西安 710061

2. 西安交通大学公共政策与管理学院 陕西西安 710049

【摘要】目的:评价陕西省疾病经济负担的变化趋势,为卫生管理者进行政策调整和干预提供理论依据。方法:采用 ARIMA 和 GM(1,1) 模型对 2016—2020 年陕西省个人现金卫生支出和其占卫生总费用比例进行测算。结果:2016—2020 年间,陕西省 OOP 占卫生总费用比重将由 32.69% 降至 26.92%。结论:陕西省居民 OOP 占卫生总费用比重在 2016—2020 年不断下降,表明陕西省居民个人卫生支出负担正逐渐减轻,卫生筹资结构逐步趋于合理。建议进一步加大政府和社会卫生投入,合理提高社会医疗保险报销比例,充分发挥社会医疗保险的作用。

【关键词】个人卫生支出;疾病经济负担;时间序列预测

中图分类号:R197 文献标识码:A doi:10.3969/j.issn.1674-2982.2018.07.004

Prediction on the personal out-of-pocket healthcare expenses in Shaanxi Province based on ARIMA and GM (1, 1) model

YU Fei¹, GENG Shun-li¹, GAO Jian-min², FAN Xiao-jing¹, DONG Wan-yue¹, LV Qiang¹

1. Xi'an JiaoTong University Health Science Center, Xi'an Shanxi 710061, China

2. School of Public Policy and Administration, Xi'an JiaoTong University, Xi'an Shanxi 710049, China

【Abstract】 Objective: This paper evaluates the changing trend of medical economic burden in Shaanxi province and provides theoretical basis and policy suggestions for policy adjustment and intervention of health managers. Methods: To forecast and evaluate OOP and account for the proportion of total health expenditure in 2016—2020 in Shaanxi Province through the methods of ARIMA, GM (1,1). Results: The forecast shows that from 2016 to 2020, the proportion of OOP in NHE will reduce from 32.69% to 26.92%. Conclusions: Shaanxi Province residents' OOP accounting for the proportion of total health expenditure in 2016—2020 will continue to decline. It shows that the burden of personal health expenditure is gradually alleviating in Shaanxi residents, and the health financing structure is becoming more and more reasonable. We must further increase government and social health input, reasonably improve the proportion of reimbursement of social medical insurance, and give full play to the supplementary role of social medical insurance.

【Key words】 Out-of-Pocket payments; Economic burden of disease; Time series prediction method

居民个人现金卫生支出(OOP)是指居民在接受医疗卫生服务时所承担的现金费用,包括直接购买卫生服务的花费以及在各种医疗保障制度下共付的费用^[1],它是卫生筹资来源的重要组成部分。研究 OOP 占卫生总费用比重,用来衡量不同经济水平下居民发生“因病致贫”、“因病返贫”风险的大小^[2];

是衡量政府对卫生领域投入力度、评价新一轮医疗体制改革成效的重要指标。新医改政策在卫生筹资领域提出要加强各级政府的医疗卫生投入力度,降低 OOP 占卫生总费用的比例,明确要求要将该比例降到 30% 以下,然而现实情况与该目标仍然存在一定差距。本研究在“十三五”开端,新医改实施 8

* 作者简介:于菲,女(1993 年—),硕士研究生,主要研究方向为卫生经济与卫生政策。E-mail: 276531617@qq.com
通讯作者:高建民。E-mail: gaojm@mail.xjtu.edu.cn

年之际,对 2016—2020 年陕西省人均实际 OOP 及其占人均实际卫生总费用比重进行科学预测,评价陕西省居民疾病经济负担的变化趋势,为卫生管理者进行政策调整和干预提供理论依据和政策建议。

1 资料与方法

1.1 资料来源

利用 1995—2015 年《陕西省卫生计生统计年鉴》、《陕西省卫生总费用研究报告》、《陕西省统计年鉴》和陕西省统计局网站发布的相关统计公报中卫生总费用相关时间序列数据。

1.2 研究方法

本研究应用 ARIMA 模型对陕西省居民人均实际 OOP 进行预测;随后,利用灰色系统理论的 GM(1,1) 模型对人均卫生总费用进行预测。最后预测二者比值。

1.2.1 ARIMA 模型

ARIMA 模型预测精度较高,是发展比较成熟的时间序列预测模型,因此被广泛应用于卫生总费用、卫生人力资源等卫生领域的预测分析。^[3] ARIMA(p, d, q) 模型的实质是对原序列进行 d 阶差分后,把 $\Delta^d y_t$ 作为因变量所建立的 ARMA(p, q) 模型,其中 p 为自回归项, q 为移动平均项。^[4]

ARIMA 模型的一般表达式为:

$$\Delta^d y_t = \varphi_1 \Delta^d y_{t-1} + \varphi_2 \Delta^d y_{t-2} + \dots + \varphi_p \Delta^d y_{t-p} + u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2} + \dots + \theta_q u_{t-q} \quad (1)$$

本研究中,利用 1995—2015 年陕西省人均实际 OOP 数据进行 ARIMA 模型的拟合,利用 Eviews6.0 对数据进行处理,预测陕西省 2016—2020 年人均实际 OOP 预测。

1.2.2 灰色预测模型

在医疗卫生领域中,卫生总费用的测算受到多方因素的影响,具有典型灰色系统的特征。^[5] 本研究采用基于灰色系统理论的 GM(1,1) 模型对陕西省 2016—2020 年人均实际卫生总费用进行预测,利用 GTMS3.0 对数据进行处理和建模。

模型的一般建模步骤为:

(1) 数据的预处理:

为了弱化原始数列内在的随机性,累加生成 (AGO) 获得灰色数列。

设原始数列为: $x^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n))$, 令:

$$x^{(1)}(k) = \sum_{i=1}^k x^{(0)}(i), k = 1, 2, \dots, n \quad (2)$$

$$x^{(0)} = (x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n)) \quad (3)$$

称所得到的数列 $x^{(1)}$ 为原始数列 $x^{(0)}$ 的一次累加数列 (1-AGO)。

(2) 构建微分方程

设 $z^{(1)} = (z^{(0)}(1), z^{(0)}(2), \dots, z^{(0)}(n))$, 令:

$$z^{(1)}(k) = \frac{1}{2}x^{(1)}(k) + \frac{1}{2}x^{(1)}(k-1), \quad k = 2, 3, \dots, n \quad (4)$$

称 $z^{(1)}(k)$ 为 $x^{(1)}$ 数列的紧邻均值生成数列,得到 GM(1,1) 模型的灰微分方程为:

$$x^{(0)}(k) = az^{(1)}(k) = b \quad (5)$$

式中: a 被称为发展系数, b 为灰色作用量。

(3) 参数估计

为了求得参数 a, b 的值,构建矩阵向量,表示如下:

$$u = \begin{bmatrix} a \\ b \end{bmatrix}, Y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix}, B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix} \quad (6)$$

利用最小二乘估计法即可求得向量的估计值。

(4) 预测值的还原

将参数 a, b 值带入灰微分方程,则可解得:

$$\hat{x}^{(1)}(k+1) = \left(x^{(0)}(1) - \frac{a}{b}\right)e^{-ak} + \frac{a}{b}, \quad k = 1, 2, \dots, n-1 \quad (7)$$

$$\hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k), \quad k = 1, 2, \dots, n-1 \quad (8)$$

其中, $\hat{x}^{(1)}(k+1)$ 为生成数列的预测值,对其进行累减还原,最终可得原始数列的预测值 $\hat{x}^{(0)}(k+1)$ 。

(5) 模型的检验

利用后验差比值 c 和小概率误差 p 两个指标对模型进行预测精度检验,从而判断模型预测的可靠性。

1.3 统计软件

使用 Excel 2010、Eviews 6.0、SPSS18.0、GTMS3.0 等统计分析软件进行处理和分析。

2 结果

在进行分析前,本研究首先利用 GDP 平减指数 (以 1978 年作为基年) 来消除价格因素对卫生支出的作用,进一步通过对 OOP 取人均值来消除人口因素对其的作用,得到陕西省居民人均实际 OOP 数据。

2.1 基于 ARIMA 模型的人均实际 OOP 预测

2.1.1 序列的平稳化处理和平稳性检验

首先对原始 OOP 序列进行自然对数转换,消除可能存在的异方差,生成新的序列为 lnOOP,通过 ADF 单位根检验来确定序列的平稳性。由检验结果可知:序列(lnOOP 序列)ADF 值大于 5% 检验临界值,*P* 值为 0.999 3,明显存在单位根,为非平稳序列;对 lnOOP 序列进行一阶差分转换后的 dlnOOP 序列则通过了单位根检验,ADF 值小于 5% 检验临界值,*P* 值为 0.000 1,判定为平稳序列,因此可以对 dlnOOP 序列进行 ARIMA 建模。

2.1.2 模型定阶及参数估计

绘制一阶差分 dlnOOP 序列的自相关和偏相关图(图 1),发现在 12 个滞后期内,自相关函数值和偏相关函数值从 2 阶起逐渐衰减,所以初步确定模型为 ARIMA(2,1,2)模型,充分考虑其他可能的阶数,最终模型的确定需根据赤池信息准则(AIC 准则)和贝叶斯信息准则(BIC 准则)来确定,一般认为 AIC 和 BIC 值越小,模型对数据的解释程度越高,拟合优良性越好,由表 1 可知,本研究最终选择 ARIMA(1,1,2)模型。

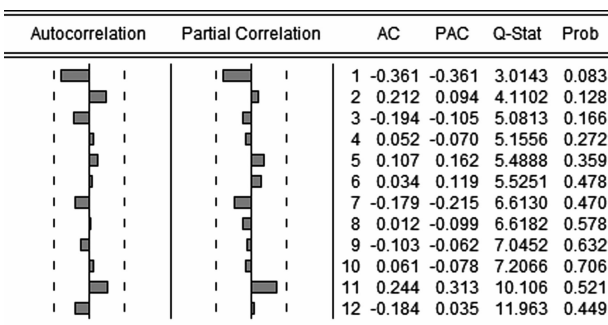


图 1 lnOOP 序列一阶差分自相关图

表 1 ARIMA 模型的选择

模型	AIC 值	BIC 值
ARIMA(1,1,1)	-2.249	-2.150
ARIMA(2,1,1)	-2.429	-2.281
ARIMA(1,1,2)	-2.686	-2.537
ARIMA(2,1,2)	-2.337	-2.140

ARIMA(1,1,2)模型参数的最小二乘估计结果如下(括号内为各系数的 *t* 值):

$$d\ln OOP_t = 0.990d\ln OOP_{t-1} + u_t - 1.878u_{t-1} + 0.931u_{t-2} \quad (9)$$

(123.948) (-63.823) (35.548)

且模型中的三个参数均通过显著性检验。

2.1.3 白噪声检验

对 ARIMA(1,1,2)模型回归后的残差进行白噪声检验,结果如图 2 所示,发现自相关的 Q 检验在 12 个滞后期内,残差没有显著的自相关,各滞后期 *P* 值

均大于 0.05,表明模型回归后的残差为白噪声序列,ARIMA(1,1,2)模型已将序列中的信息提取完毕。

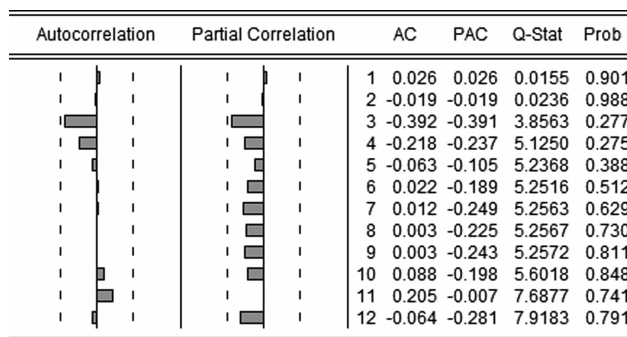


图 2 ARIMA 模型残差自相关图

2.1.4 模型拟合效果评价

利用 ARIMA(1,1,2)模型对 1995—2015 年人均实际 OOP 数据进行样本内拟合预测,并绘制拟合效果图(图 3),发现 ARIMA(1,1,2)模型的拟合效果较优。利用均方根误差(RMSE)、平均绝对误差(MAD)、平均绝对百分误差(MAPE)三个指标对该模型预测精度进行评价。其中: x_t 为实际值, \hat{x}_t 为预测值。

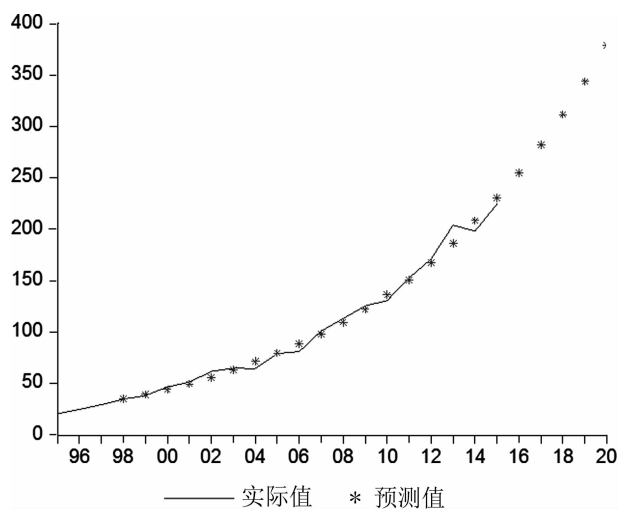


图 3 ARIMA(1,1,2)模型拟合效果图

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (x_t - \hat{x}_t)^2} \quad (10)$$

$$MAD = \sum_{t=1}^n \frac{|x_t - \hat{x}_t|}{n} \quad (11)$$

$$MAPE = \sum_{t=1}^n \left| \frac{x_t - \hat{x}_t}{x_t} \right| \times \frac{100}{n} \quad (12)$$

经计算该模型 RMSE = 6.08, MAD = 4.10, MAPE = 4.38,三个指标都较小,说明模型预测精度较高,可以用来做样本外的短期预测,进一步对陕西省 2016—2020 年人均实际 OOP 进行预测(表 2)。

表2 陕西省 2016—2020 年人均实际 OOP ARIMA(1,1,2)模型预测结果(元)

年份	实际值	预测值	年份	实际值	预测值
1995	20.55	—	2008	113.14	109.53
1996	24.77	—	2009	125.42	122.18
1997	29.47	—	2010	130.56	136.28
1998	34.84	35.04	2011	152.71	150.88
1999	38.31	39.32	2012	171.18	167.54
2000	46.58	43.97	2013	204.10	186.15
2001	51.14	49.50	2014	198.16	208.48
2002	61.79	55.68	2015	224.59	230.65
2003	64.73	63.17	2016	—	254.92
2004	64.63	71.32	2017	—	281.92
2005	79.11	79.43	2018	—	311.47
2006	80.92	88.85	2019	—	343.77
2007	101.04	98.27	2020	—	379.05

2.2 基于 GM(1,1)模型的人均实际 OOP 占卫生总费用比重预测

个人现金卫生支出的绝对值在一定程度上可以反映居民疾病经济负担的大小,但由于不同国家、地区间的卫生总费用、家庭人均收入等存在较大差异,因此仅利用绝对值来描述“看病贵”的程度不具有说服力和可比性,因此国内外应用较为广泛的指标是 OOP 占卫生总费用比例^[6],常用来作为评价居民经济负担的显示性指标^[7]。

2.2.1 构建灰色生成数列

设原始数据(1995—2015 年人均实际卫生总费用)为 $x^{(0)}$,对 $x^{(0)}$ 作一次累加计算,得到累加生成数列 $x^{(1)}$,原始数列和累加生成序列如表 3 所示。

表3 原始数列和生成的一次累加数列

年份	原始数列 $x^{(0)}$	一次累加数列 $x^{(1)}$	年份	原始数列 $x^{(0)}$	一次累加数列 $x^{(1)}$
1995	40.73	40.73	2006	153.21	1 073.10
1996	45.66	86.39	2007	202.03	1 275.13
1997	52.24	138.63	2008	247.37	1 522.50
1998	61.38	200.01	2009	307.32	1 829.81
1999	65.74	265.76	2010	326.51	2 156.33
2000	76.85	342.61	2011	392.69	2 549.02
2001	89.02	431.63	2012	450.60	2 999.62
2002	108.30	539.93	2013	525.52	3 525.14
2003	121.32	661.25	2014	582.16	4 107.30
2004	120.42	781.67	2015	684.81	4 792.11
2005	138.22	919.89			

2.2.2 构建 GM(1,1)模型

利用最小二乘估计法对模型参数 a (发展系数)、 b (灰色作用量)进行估计,结果为 $a = -0.147$

$8, b = 31.6402$ 。于是得到 GM(1,1)的白化微分方程和时间响应式分别为:

$$\frac{dx^{(1)}}{dt} - 0.1478dx^{(1)} = 31.6402 \quad (13)$$

$$x(k+1) = 254.8044e^{0.1478k} - 214.0744 \quad (14)$$

2.2.3 数据拟合、预测精度评价及样本外预测

利用 GM(1,1)模型对陕西省 1995—2015 年人均实际卫生总费用进行预测拟合,拟合效果如图 4。经过计算,该预测模型的平均绝对百分误差为 5.71,后验差比值 c 为 0.064,小概率误差 p 值为 1,故该模型的预测精度较高。故利用该模型对 2016—2020 年的人均实际卫生总费用进行预测,并计算 2016—2020 年个人现金卫生支出占卫生总费用比例,结果显示陕西省 OOP 占卫生总费用比例在 2016—2020 年将延续之前的下降趋势,并有望在 2018 年降到 30% 以内(表 4)。

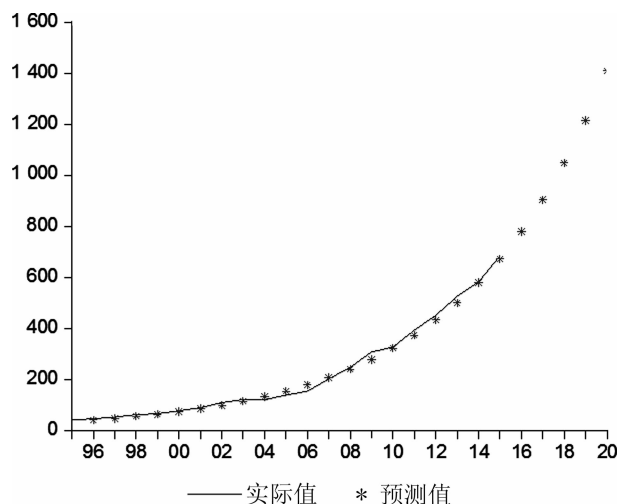


图4 GM(1,1)模型拟合效果

表4 2016—2020 年人均实际卫生费用及 OOP 占卫生总费用比例预测值

年份	人均实际卫生总费用(元)	OOP 占卫生总费用比例(%)
2016	779.78	32.69
2017	903.96	31.19
2018	1 047.93	29.72
2019	1 214.82	28.30
2020	1 408.29	26.92

3 讨论

3.1 2016—2020 年陕西省居民 OOP 占卫生总费用比重变化

世界卫生组织在《西太平洋地区和东南亚地区国

家卫生筹资策略(2006—2010)》中指出:只有居民个人现金卫生支出占卫生总费用比重在 30% 以下时,人民群众才能够通过公共筹资渠道享受比较公平且优质的医疗卫生服务。^[9]由预测结果可知,2016—2020 年陕西省居民 OOP 占卫生总费用比重会延续之前下降的趋势,预计到 2018 年该比重会首次降到 30% 以内的目标,到 2020 年该比重将会进一步降低至 26.92%。这说明陕西省居民个人卫生支出负担正逐渐减轻,筹资结构通过不断调整正逐步趋于合理。

3.2 继续加大对医疗卫生的投入力度

尽管陕西省居民个人现金支出占卫生总费用的比例呈逐年下降趋势,并且有望于 2018 年实现世界卫生组织提出的 30% 以下的目标。但仍需进一步调整和优化陕西省的卫生筹资结构,提高政府和社会卫生投资,鼓励社会资金进入,以促进医疗卫生服务的可及性,增加居民对卫生服务的利用,减少居民的自付比例,进一步优化陕西省卫生筹资结构,逐步摆脱对个人现金筹资渠道的过分依赖,让居民的个人现金卫生支出比例回归到合理的水平。

3.3 合理提高社会医疗保险补偿比例,发挥商业健康保险补充作用

社会医疗保险和商业健康保险是对医疗费用的有效分担^[10],医保补偿很大程度上降低了居民的疾病经济负担。不断提高社会基本医疗保险的报销比例,以减轻参保人员的经济负担;同时推进城乡医保并轨,杜绝重复参保现象,避免政府重复补助。商业健康保险应加强专业化管理和险种改革,刺激居民的购买需求,鼓励政府、医院和商业健康保险合作,减轻政府卫生筹资压力,分担风险,充分发挥其补充作用,

进一步降低居民个人卫生支出,减轻疾病经济负担。

作者声明本文无实际或潜在的利益冲突。

参 考 文 献

- [1] 周书美, 范艳存, 孙静, 等. 2007—2013 年内蒙自治区居民个人现金卫生支出研究[J]. 中国卫生经济, 2015, 34(12): 63-65.
- [2] 徐明江. 1978 ~ 2011 年我国个人卫生支出的影响因素研究[D]. 南宁: 广西医科大学, 2014.
- [3] Box G E P. Time seriesanalysis: forecasting and control/-Rev. ed[M]. Holden-day, 1976.
- [4] 童光荣, 何耀. 计量经济学实验教程[M]. 武汉: 武汉大学出版社, 2008.
- [5] 刘思峰. 灰色系统理论及其应用[M]. 北京: 科学出版社, 2010.
- [6] 张录法. 新医改后中国个人现金卫生支出的趋势分析[J]. 统计与信息论坛, 2010(12): 52-56.
- [7] 张毓辉, 翟铁民, 魏强, 等. 个人卫生支出比重与居民疾病经济负担关系的案例研究[J]. 卫生经济研究, 2011(6): 18-21.
- [8] 徐高强. 山西省煤矿区矿井水涌水量预测模型研究[D]. 太原理工大学, 2008.
- [9] 高建民, 张文, 杨进. 陕西省卫生总费用筹资水平及结构研究[J]. 中国卫生经济, 2011, 30(5): 19-21.
- [10] 刘芳芳. 浅析我国商业健康保险对社会医疗保险的补充作用[J]. 中国卫生政策研究, 2010, 3(7): 38-43.

[收稿日期: 2017-07-23 修回日期: 2018-03-24]

(编辑 薛云)