

ARIMA 季节乘积模型在儿童肺炎 门急诊人次预测中的应用

杨蕾¹, 吴文华², 任泉³, 王艳杰¹, 刘早玲¹

1. 新疆医科大学公共卫生学院, 新疆 乌鲁木齐 830011;

2. 新疆医科大学第一附属医院信息中心, 新疆 乌鲁木齐 830054; 3. 乌鲁木齐市气象局, 新疆 乌鲁木齐 830002

摘要: **目的** 探索自回归差分移动平均 (autoregressive integrated moving average, ARIMA) 季节乘积模型在预测儿童肺炎门急诊人次的应用, 为合理利用医疗资源提供科学依据。 **方法** 收集乌鲁木齐市两家三级甲等医院 2011–2016 年儿童肺炎逐月门急诊人次数据, 使用 R 3.4.1 软件进行模型的识别、参数估计与检验, 建立 ARIMA 季节乘积模型对 2011 年 1 月–2016 年 6 月儿童肺炎逐月门急诊人次进行拟合, 并利用 2016 年 7–12 月数据计算预测值与实际值的平均预测相对误差来评价预测效果。 **结果** ARIMA(0,1,2)(1,1,0)₁₂ 模型是拟合儿童肺炎门急诊人次的最佳预测模型, 平均相对误差为 9.82%。 **结论** ARIMA 季节乘积模型有较好的拟合和短期预测效果, 能为医院合理利用医疗资源提供参考依据。

关键词: 儿童肺炎; 门急诊人次; ARIMA 季节乘积模型; 预测

中图分类号: R181.2 **文献标识码:** A **文章编号:** 1006–3110(2019)01–0033–04 **DOI:** 10.3969/j.issn.1006–3110.2019.01.009

Application of ARIMA seasonal product model to predicting the number of outpatient and emergency department visits for childhood pneumonia

YANG Lei¹, WU Wen-hua², REN Quan³, WANG Yan-jie¹, LIU Zao-ling¹

1. School of Public Health, Xinjiang Medical University, Xinjiang, Urumqi 830054, China;

2. Center for Medical Information, the First Affiliated Hospital of Xinjiang Medical University, Urumqi, Xinjiang 830011, China;

3. Urumqi Weather Bureau, Urumqi, Xinjiang 830002, China

Abstract: **Objective** To explore the application of autoregressive integrated moving average (ARIMA) seasonal product model to forecasting the number of outpatient and emergency department visits for childhood pneumonia, and to provide a scientific basis for the rational utilization of medical resources. **Methods** We collected the data regarding the number of monthly outpatient and emergency department visits for childhood pneumonia from two tertiary comprehensive hospitals in Urumqi City from 2011 to 2016. R 3.4.1 software was used for model identification, parameter estimation and verification. ARIMA seasonal product model was established to fit the number of monthly outpatient and emergency department visits for childhood pneumonia from January 2011 to June 2016, and the prediction effect was evaluated through checking the average relative error between the predicted and actual values by using the monthly data from July to December in 2016. **Results** The ARIMA (0,1,2) (1,1,0)₁₂ model was the best prediction model, with an average relative error of 9.82%. **Conclusions** ARIMA seasonal product model has better fitting and short-term prediction effect, which can provide references for hospitals to rationally utilize medical resources.

Key words: childhood pneumonia; the number of outpatient and emergency department visits, ARIMA seasonal product model; prediction

肺炎是婴幼儿时期常见和多发的感染性疾病,是发展中国家儿童死亡的主要原因。根据联合国儿童基金会的数据统计,每年全世界约有 350 万例 5 岁以下儿童死于肺炎,占 5 岁以下儿童总死亡数的 28%^[1]。我国仍属于儿童肺炎高负担国家,每年的肺炎患儿总计 2 100 万人次,是全国 5 岁以下儿童第二位的死亡

原因^[2],尤其是西部及边远地区,5 岁以下儿童肺炎的发病率、死亡率和病死率均高于全国平均水平。为了解乌鲁木齐市儿童肺炎门急诊人次的变化规律,本研究收集 2011–2016 年乌鲁木齐市两家三级甲等医院儿科肺炎月门急诊人次数据,采用 ARIMA 乘积季节模型建立乌鲁木齐市儿童肺炎门急诊人次预测模型,为有关部门预防儿童肺炎提供依据。

基金项目: 新疆维吾尔自治区自然科学基金 (2016D01C188); 新疆医科大学科研创新基金 (XYDCX201512)

作者简介: 杨蕾 (1981–), 女, 硕士, 讲师, 研究方向: 环境与人类健康。

1 资料与方法

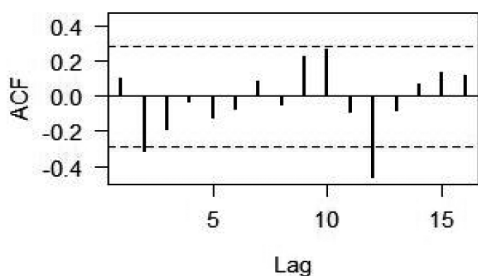
1.1 资料来源 以新疆医科大学第一附属医院和乌鲁木齐市北门儿童医院 2011-2016 年儿科门急诊就诊的患儿为研究对象,收集第一诊断为肺炎的患儿信息,以月门急诊人次构成时间序列。

1.2 ARIMA 乘积季节模型的构建 ARIMA 乘积季节模型即 $ARIMA(p,d,q)(P,D,Q)_s$ 模型为^[8]:

$$\phi_p(B)\Phi_p(B^s)(1-B)^d(1-B^s)^D Y_t = \theta_q(B)\Theta_q(B^s)e_t$$

其中 p 、 d 和 q 分别为非季节自回归阶数、差分次数和滑动平均的阶数; P 、 D 和 Q 分别表示季节自回归阶数、季节差分次数和季节滑动平均阶数, B 为后移算子, s 为季节周期。

ARIMA 乘积季节模型建模步骤:(1)非平稳序列的平稳化:通过观察原始时间序列图判断其是否需要数据转换或差分,对差分后的序列做 ADF 单位根检验, $P < 0.05$ 时可认为序列平稳;(2)ARIMA 季节乘积模型的识别:通过观察新序列的自相关系数(ACF)图和偏自相关系数(PACF)图,选择适当阶数的 ARIMA(p,d,q)(P,D,Q) $_s$ 模型;(3)参数估计:运用最小二乘估计法充分利用序列的信息对模型中未知参数进行估计;(4)模型与参数显著性检验:如果模型残差为白噪声序列则说明模型合理。参数的显著性检验就是要检验每一个未知参数是否显著,使模型最精简,当 $P < 0.05$ 时可认为参数显著;(5)模型优化:在所有通过模型检验的 ARIMA(p,d,q)(P,D,Q) $_s$ 模型中根据赤池信息准则(AIC)、基于改进的 Akaike 信息准则(AICC)及贝叶斯信息准则(BIC)函数值最小原则优化模型;(6)模型拟合与预测:以 2011 年 1 月-2016 年 6 月逐月数据进行模型拟合,利用 2016 年 7-12 月逐月数据检验模型预测效果,根据预测值与实际值的平均预测相对误差评价预测效果。



1.3 统计学处理 采用 R 3.4.1 统计软件进行数据处理及建模分析,主要采用 stats、forecast 和 tseries 程序包中的函数进行时序数据的描述和预测。

2 结果

2.1 2011-2016 年儿童肺炎月门急诊人次情况 绘制 2011 年 1 月-2016 年 6 月乌鲁木齐市两家三甲医院儿童肺炎门急诊人次的时序图,见图 1。该时序图有如下特点:(1)序列存在明显的周期性。各月门急诊人次呈现出明显的季节性,冬春高发,其中 12-5 月门急诊人次最多,6-10 月最低,表现出每 12 个月出现 1 次明显的高峰。(2)序列整体呈逐年上升趋势。以上特点提示原始序列是非平稳的时间序列。

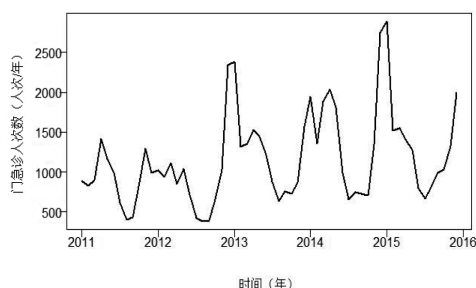


图 1 儿童肺炎门急诊人次时间序列图

2.2 ARIMA 季节乘积模型的建立

2.2.1 序列的平稳化 2011-2016 年儿童肺炎的月门急诊人次为具有明显的长期趋势与季节性的非平稳序列,建模前先对该序列进行一阶差分和一阶周期为 12 的季节性差分处理使序列平稳化,差分后序列的自相关(ACF)图和偏自相关(PACF)图(图 2),并对差分后序列进行 ADF 单位根检验,结果显示 $P = 0.0102$,说明差分后序列平稳,可以构建 ARIMA 季节乘积模型。

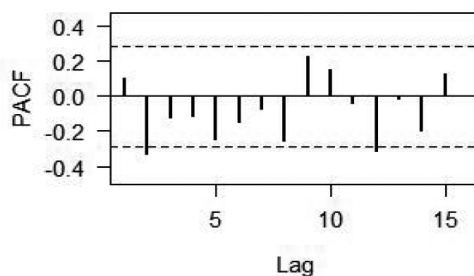


图 2 差分序列的 ACF 图和 PACF 图

2.2.2 ARIMA 乘积季节模型识别 根据儿童肺炎门急诊人次差分后序列的自相关图和偏自相关图初步判断非季节性 ARIMA 模型中 p, q 值分别为 0 和 2。由于原始非平稳序列经过一阶差分和一阶季节性差分处理后达到平稳状态,因此可以确定 ARIMA(p, d, q)

(P, D, Q) $_{12}$ 模型中 $d=1, D=1$ 。季节模型的参数 P, Q 一般不超过二阶,因而分别取 0、1、2 由低阶到高阶逐个测试,通过对模型的残差进行自相关检验,根据拟合优度检验结果以 AIC、BIC 和 AICC 最小准则来选择模型。

2.2.3 参数估计、模型检验及优化 经过筛选符合条件的 ARIMA 季节乘积模型有九个,最终确定最佳模型为 $ARIMA(0,1,2)(1,1,0)_{12}$,见表 1。该模型的 $AIC=765.27$, $AICC=766.1$, $BIC=773.15$,模型中各参数检验均有统计学意义 ($P<0.05$),见表 2。采用 Ljung-Box 方法对模型残差进行检验,检验统计量为 6.302 2, $P=0.390 2$,同时对残差序列做自相关图,可见残差序列的自相关系数截尾,说明序列信息已经被充分提取,模型残差为白噪声,见图 3。

表 1 残差符合条件的各 ARIMA 季节乘积模型及 AIC、AICC 和 BIC 的值

模型	AIC	AICC	BIC
$ARIMA(0,1,2)(0,1,0)_{12}$	777.1	777.5	783.0
$ARIMA(0,1,2)(1,1,0)_{12}$	765.2	766.1	773.1
$ARIMA(0,1,2)(0,1,1)_{12}$	769.3	770.1	777.1
$ARIMA(0,1,2)(1,1,1)_{12}$	766.7	768.1	776.6
$ARIMA(0,1,2)(0,1,2)_{12}$	767.9	769.2	777.8
$ARIMA(0,1,2)(2,1,0)_{12}$	766.7	767.9	776.5
$ARIMA(0,1,2)(1,1,2)_{12}$	768.7	770.5	780.5
$ARIMA(0,1,2)(2,1,1)_{12}$	768.6	770.5	780.4
$ARIMA(0,1,2)(2,1,2)_{12}$	770.6	773.1	784.4

表 2 $ARIMA(0,1,2)(1,1,0)_{12}$ 模型参数检验

模型参数	回归系数	标准误	t 值	P 值
MA(1)	-0.300 6	0.121 5	-2.474	0.016 1
MA(2)	-0.562 4	0.118 5	-4.745	<0.000 1
SAR(1)	-0.551 8	0.121 1	-4.558 2	<0.000 1

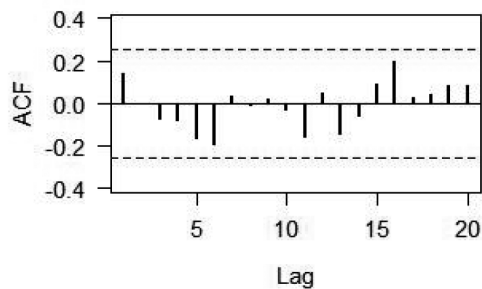


图 3 残差的自相关图

2.2.4 模型预测及应用 运用 $ARIMA(0,1,2)(1,1,0)_{12}$ 模型对 2016 年 7-12 月逐月门急诊人次进行短期预测,结果显示各月实测值均落入预测值的 95% 置信区间范围,平均相对误差为 9.82%,模型拟合效果较好。见表 3。

表 3 2016 年 7-12 月儿童肺炎逐月门急诊人次实际值与预测值比较

时间(年.月)	实际值	预测值	相对误差(%)	预测值的 95%CI
2016.7	645	705	9.3	122~1 289
2016.8	620	685	10.4	27~1 397
2016.9	698	789	13.0	72~1 505
2016.10	794	788	-0.7	67~1 510
2016.11	1 323	1 199	-9.4	473~1 924
2016.12	2 624	2 200	-16.1	1 469~2 930

3 讨论

儿童肺炎是严重的儿童呼吸道疾病,我国卫生部把肺炎列为了包括腹泻、佝偻病和贫血在内的小儿四病防治方案中的首位。尽管随着安全有效的抗生素和疫苗的使用,儿童肺炎的预防和治疗取得了卓有成效的成绩,但在全球范围内,肺炎仍是 5 岁以下儿童死亡的首位独立死因,每年导致约 200 万人死亡^[3],肺炎是中国成人和儿童的主要死因之一^[4]。Leni Kang 等^[5]发表的最新研究表明,2014-2015 年间我国 5 岁以下儿童肺炎的死亡率 162.8/10 万,并且由于经济条件、医疗保障水平和环境条件等的差异,不同地区儿童肺炎发病率与死亡率呈现明显差异。西部及偏远地区儿童肺炎的死亡率最高,分别是中部地区的 2.5 倍和东部地区的 7.5 倍^[6]。我国城市 5 岁以下儿童肺炎的死亡率为 64/10 万,农村为 260.9/10 万^[5]。全球估计中国 5 岁以下儿童每年有 2 110 万临床肺炎新发病例(0.22 次/人年),在疾病负担上仅次于印度(4 300 万新发病例,0.37 次/人年)^[7-8]。有研究显示我国肺炎发病整体上呈现稳定或者下降的趋势,但是肺炎依然是威胁我国儿童健康的重要公共卫生问题。

求和自回归移动平均 (ARIMA) 模型是由美国学者 Box 和英国统计学家 Jenkins 于 20 世纪 70 年代初提出的经典时间序列预测方法,其基本思想是将预测对象随时间推移而形成的数据视为一组依赖于时间 t 的随机变量,是自回归模型与移动平均模型的有效组合。ARIMA 模型可以分为简单模型和乘积季节模型,当发现序列在短期相关性和季节效应之间具有乘积关系时,用 ARIMA 乘积季节模型可以更充分地提取趋势信息和季节信息,较好地进行序列拟合及预测,是实用性强、短期预测精确度高的时间序列预测方法之一,已被广泛应用于疾病发病率、门诊或出院量、住院费用等预测^[9-11]。ARIMA 乘积季节模型主要拟合具有趋势性和季节性规律波动的序列,其结合了自回归和移动平均法的长处,具有不受数据类型束缚和适用性强的特点。本文通过对某三甲医院 2011-2016 年儿童肺炎逐月门急诊人次建立 ARIMA 季节乘积模型,序列经过平稳化、模型识别、参数估计及优化等步骤,最终选取最优模型 $ARIMA(0,1,2)(1,1,0)_{12}$,结果表明模型能较好地拟合出儿童肺炎门急诊人次和季节趋势,拟合值与实际值的动态变化规律相同。通过对儿童肺炎门急诊人次的预测,可以优化资源分布,合理利用有限的儿科医疗资源。

参考文献

吉林市中小學生視力不良現狀及 增長速度趨勢分析

周歌¹, 吳方園¹, 劉美田¹, 錢曉波², 張秀敏¹, 劉紅箭¹

1. 吉林大學公共衛生學院, 吉林 長春 130021; 2. 吉林省健康教育中心, 吉林 長春 130062

摘要: **目的** 了解吉林市中小學生視力不良現狀並擬合視力不良及各程度不良檢出率隨年齡增長速度曲線模型, 為制定視力不良防控策略提供科學依據。 **方法** 於 2015 年 4-5 月, 採用多階段分層整群隨機抽樣方法, 抽取吉林市船營區、丰满區、蛟河市和永吉縣所在地共 9 676 名中小學生進行視力檢測和問卷調查, 分析學生視力不良檢出情況。利用曲線擬合的方法獲得視力不良及各程度不良的增長速度模型, 分析其增長速度在城鄉男女生間的變化規律。 **結果** 吉林市中小學生視力不良檢出率為 77.9%, 輕、中和重度視力不良檢出率分別為 12.1%、17.4% 和 48.5%。多因素 logistic 回歸分析顯示, 視力不良及各程度不良檢出率的性別、城鄉差異有統計學意義 ($P < 0.05$)。視力不良及各程度不良增長速度模型顯示, 城鄉男女生視力不良檢出率增長速度曲線呈“拋物線”形, 在學齡初期和青春晚期出現“交叉現象”, 視力不良率增長速度的年齡峰值提前; 輕度視力不良率增長速度曲線在 7~15 歲呈負增長, 而城市女生、鄉村男生、鄉村女生在進入青春晚期後其增速曲線又出現增長傾向; 中度視力不良率增長速度曲線呈“U”形變化, 在 10~15 歲呈負增長; 重度視力不良率增長速度整體呈下降趨勢, 在學齡初期增速最快, 增速曲線在青春期中出現“交叉現象”。 **結論** 吉林市中小學生視力不良率隨年齡增長明顯, 增長速度曲線的年齡峰值不斷提前, 呈進一步低齡化發展, 視力不良防控重點應向學齡初期及學前傾斜。

關鍵詞: 中小學生; 視力不良; 檢出率; 增長速度

中圖分類號: R778.1⁺1 **文獻標識碼:** A **文章編號:** 1006-3110(2019)01-0036-06 **DOI:** 10.3969/j.issn.1006-3110.2019.01.010

Current status of poor vision and its trend of increasing velocity among primary and middle school students in Jilin City

ZHOU Ge¹, WU Fang-yuan¹, LIU Mei-tian¹, QIAN Xiao-bo², ZHANG Xiu-min¹, LIU Hong-jian¹

1. School of Public Health, Jilin University, Changchun, Jilin 130021, China

2. Jilin Provincial Health Education Center, Changchun, Jilin 130062, China

Corresponding author: LIU Hong-jian, E-mail: hongjianliu63@163.com

Abstract: **Objective** To analyze the prevalence of poor vision and create curve models of increasing velocity with age for

基金项目: 吉林省卫生计生软科学课题项目 (2015R022)

作者简介: 周歌 (1993-), 女, 山东济宁人, 硕士在读, 研究方向: 慢性病流行病学。

通信作者: 刘红箭, E-mail: hongjianliu63@163.com。

- [1] Black RE, Cousens S, Johnson HL, et al. Global, regional, and national causes of child mortality in 2008: a systematic analysis[J]. Lancet, 2010, 375(9730): 1969-1987.
- [2] 冯江, 袁秀琴, 朱军, 等. 中国 2000-2010 年 5 岁以下儿童死亡率和死亡原因分析[J]. 中华流行病学杂志, 2012, 33(6): 558-561.
- [3] Rudan I, O'Brien KL, Nair H, et al. Epidemiology and etiology of childhood pneumonia in 2010: estimates of incidence, severe morbidity, mortality, underlying risk factors and causative pathogens for 192 countries[J]. J Glob Health, 2013, 3(1): 010401-010415.
- [4] Duan X. Major causes of death among men and women in China [J]. NEJM, 2015, 9(15): 1124-1134.
- [5] Kang L, He C, Miao L, et al. Geographic disparities of pneumonia-specific under-5 mortality rates in Mainland China from 1996 to 2015: a population-based study[J]. J Infect Dis, 2017, 59(C): 7-13.
- [6] Liu L, Oza S, Hogan D, et al. Global, regional, and national causes of under-5 mortality in 2000-15: an updated systematic analysis with im-

- lications for the Sustainable Development Goals [J]. Lancet, 2016, 388(10063): 3027-3035.
- [7] 官旭华, Silk BJ, Li W, et al. 中国大陆肺炎发病率与死亡率: 1985-2008 年中英文文献的系统分析[J]. 公共卫生与预防医学, 2011, 22(1): 14-19.
- [8] Rudan I, Boschi-Pinto C, Biloglav Z, et al. Epidemiology and etiology of childhood pneumonia [J]. Bull World Health Organ, 2008, 86(5): 408-416.
- [9] 姜鹏威, 吴秀峰, 张学良, 等. 基于 ARIMA 乘积季节模型的新疆布鲁氏菌病流行趋势分析[J]. 新疆医科大学学报, 2017, 40(1): 86-90.
- [10] 张曼, 王云霞, 王凯, 等. ARIMA 模型预测乌鲁木齐市吸毒人群艾滋病病毒感染状况[J]. 实用预防医学, 2018, 25(3): 310-312.
- [11] 郑慧敏, 薛允莲, 黄燕飞, 等. ARIMA 模型在深圳市法定传染病发病趋势预测的应用[J]. 实用预防医学, 2016, 23(2): 240-243.

收稿日期: 2018-03-15