

# ARIMA( $p, d, q$ ) 模型在不同地区来源 门诊量预测中的应用

徐永强<sup>1</sup> 周 静<sup>2</sup> 陈学军<sup>1</sup> 潘 安<sup>1</sup>

**【摘要】** 目的 应用时间序列 ARIMA 预测模型建立医院总的月门诊量的预测模型,利用该模型对不同地区来源的门诊量进行预测,为优化医疗资源配置提供科学的统计学依据。方法 运用 SAS 9.0 软件对原始数据进行均值零化、正态性检验,并通过差分使原始序列平稳,采用条件最小二乘法估计模型参数,依据 AIC 和 SBC 准则确定模型阶数,建立 ARIMA 预测模型。结果 具体的 ARIMA 预测模型为  $(1 - 0.461\ 96B)(1 - B)x_t = (1 - 0.641\ 93B^{12})\alpha_t$ 。结论 所建立模型 ARIMA(1,1,12)可以对本院总的月门诊量及部分地区来源的月门诊量进行准确预测,使得某些地区群众看病难问题的解决得到理论上的统计学依据。

**【关键词】** ARIMA 模型 时间序列 门诊量 预测

中图分类号:R195.1 文献标识码:A 文章编号:1006-5253(2009)03-0202-05

**Applications of ARIMA( $p, d, q$ ) model on predicting mensual quantity of outpatients from different counties** XU Yong-qiang\*, ZHOU Jing, CHEN Xue-jun, et al. \*Wenzhou Medical College 2nd Affiliated Hospital, Wenzhou 325000, China

**【Abstract】 Objective** To discuss the application of time series ARIMA predictive model on mensual quantity of outpatients (MQOPTS), to predict mensual quantity of outpatients from different counties and to provide scientific, statistical foundation for optimization of medical treatment resource collection. **Methods** The total quantity of outpatients within 4 years was analyzed with SAS 9.0 software and predicted with ARIMA( $p, d, q$ ) model. ARIMA predictive model was fitted and the order of model was confirmed through Akaike Information Criterion and Schwarz Bayesian Criterion. **Results** The best ARIMA model is  $(1 - 0.461\ 96B)(1 - B)x_t = (1 - 0.641\ 93B^{12})\alpha_t$ . **Conclusion** Time series of MQOPTS can be predicted by the ARIMA(1,1,12) model with high prediction precision and this model may provide statistical foundation for solving questions of those who have difficulties to see a doctor because of living in a remote area.

**【Key words】** ARIMA model Time series Mensual quantity of outpatients Prediction

对群众就医问题的研究及医院接诊情况的分析均离不开科学的统计预测。医院门诊量是一个集中反映众多患者在某院就医频率的重要指标。通过对医院月门诊量进行短期预测,医院自身及相关医疗机构可以就合理配置医疗卫生资源、提高医疗服务水平、及时处理患者就医高峰以及提高医疗社会效益等方面都有重要意义。医院门诊量数据构成了一组按时间顺序排列的数据,即时间序列<sup>[1]</sup>。迄今为止最通用的时间序列预测法是以美国统计学家 Geogre E. Box 和 Gwilym M. Jenkins 的名字命名的预测方法,即博克斯-詹金斯法,简写作 B-J 法。是通过利用自回归移动平均模型即 ARMA 模型(Autoregressive moving-average model)进行预测。运用 ARMA 模型的前提条件是建立平稳的时间序列,而实际工作中,时间序列受各种因素影响,所以多为非平稳的。通过差分处理得到平稳时间序列建立的模型称为自回归移动积分模型,即 ARIMA 模型(Autoregressive integrated moving-average model)<sup>[2]</sup>。笔者根据所在单位近 5 年的资料,利用 ARI-

MA 模型进行月门诊量(Mensual quantity of outpatients, MQOPTS)的预测,以探讨较为适宜的预测方法。

## 1 资料

为温州医学院附属第二医院 2002-04—2007-03 的月门诊量(次月第一日的数据反映的是上一月门诊量)。资料来源为该院信息科(总的门诊量数据如表 1 所示,因限于篇幅,来自温州及其各县市区的门诊量未给出),数据确凿可靠。

## 2 方法及结果

### 2.1 ARIMA( $p, d, q$ )模型简介

对于  $d$  阶非平稳序列  $x_t$  而言,  $\nabla^d x_t$  是一个平稳序列<sup>[3]</sup>,设其适合 ARMA( $p, q$ )模型,即:

$$\begin{aligned}\Phi(B)(\nabla^d x_t) &= \Theta(B)\alpha_t \\ \Phi(B)(1-B)^d x_t &= \Theta(B)\alpha_t\end{aligned}\quad (1)$$

作者单位:1 325000 温州医学院附属第二医院  
浙江省温州市

2 温州医学院附属第一医院

其中： $\Phi(B) = 1 - \varphi_1 B - \varphi_2 B^2 - \cdots - \varphi_p B^p$ ， $\Theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q$ ，称模型(1)为 ARIMA( $p, d, q$ )模型。其中  $p$  为自回归阶数， $d$  为差分阶数， $q$  为移动平均阶数， $B$  为后移算子， $\alpha_t$  为白噪声序列。

表 1 医院月门诊量

日期	月门诊量 /人次	日期	月门诊量 /人次	日期	月门诊量 /人次	日期	月门诊量 /人次	日期	月门诊量 /人次
2002-05-01	78 162	2003-05-01	59 616	2004-05-01	71 049	2005-05-01	11 718	2006-05-01	127 424
2002-06-01	78 163	2003-06-01	70 656	2004-06-01	11 004	2005-06-01	12 350	2006-06-01	139 252
2002-07-01	78 164	2003-07-01	94 565	2004-07-01	12 626	2005-07-01	13 184	2006-07-01	162 543
2002-08-01	84 531	2003-08-01	97 661	2004-08-01	12 437	2005-08-01	14 181	2006-08-01	152 129
2002-09-01	64 970	2003-09-01	88 360	2004-09-01	10 325	2005-09-01	119 527	2006-09-01	132 245
2002-10-01	71 596	2003-10-01	89 330	2004-10-01	10 407	2005-10-01	115 777	2006-10-01	131 581
2002-11-01	75 255	2003-11-01	84 147	2004-11-01	10 664	2005-11-01	122 364	2006-11-01	137 529
2002-12-01	73 437	2003-12-01	82 360	2004-12-01	10 057	2005-12-01	112 084	2006-12-01	130 845
2003-01-01	68 629	2004-01-01	74 149	2005-01-01	96 082	2006-01-01	106 558	2007-01-01	131 755
2003-02-01	58 210	2004-02-01	82 910	2005-02-01	86 369	2006-02-01	107 532	2007-02-01	114 150
2003-03-01	81 586	2004-03-01	10 811	2005-03-01	12 358	2006-03-01	141 863	2007-03-01	133 588
2003-04-01	84 901	2004-04-01	10 747	2005-04-01	12 603	2006-04-01	129 373	2007-04-01	141 810

2.2 建立预测模型

应用 SAS 9.0 软件进行分析绘图(因该版本软件在生成的图中尚不支持汉字输入,故在本文中的图例中温州及其各地市的名称等文字由 photoshop6.0 软件写入)。

2.2.1 模型识别 该组数据(2002-04—2007-03 期间的门诊量)由该院信息科提供,并交由 SAS 程序汇总(程序省略),数据客观、确凿,故不再对时间序列的独立性进行检验。该资料为非平稳的时间序列(参考图 3),具有明显的增长趋势。对 2002-04—2006-09 的数据进行分析,2006-10—2007-03 的数据用来初步检验预测结果。利用 proc ttest 对该数据进行和总体均数 0 相比较的  $t$  检验(均值为 1 003 349,  $P$  值小于 0.000 1)发现,其均值不为 0。进行均值零化处理后,再次进行的  $t$  检验表明该时间序列的均值零化已完成(均值为 0.001 1,  $P$  值为 1.000 0),并且经 proc univariate 进行正态性分布的检验发现,该时序服从正态性分布( $W$  值为 0.978 583,  $P$  值为 0.442 1)。利用 SAS 软件强大的运算功能,proc arima 程序应用 identify 语句对时间序列进行一阶差分,从 1 ~ 15 对滞后值进行取值,并应用 stationarity = (rw) 语句检验平稳性,最终选择一阶差分的滞后值为 12, lag 值为 0、1、2 时的 Tau 值分别为 - 1.69 ( $P=0.098\ 9$ )、- 1.24 ( $P=0.223\ 9$ )、- 1.01 ( $P=0.316\ 8$ ),输出的自相关函数图和偏自相关函数图均为拖尾,且白噪声检验拒绝序列为白噪声的原假设( $\chi^2=17.06, P=0.009\ 1$ ),因此时序有一个 ARIMA 模型<sup>[4-5]</sup>。

2.2.2 模型定阶即检验 本文针对 ARIMA( $p, 1, q$ )模型,利用 SAS 软件强大的统计运算功能,  $p$  值和  $q$  值均从 0 ~ 15 分别进行取值,在残差的自相关检验为白噪声的基础上,采用 AIC 准则(Akaike's information criterion)和 SBC 准则(Sawa's bayesian criterion)判定模型阶数,最终确定:  $p=1, q=12$ 。参数估计采用系统默认的条件最小二乘法,计算各参数的  $P$  值均小于 0.05,结合程序所得出的 Autoregressive factors: 1 - 0.461 96  $B$ , Moving average factor: 1 - 0.641 93  $B^{12}$  (均由 proc arima 程序的 estimate 语句来完成),从而建立模型 ARIMA(1, 1, 12),其具体模型为:

$$(1 - 0.461\ 96B)(1 - B)x_t = (1 - 0.641\ 93B^{12})\alpha_t$$

2.3 利用确定的模型进行预测

预测值与实际值的比较由 proc arima 程序的 forecast 语句完成<sup>[6]</sup>。预测结果如图 1 所示。

2.4 利用模型 ARIMA(1, 1, 12)对来自温州各地市的患者数据进行分析和预测

温州地区包括乐清、永嘉、洞头、温州、平阳、瑞安、文成、泰顺和苍南共 9 地市,温州行政区划及 2002-04—2007-03 期间内各县市区来医院就诊总人次及其地理位置示意图(由 proc gmap 完成)如图 2 所示。在 2002-04—2007-03 期间,温州市的各个县市区每月来本院就诊的人数不尽相同(如图 3 所示),其预测情况不再赘述。其中,温州市区来源门诊量的预测结果如图 4 所示(该模型下,因为洞头和泰顺的时间序列不服从正态性分布,平阳和文成的时序数据不稳定,故没有进行预测)。

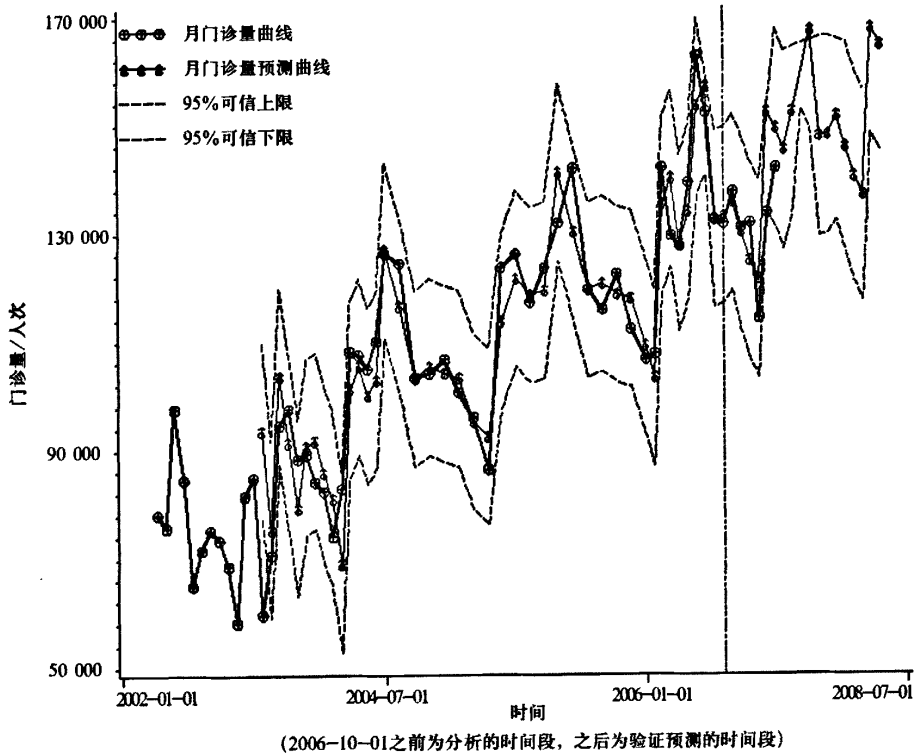


图 1 2002-04—2007-03 医院月门诊量及其预测曲线

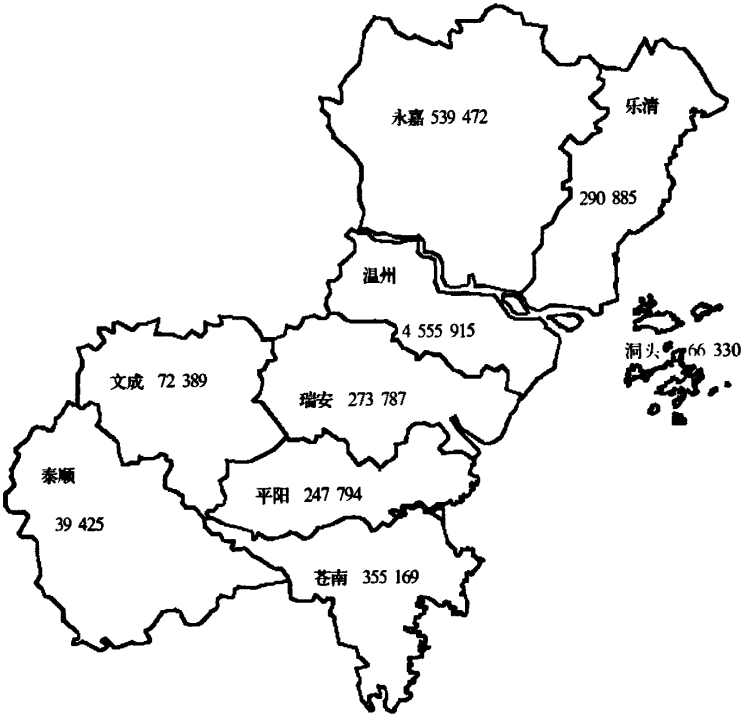
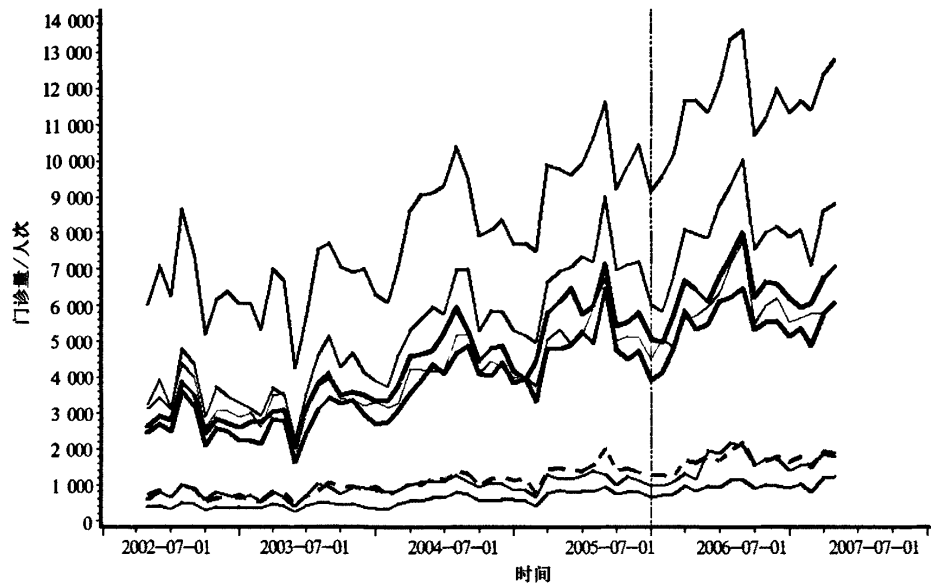
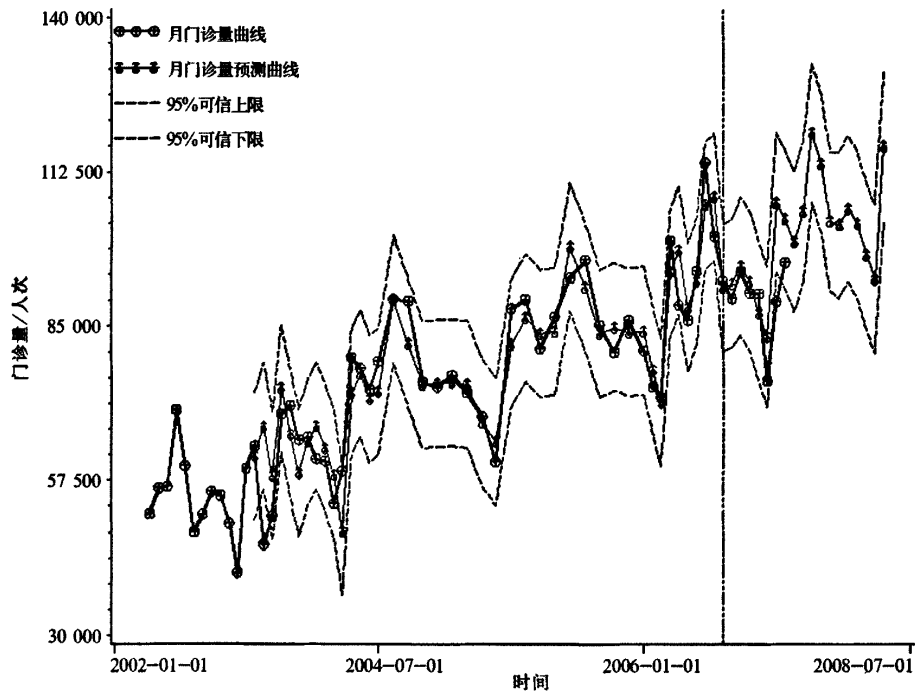


图 2 2002-04—2007-03 各县市区来医院就诊总人次及其地理位置示意图



(在2006-01-01点, 自上而下的曲线分别代表永嘉、苍南、乐清、瑞安、平阳、文成、洞头、泰顺)

图3 温州所辖8个县的来院门诊量时间序列图



(2006-10-01之前为分析的时间段, 之后为验证预测的时间段)

图4 来自温州市区的月门诊量及其预测曲线

3 讨论

现实生活中有很多数据是定期采集的观测数据, 按时间的顺序记录, 属于时间序列。采集此类数据的目的便是对时间序列进行观察、研究、分析, 找寻其中变化的规律, 预测将来的走势, 即属于时间序列分

析<sup>[7-8]</sup>。而 ARIMA 模型考虑了季节因素对时间平稳性造成的影响, 所以可以得到与实际较为吻合的预测效果<sup>[9]</sup>。

进行时间序列分析, 关键是正确建模。时间序列建模就是建立数据序列的过去值与未来值之间的联

系<sup>[10]</sup>。因为在建模的过程中可以用一系列的统计方法检验模型的适用性,以不断调整模型的阶数,直至达到满意的结果。所以在实际应用中,一般应先建立较为简单的模型。有人认为尽可能的采取  $AR(p)$  模型来拟合动态数据,可节省建模和参数估计的计算量,即使真正模型为  $MA(q)$  或  $ARMA(p,q)$  模型也可用高阶的  $AR$  模型来逼近<sup>[11]</sup>。本文选用了  $ARIMA(p,d,q)$  模型来研究笔者单位 2002-04—2006-09 之间医院门诊量的季节性变化规律。虽然由于受 SARS 疫情的影响,2003 年的门诊量较前一年明显下降,数据的增长趋势出现变化,对数据的稳定性产生了影响,然而该模型仍能较准确地进行短期预测。提示此模型对医院门诊量的长期趋势预报有一定参考价值。该模型对温州各个县市区进行预测时,部分统计过程却出现了错误的结果,例如平阳和文成两地的数据在该模型下不具有稳定性,程序得出的预测图形(文中未附)虽形象颇佳,实为一纸空文。提示用  $ARIMA$  模型作实际预测时,虽能部分发挥作用(如温州市、瑞安市、永嘉、苍南及乐清等地区的数据资料),但是当观测对象或观测范围发生变化时,还需根据实际的时间序列重新建立模型,做新的预测。

通过对时间序列的分析和预测,我们可大体观察出医院总体门诊量的变化趋势,及短期内的具体预测值,为医疗卫生部门进行优化资源配置提供初步的统计依据。文中对各县市区的分析,揭示了部分地区在一年中的某些时段很少来院就诊,且不同地区来诊人数高峰和低谷的时间先后略有不同。根据预测结果,结合医疗部门具体情况,临时派遣部分医护力量在可能的就诊低谷时段,按照来诊人数低谷的先后,对特定地区的社区医疗部门进行支援,增加其诊疗力量,使患者可以就近及时就医,可能会在一定程度上有针对性地解决部分地区群众难以及时就医的困难。另外,结合地图,通过对时间序列的分析和预测还可以发现其

他一些鲜有关注的问题。因为温州地形复杂,以山居多,不少患者发病的初期往往由于各种因素不能及时诊治。例如泰顺、文成和洞头 3 县离温州市并不是很远,但是要么群山相隔,要么隔海而望,交通不便,所以来温州就诊的人数远远低于其他各县,提示该 2 县的社区医疗力量是增援的主要对象。瑞安与温州比邻,但来院门诊量却远远不能与远在最南端的苍南相比,通过预测图可以看出,近年来两者差距似有继续增大的趋势。这与瑞安当地医疗力量的加强不无关系,却也提醒我院需要从改善就医环境氛围、进一步提高医疗质量、扩大技术优势等方面继续努力。

### 参 考 文 献

- [1] 王春平,王志峰,单杰,等. 随机时间序列分析法在传染病预测中的应用[J]. 中国医院统计,2006,13(9):229-232.
- [2] Box GEP, Jenkins GM, Reinsel GC. Time series analysis forecasting and control[M]. 北京:中国统计出版社,1997:377-424.
- [3] 周忠彬,吕红梅,邹鄂.  $ARIMA$  干预模型在医院门诊量预测中的应用[J]. 中国医院统计,2008,15(6):110-112.
- [4] 黄彦,易东,田考聪.  $ARIMA$  预测模型的 SAS 程序实现及其应用[J]. 激光杂志,2007,28(1):96.
- [5] 周展凤,杨小敏,刘小冬. 某脑科医院近 20 年神经精神疾病门诊量动态分析[J]. 中国慢性病预防与控制,2005,13(5):211-213.
- [6] 丁守奎,康家琦,王洁贞.  $ARIMA$  模型在发病率预测中的应用[J]. 中国医院统计,2003,10(1):23-26.
- [7] 潘华志,耿则勋,夏治国,等. SAS/ETS 在变形监测数据处理中的应用[J]. 测绘工程,2007,16(1):59-65.
- [8] 高惠旋. SAS 系统·SAS/ETS 软件使用手册[M]. 北京:中国统计出版社,1998:65-115.
- [9] 梁桂玲,刘颜,邓涸沐.  $ARIMA$  模型应用于月门诊量预测[J]. 中国医院统计,2006,13(1):24-26.
- [10] 方永兵. 基于时间序列分析的地表沉降数据分析处理[J]. 山东理工大学学报:自然科学版,2007,21(5):32-34.
- [11] 徐静.  $ARMA$  模型及其应用[J]. 立信会计高等专科学校学报,2001,15(3):21-24.

(收稿日期:2008-11-27;修回日期:2009-05-14)

(上接第 201 页)

- [3] 龚晓燕. Ridit 分析法在医疗质量管理中的应用[J]. 中国医院统计,2003,10(4):226.
- [4] 欧新元,张红霞. 应用 Ridit 法评价医疗疗效[J]. 中国医院统计,2004,11(1):52.
- [5] 罗明奎,蔡昌启,雷玉洁,等. 等级资料 Ridit 分析及正确使用[J]. 中国卫生统计,2003,20(4):252-254.
- [6] 黄建平,姜新莉,苏静. 近 10 年我院的疗效评价及影响因素分析[J]. 中国医院统计,2006,13(2):150-152.
- [7] 曹阳,张罗漫. 医疗治疗的综合评价方法[J]. 中国卫生统计,1998,15(5):34-43.

- [8] 彭忠生. 我院医疗质量分析及思考[J]. 中国医院统计,1994,1(2):92-93.
- [9] 刘朝辉,胡伟灿. 运用 Ridit 法综合评价医院医疗质量[J]. 中国卫生统计,2005,12(6):432.
- [10] 徐雁,耿兴斌. Ridit 分析在临床治疗效果评价中的应用[J]. 数理医学杂志,1995,8(3):259-260.
- [11] 朱本耀. 运用 Ridit 分析法评价及分析医院医疗质量[J]. 中国卫生统计,2006,23(3):280-281.
- [12] 章扬熙. 医学科研设计与卫生统计[M]. 郑州:郑州大学出版社,2005:137-141.

(收稿日期:2008-12-15;修回日期:2009-03-10)

作者: 徐永强, 周静, 陈学军, 潘安

作者单位: 徐永强, 陈学军, 潘安(浙江省温州市, 温州医学院附属第二医院, 325000), 周静(温州医学院附属第一医院)

刊名: 中国医院统计

英文刊名: CHINESE JOURNAL OF HOSPITAL STATISTICS

年, 卷(期): 2009, 16(3)

被引用次数: 1次

## 参考文献(11条)

1. 王春平, 王志锋, 单杰, 等. 随机时间序列分析法在传染病预测中的应用[J]. 中国医院统计, 2006, 13(9): 229-232.
2. Box GEP, Jenkins GM, Reinsel GC. Time series analysis forecasting and control[M]. 北京: 中国统计出版社, 1997: 377-424.
3. 周忠彬, 吕红梅, 邹郢. ARIMA干预模型在医院门诊量预测中的应用[J]. 中国医院统计, 2008, 15(6): 110-112.
4. 黄彦, 易东, 田考聪. ARIMA预测模型的SAS程序实现及其应用[J]. 激光杂志, 2007, 28(1): 96.
5. 周展凤, 杨小敏, 刘小冬. 某脑科医院近20年神经精神疾病门诊量动态分析[J]. 中国慢性病预防与控制, 2005, 13(5): 211-213.
6. 丁守蛮, 康家琦, 王洁贞. ARIMA模型在发病率预测中的应用[J]. 中国医院统计, 2003, 10(1): 23-26.
7. 潘华志, 耿则勋, 夏治国, 等. SAS/ETS在变形监测数据处理中的应用[J]. 测绘工程, 2007, 16(1): 59-65.
8. 高惠旋. SAS系统·SAS/ETS软件使用手册[M]. 北京: 中国统计出版社, 1998: 65-115.
9. 梁桂玲, 刘颜, 邓泗沐. ARIMA模型应用于月门诊量预测[J]. 中国医院统计, 2006, 13(1): 24-26.
10. 方永兵. 基于时间序列分析的地表沉降数据分析处理[J]. 山东理工大学学报: 自然科学版, 2007, 21(5): 32-34.
11. 徐静. ARMA模型及其应用[J]. 立信会计高等专科学校学报, 2001, 15(3): 21-24.

## 相似文献(10条)

1. 期刊论文 华来庆, 申广荣, 熊林平, 孟虹, 赵胜荣, 胡亚萍, HUA Lai-qing, SHEN Guang-rong, XIONG Lin-ping, MENG Hong, ZHAO Sheng-rong, HU Ya-ping ARIMA模型在黄瓜霜霉病疾病指数时间序列建模中的应用研究 - 第二军医大学学报2006, 27(7)

目的: 探索黄瓜霜霉病疾病指数时间序列预测方法. 方法: 采用黄瓜霜霉病病情指数时间序列进行研究, 通过模型识别、残差方差比较、参数估计及其检验、观察参数之间相关系数矩阵、白噪声检验、模型的拟合度分析等过程. 结果: 建立了ARIMA(2, 2, 0)模型:  $(1+0.487B+0.554B^2)(1-B)^2Y_t=at$ . ARIMA(2, 2, 0)模型的预测值误差平方和SSE=0.001822, 根均方误差RMSE=0.008537, 且验证数据的预测值与原值吻合较好. ARIMA(2, 2, 0)模型为本研究获得的预测效果较好的一维时间序列模型, 适合于黄瓜霜霉病的中期、后期预测. 结论: 通过残差方差定阶法缩小模型选择范围, 再结合模型的参数估计、相关系数矩阵、白噪声检验和拟合优度检验最后确定模型的思路, 有利于快速准确地找到合适的模型.

2. 学位论文 景立伟 传递函数模型及其在医院业务收入分析中的应用 2006

时间序列是按照时间顺序取得的一系列观察值, 时间序列数据的本质特征就是相邻观察值之间的依赖性. 所以在对时间序列数据进行分析时, 一般的回归模型难以体现变量自身前后及应变量与自变量过去的依赖关系. ARIMA模型是对变量自身前后依赖性进行分析的技术, 它描述了变量自身当前与过去的统计依赖关系, 在显示变量的动态系统(dynamical system)演变规律方面有着较为丰富的结构. 但是由于ARIMA模型没有考虑其它变量对应变量的作用, 所以也无法表达系统中变量间相互影响的关系. 传递函数模型可在一定程度上弥补这两种模型的欠缺, 并且模型构造灵活, 可以用较少的参数建立阶数较高的模型.

当ARIMA模型包括其它时间序列作为输入变量时, 被称为传递函数模型(transfer function model)、多变量时间序列模型(multivariate timeseries model)、ARIMAX模型或Box-Tiao模型. 传递函数模型是ARIMA模型的自然推广, Pankratz统称这种包含其它时间序列作为输入变量的ARIMA模型为动态回归. 传递函数模型由George E. P. Box和Gwilym M. Jenkins于七十年代在ARIMA模型的基础上提出, Box-Tiao也曾经对其作过详细讨论. 传递函数模型在形式上可以看作是ARIMA模型与回归模型的结合, 这种模型能清晰的刻画变量间动态影响的过程. 作为多元回归模型的推广, 传递函数模型建模的假定较少, 容易得到满足, 在现实的系统中有着广泛的应用.

使用传递函数模型对时间序列进行分析, 最大的意义在于提高预测精度. 在利用一个时间序列在某时刻的有效观察值去预报在某个未来时刻该序列的值时, 若能确定出输入时间序列及输入时间序列和输出时间序列之间的动态关系, 则可以输入时间序列和输出时间序列的过去值对输出时间序列作出预报. 由于传递函数模型是在利用ARIMA模型理论上, 同时将输入序列的变化对输出序列预测值的影响充分的加以考虑, 所以可以使预报误差大大降低.

本文中, 我们将利用传递函数模型对医院业务收入的影响因素进行分析, 同时还要对医院业务收入进行预测, 并将传递函数模型的预测结果与实际值和ARIMA模型的预测结果进行比较.

3. 期刊论文 张奕韬, ZHANG Yi-tao 基于ARIMA模型的外汇汇率时间序列预测研究 - 华东交通大学学报2009, 26(5)

利用数据挖掘技术分析外汇汇率时间序列, 从时间序列中获得正确的、隐含的、潜在的信息对于金融领域研究具有重要的现实意义. 通过数据挖掘中的ARIMA模型, 以某银行的外汇汇率时间序列为研究对象, 采用差分方法和建模规则, 对外汇的卖出价进行了建模与预测. 通过与逐步自回归预测模型相比较

, ARIMA模型对外汇率率时间序列数据具有很强的预测能力.

#### 4. 期刊论文 [刘爱萍. 郜文明](#) [ARIMA模型在省级全社会固定资产投资预测中的应用](#) -[河南金融管理干部学院学报](#)

2008, 26 (4)

ARIMA模型较好地解决了非平稳时间序列的建模问题, 并且在时间序列的短期预测方面有很好的表现. 借助于EViews等统计软件, 可以方便地将ARIMA模型用于时间序列问题的研究和预测. 利用河南省1989至2006年的全社会固定资产投资总额数据, 运用计量经济学软件EViews, 基于时间序列分析方法建立相应的ARIMA模型, 进行预测分析, 为各级政府和企事业单位相关的管理决策, 提供数量化的参考信息.

#### 5. 学位论文 [王卫琪](#) [ARIMA模型及其在医疗资源消费预报中的应用](#) 2008

时间序列是按照时间顺序取得的一系列观察值, 时间序列数据的本质特征就是相邻观察值之间的依赖性. 所以在对时间序列数据进行分析时, 一般的回归模型难以体现变量自身前后及应变量与自变量过去的依赖关系. 自回归求和滑动平均 (ARIMA) 模型是对变量自身前后依赖性进行分析的技术, 它描述了变量自身当前与过去的统计依赖关系, 在显示变量的动态系统 (dynamical system) 演变规律方面有着较为丰富的结构.

ARIMA模型由George E. P. Box和Gwilym M. Jenkins首次系统提出, 其建模的假定较少, 容易得到满足, 在现实的系统中有着广泛的应用.

使用ARIMA模型对时间序列进行分析, 最大的意义在于预报. 本文中, 将利用ARIMA模型对医疗资源消费进行预报, 并将预报结果与实际值和普通回归模型的预报结果进行比较.

#### 6. 会议论文 [李志刚. 荣长军. 何志文. 易世华](#) [磁暴地磁水平分量ARIMA模型建模](#) 2007

本文考虑到非线性分析对时间序列平稳性的要求, 以及磁暴地磁水平分量时间序列明显的时间趋势, 探讨了用ARIMA模型对磁暴地磁水平分量时间序列进行建模和一步预测与两步预测, 为进一步通过模型预测磁暴的演化提供了一种思路.

#### 7. 期刊论文 [饶文章. 徐水太. 熊灵燕. Rao Yunzhang. Xu Shuitai. Xiong Lingyan](#) [基于ARIMA模型的矿区重金属污染时间序列预测](#) -[金属矿山](#)2010, "" (6)

矿区重金属污染具有时间序列的特征, 因此可以采用时间序列ARIMA模型对重金属污染进行预测. 对南方某铜硫矿, 在1995年1月-2008年6月重金属月监测数据的基础上, 运用ARIMA模型建立了矿区尾矿库废水总排放口Zn浓度的预测模型, 结果表明, ARIMA (1, 1, 2) 模型能较好地拟合2008年1月-2008年6月重金属污染变化规律, 经实际计算结果验证所建模型, 误差在5%左右, 经检验其精度满足要求. 预测结果显示, 该矿区未来重金属Zn仍然处在污染状态.

#### 8. 学位论文 [漆莉](#) [重庆市2004~2006年流感流行特征分析及ARIMA模型在传染病发病预测中的应用](#) 2007

目的: 本研究一方面为了掌握我市近几年的流感流行状况, 为制定流感防控策略提供科学依据; 另一方面, 探讨ARIMA模型在传染病发病预测方面的应用, 以流感、痢疾为代表, 对其进行分析、评价, 寻求该方法的最佳适用范围与适用条件.

方法: (1) 通过哨点医院监测和暴发疫情现场调查处理, 收集我市2004年4月至2006年3月部分地区的流感样病例 (ILI) 发病资料和全市ILI/流感的暴发疫情资料, 并采集ILI标本进行实验室检测, 对所有流行病学资料和病原学检测结果进行描述、分析. (2) 通过全市各医疗机构收集我市2002年1月至2006年12月的流感和痢疾发病资料, 用SPSS统计软件对流感和痢疾的月发病数建立ARIMA模型, 用Q统计量检验法对模型的适应性进行检验, 并用组外回代的方法检验其预测精度.

结果: (1) 2004年4月至2006年3月, 我市流感监测哨点医院ILI就诊百分比 (ILI%) 出现2004年4月和10月、2005年4月和8月及11月、2006年4月和7月几个高峰, 流感病毒分离高峰与ILI%高峰基本一致, 稍有滞后. 2005年、2006年ILI%夏季高峰高于春季高峰, 但病原学检测没有显示在夏季分离到更多的流感病毒. (2) 我市流感季节类型明显, 2004年8月、2005年2月和9月、2006年2月和7月出现优势毒株的变化, 随后1、2个月, 出现ILI%高峰.

(3) ILI/流感暴发疫情全年均有, 春季和秋季略有增加; 疫情发生地以中小学校为主 (占91.20%), 尤其是偏远地区更易发生; 疫情类型以A型为主, H1N1型和H3N2型引起的暴发交替出现, 且与哨点医院收集的ILI标本毒株型别变化一致. (4) 我市流感月发病数时间序列的自相关分析图显示: 我市流感没有固定的季节性, 用ARIMA模型法建立预测模型, 只能进行3个月以内的预测, 预测区间范围较大, 缺乏实际意义. 痢疾月发病数时间序列的自相关分析图显示: 我市痢疾具有明显的季节性, 建立季节复合模型后预测效果比较理想, 可进行至少1年的预测; 用48个数据建模, 6个月预测效果的拟合度 $R^2 > 0.9733$ , 12个月预测效果的 $R^2 > 0.8574$ ; 36个数据和42个数据也能取得比较理想的预测效果, 但 $R^2 > 0.48$ 个数据相同时间预测效果的 $R^2 > 0.48$ 要低些.

结论: (1) 我市可能存在春季和秋季两个流感流行高峰, 应加强春秋两季的流感疫苗接种, 推荐流感疫苗最佳接种时间为每年2、3月或9、10月; 加强我市夏季呼吸道疾病的鉴别诊断, 尽量避免误诊; 每年2月、8月前后加强我市流感病毒的分离工作, 及早发现毒株变异, 当优势毒株出现变化时应警惕ILI流行高峰的出现; 我市流感暴发疫情的防控应以农村中小小学为重点, 哨点医院监测到A型毒株型别出现变化时, 应警惕新型毒株引起的暴发疫情高峰的出现. (2) ARIMA模型法可以对疾病的季节性进行分析, 对不同疾病的预测效果存在较大差异, 当疾病发病比较稳定时预测效果可能更好; 在用ARIMA模型法进行预测时, 应根据资料选用不同的模型类型, 所需建模数据可以少于通常要求的50个, 但建模数据越多预测效果越好, 随着预测时间的推移, 预测效果有所降低, 因此运用该方法进行预测时应不断增加新的观测数据, 重新建立预测模型.

#### 9. 期刊论文 [刘付程. 刘吉堂. 苏伟. 郭衍游. LIU Fu-cheng. LIU Ji-tang. SU Wei. GUO Yan-you](#) [基于季节ARIMA模型的](#)

#### [近海表层水温时间序列模拟与预测](#) -[淮海工学院学报 \(自然科学版\)](#) 2009, 18 (4)

基于1996-2007年逐月时间序列数据, 采用季节ARIMA模型对连云港近海表层水温时间序列进行模拟, 并依据残差不相关和简洁性原则确定模型的结构, 建立最优预测模型ARIMA (1, 0, 1) (0, 1, 0) 12. 运用该模型对2008年逐月表层水温进行预测, 预测值与实际值的变动趋势基本一致, 且平均相对误差仅为3.5%. 在此基础上对2009年连云港近海逐月表层水温进行预报, 预报结果符合该海域表层水温的逐月变化趋势, 表明模型用于近海表层水温预报是可行的.

#### 10. 期刊论文 [孙彩云. 杨晓静. SUN Caiyun. YANG Xiaojing](#) [乘积ARIMA模型的建立及应用](#) -[华北科技学院学报](#)

2008, 5 (2)

ARIMA模型作为统计预测中的一种重要模型, 被广泛运用于各个领域. 本文在非季节性 & 季节性ARIMA模型的基础上介绍了乘积ARIMA模型的一般表述, 提供了建立这类模型及使用其进行预报的一般过程. 并以某超市月顾客数量数据进行了实例分析. 分析结果表明, 当一时间序列不仅含有季节性成分, 还混有非季节性成分时, 用乘积ARIMA模型的预测精度较高. 但该模型只适用于短期预测.

本文链接: [http://d.g.wanfangdata.com.cn/Periodical\\_zgyytj200903003.aspx](http://d.g.wanfangdata.com.cn/Periodical_zgyytj200903003.aspx)

授权使用: 南昌大学图书馆 (wfncdxtsg), 授权号: 2b2a99c9-94ad-423b-8680-9e2b010b29d4

下载时间: 2010年11月11日