

桂林市汽车销售量的时间序列预测模型*

The Forecasting Model of Time Series for the Auto Sales in Guilin

梁鑫, 庞丽, 彭冬梅

LIANG Xin, PAN G Li, PENG Dong-mei

(广西师范大学数学科学学院, 广西桂林 541004)

(College of Mathematics, Guangxi Normal University, Guilin, Guangxi, 541004, China)

摘要: 选取 1999 年 1 月到 2007 年 3 月桂林市各季度的汽车销售量原始数据, 在 SPSS 系统下, 运用滑动求和自回归 (ARIMA) 方法及非参数方法建立桂林市汽车销售量时间序列模型 $ARIMA(p, d, q)$, 从模型识别、参数估计、适应性检验和实际拟合 4 个方面来确定模型的参数 (p, d, q) , 并对模型的预测效果进行检验。结果表明, $ARIMA(0, 2, 2)$ 模型能够较好地包含桂林市汽车销售量的发展趋势, 该模型对 2007 年二季度至 2008 年二季度汽车销售量的预测值与实际值的误差小, 相对误差可以控制在 3% 以内。

关键词: 预测模型 时间序列 汽车销售量 非参数统计

中图分类号: O212 文献标识码: A 文章编号: 1005-9164(2008)04-0386-03

Abstract Under the SPSS system, this paper carries out the analysis by using the data of the auto sales in Guilin from January 1999 to March 2007, to confirm the time series model that mostly fit in with the development rule of the auto sales in Guilin from four aspects: model identification, parameter estimation, diagnostic checking and actual fitting. The research presents the auto sales from the second quarter in 2007 to the second quarter in 2008, compared with the actual value, the comparative error of the forecast value of the $ARIMA(0, 2, 2)$ model is less than the 3%. Therefore, the $ARIMA(0, 2, 2)$ model fits the development rules of the auto sales in Guilin.

Key words forecasting model, time series auto, sales, non-parameter statistics

近年来, 国内外的汽车市场无论是产量还是销售量都保持着较快的增长速度, 国内汽车行业的市场占有率在整体上也呈现出上升趋势。为了能在激烈的竞争中占有一席之地, 各汽车厂商都将主要精力集中于未来的变化上; 而汽车销售量的增多, 必然引起相应行业部门政策的改变, 如汽车生产商有效成本的投入, 国家对公路以及城市道路的合理建设, 交通设施的改进和完善, 环境部门对城市环境的检测等等。因此, 建立一个能够精确预测汽车销售量的模型已经成为一种必然趋势。很多学者对汽车销售

量进行了研究, 如文献 [1] 利用多元线性回归模型对汽车销售量进行分析, 文献 [2] 利用 ARM A 模型与 RBF 神经网络相结合, 建立了车销售量的预测模型。但是这些研究没有在消除销售量之间相关性的基础上进行建模, 容易造成模型的预测效果不佳。本文基于 SPSS 系统^[3], 运用滑动求和自回归 (ARIMA) 方法及非参数方法^[4]对桂林汽车销售量的发展趋势进行研究, 以为建立汽车销售量的预测模型提供参考, 同时也为相关部门的决策提供参考依据。

1 数据来源及数据分析

以 1999 年 1 月至 2007 年 3 月期间各季度的桂林市汽车季度销售量为原始数据 (数据来源于桂林市交警大队), 对各季度的汽车销售原始数据做时间序列

收稿日期: 2008-03-18

修回日期: 2008-09-26

作者简介: 梁鑫 (1978-), 男, 讲师, 主要从事概率统计研究工作。

* 广西自然科学基金项目 (0728091), 广西师范大学青年科学基金项目及学生学术科技创新基金项目 (2007) 资助。

图(图 1)。从图 1 可以看到,随着我国改革开放和人民生活水平的不断提高,桂林市在过去 8 年的汽车销售量虽然有波动,但在总体上呈现出不断增长的趋势。因此该汽车销售量序列为一非平稳时间序列。

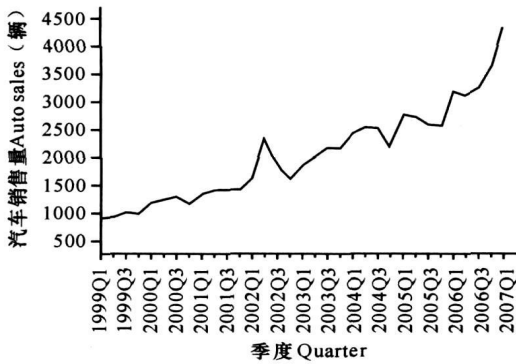


图 1 1999年至 2007年桂林市季度汽车销售量

Fig. 1 The auto sales of quarter in Guilin from 1999 to 2007

2 汽车销售量时间序列预测模型的建立及检验

根据 ARIMA 模型的要求,从模型识别、参数估计、适应性检验和实际拟合 4 个方面来确定最符合汽车销售量发展趋势的模型。

2.1 ARIMA(p, d, q)模型

设 $X = (x_1, x_2, \dots, x_t, \dots, x_n)$ 为一个在等间隔时间 $1, 2, \dots, t, \dots, n$ 采集得到的时间序列, t 为第 t 个时间点, x_t 为 t 时刻汽车销售量的时间序列值。

ARIMA(p, d, q)模型(即滑动求和自回归模型)的一般形式^[3]为:

$$\Phi(B)(\nabla^d x_t) = \Theta(B)e_t,$$

或者

$$\Phi(B)(1-B)^d x_t = \Theta(B)e_t,$$

其中 B 为后移算子, $Bx_t = x_{t-1}$, ∇ 为差分符号, $\nabla = 1 - B$, d 为差分阶数, $\Phi(B) = 1 - h_1 B - h_2 B^2 - \dots - h_p B^p$ 为自回归算子, p 为自回归阶数, h_1, h_2, \dots, h_p 为自回归部分参数; $\Theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$ 为移动平均算子, q 为移动平均阶数, $\theta_1, \theta_2, \dots, \theta_q$ 为移动平均部分参数, e_t 为白噪声序列。

2.2 模型的识别与定阶

由于桂林市的汽车季度销售量为非平稳时间序列,需要先对该时间序列进行差分平稳化处理以消除线性趋势。因此,我们对 ARIMA(p, d, q)模型中的 3 个参数 p, d, q 的各种可能取值在 SPSS 系统下进行重复拟合(明显不合适或提出不适合警告的取值形式予以删除),并采用最佳准则函数定价法中的最小信息准则(AIC 准则和 BIC 准则)^[3]来判定模型

的最佳阶数

选取不同的模型阶数对汽车销售量的时间序列值 $\{x_t\}$ 进行重复拟合,并计算相应模型的 AIC 值和 BIC 值。经比较,在收敛标准的最大值为 10 参数变化为 0.001%、平方和变化为 0.001% 的情况下,取 $(p, d, q) = (0, 2, 2)$ 时, AIC 值(444.871)和 BIC 值(BIC= 449.173)达到最小。因此,初步选定汽车销售量时间序列预测模型为 ARIMA(0, 2, 2)。

2.3 参数估计

对 ARIMA(0, 2, 2)模型的参数进行估计的结果如表 1 所示。表的参数估计值为预测模型的进一步建立提供必需的模型参数。

表 1 ARIMA(0, 2, 2)模型参数估计

Table 1 Parameter estimation in the ARIMA(0, 2, 2) model

参数 Parameter	MA1	MA2	CONSTANT
估计值 Estimate Value	0.946	-0.013	7.166
标准差 Std. Error	0.279	0.21	6.536

2.4 模型适应性检验

利用 SPSS 系统的卡方检验及单样本 K-S 检验的方法对 ARIMA(0, 2, 2)模型的序列原始数据与拟合数据的误差(残差)序列进行模型适应性检验的结果如表 2 所示。

表 2 残差序列的检验

Table 2 The test of the residual sequence

模型 Model	M	Q 值 Q value	$i_{0.05}^2(M)$	K-SZ 值 K-SZ value	Sig.
ARIMA(0, 2, 2)	5	2.3614	11.143	0.731	0.66

表 2 结果显示, $Q = 2.3614 < i_{0.05}^2(5)$, 说明残差序列是独立的; 而 $K-SZ$ 值为 $0.731 > Sig. = 0.66$, 说明残差序列是正态的。因此, ARIMA(0, 2, 2)模型的残差序列是白噪声序列^[5], 即 ARIMA(0, 2, 2)模型能够较好地包含桂林市汽车销售量的发展趋势。

2.5 实际拟合

由图 2 可以看出 ARIMA(0, 2, 2)模型的计算值和桂林市汽车销售量的实际值拟合效果很好,说明用 ARIMA(0, 2, 2)模型可以对桂林市汽车销售量的未来趋势进行较好地预测。

2.6 预测检验

运用 ARIMA(0, 2, 2)模型预测 2007 年季度至 2008 年季度的汽车销售量,所得的值与实际值的比较结果见表 3。

由表 3 可知,汽车销售量的实际值与根据 ARIMA(0, 2, 2)模型得到的预测值之间的误差很小,相对误差均控制在 3% 以内。因此,可以利用该模

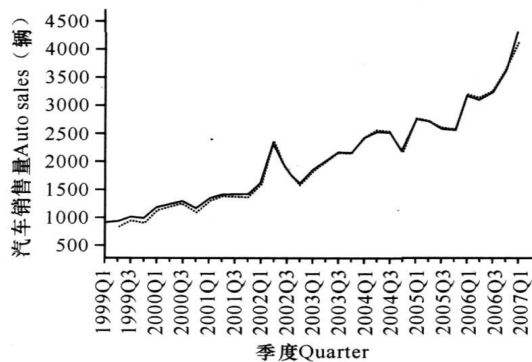


图 2 模型值与实际值

Fig. 2 The value of model and actual

— : 实际值; ····· : 模型值

— : Facting value; ····· : Modle value.

表 3 汽车销售量的实际值与预测值的比较

Table 3 The comparison about facting value and fitting value of the auto sales

季度 Quarter	实际值 (辆) Facting value	预测值 (辆) Fitting value	相对误差 Comparative error
2007 Q2	4450.38	4432.75	-0.004
2007 Q3	4691.81	4673.23	-0.004
2007 Q4	4940.41	4924.33	-0.003
2008 Q1	5196.18	5085.42	-0.021
2008 Q2	5729.21	5639.55	-0.016

3 结束语

本文建立的桂林市汽车销售量时间序列模型 ARIMA(0, 2, 2)能够较好的反映桂林市汽车销售市场中汽车季度销售量的发展规律。该模型的建立,除了可以更好地掌握汽车市场的供需关系外,还可以为有关部门预测并掌握汽车销售数量,以便及时做出相应的决策。本文基于 SPSS系统建立时间序列模型的方法,还可以对其他现实生活中具有明显时间特征和趋势的经济数据进行建模,研究其内在规律,更好地把握其未来的发展趋势。

参考文献:

- [1] 刘严萍,何全秀. SPSS预测模型在汽车市场分析中的应用 [J].河南理工大学学报, 2006, 7(3): 208-211.
- [2] 李响,宗群,童玲. 汽车销售混合预测方法研究 [J].天津大学学报, 2006, 8(3): 175-178.
- [3] 王振龙. 时间序列分析 [M].北京: 中国统计出版社, 2000.
- [4] 杨善朝,张军舰. SPSS统计软件应用基础 [M].桂林: 广西师范大学出版社, 2001.
- [5] 梁鑫,韦程东,谢佳利. 乘积季节模型在商品房市场中的研究及其应用 [J].广西师范学院学报: 自然科学版, 2006, 23(2): 8-12.

(责任编辑: 韦廷宗)

以色列科学家首次实现用气体存储图像

以色列的科学家最近成功地在原子蒸气上实现了图像存储,尽管存储的时间只有短短的 30μs,但这是人类首次成功地利用气体充当存储媒介。

以色列科学家在一束光脉冲中将一幅图片减速到每秒 8000m 的群速度。这一群速度使得图片能够存储于原子状态的气体之中长达数微秒之久。他们的实验过程是,首先将两束光线射向一个 5cm 长的气体试管上,该试管中含有 52°C 的铷气以及用于隔离作用的氦气。第一束光脉冲(即包含图像的光脉冲)只要有一半离开气体试管,立即关闭第二束光柱,这样图像的余下一半就存储于气体之中。在这里,图像是以原子的量子态编码存储的。30μs 后,再次打开第二束光柱,图像也随之恢复。

以色列科学家解释,在存储期间,整个实验系统中没有任何光场。光线携带的所有信息被转换为气体中所有原子的量子态。如果能够很容易地检测出原子的量子态的稳定水平,人们将能够看到一幅鲜明的“图像”存在于该气体之中。由于气体原子的扩散性,恢复的图像看起来可能模糊不清,而且信噪比大大降低。为了改进图像的清晰度,他们采用了一种新的技术方法,将由于原子运动造成的图像模糊化问题减少到最低程度。该技术与移相微影技术有相似之处,即是将这些图像特征的相位转移 180°,这样那些扩散到图像线条之间区域的相反相位的原子的振幅将被取消,也就不可能有光线射出而模糊图像的线条。

气体存储图像的技术,将在图像处理及相关领域有着广泛的应用。科学家们预测,利用这种方法将可能存储更为清晰的图像,包括即时图像或电影等数据。当然,这种光存储技术并不仅仅局限应用于图像存储方面,在未来的量子信息发展中,光存储技术将同样发挥着重要的作用。

(据科学网)