
ARIMA 预测模型在医院的早产儿情况统计分析中研究

莫佳琪¹, 文水珍², 刘志辉¹, 江岚¹, 杨滨波¹, 朱淳¹

(1 湖南省妇幼保健院, 湖南长沙, 410008) (2 湖南中医药高等专科学校附属第一医院, 湖南株洲, 412000)

摘要: 目的: 探索 ARIMA 预测模型在医院早产儿情况统计分析中的应用, 为医院管理、政策制订提供科学依据。方法: 运用 ARIMA 模型预测医院早产儿情况的发展趋势。结果: 构建 ARIMA (2,1,1) (0,1,1) 模型, 预测了未来两年医院早产儿情况有较大的上升趋势, 且呈现较强的周期性。结论: ARIMA (2,1,1) (0,1,1) 模型可以较好的拟合 2009—2013 年湖南省妇幼保健院医院早产儿情况的时间变化趋势, 并预测了未来两年的发展趋势。ARIMA 模型在医院早产儿情况的预测研究精度较高, 是一种理想的研究方法。

关键词: 医院; 早产儿; ARIMA 模型; 预测

Research of the premature statistical analysis through the Arima prediction modeling

Mo Jiaqi¹, Wen Shuizhen², Liu Zhijun¹, Jianglan¹, Yang binbo¹, Zhuchun¹

(1 MCH hospital of Hunan province, Changsha Hunan, 410008) (2 The First Affiliated Hospital of Hunan Traditional Chinese Medical College, Zhuzhou Hunan, 412000)

Abstract: Objective: To explore the application of Arima prediction modeling in premature statistical analysis so as to provide scientific evidence for hospital management and policy setting. Methods: Predicting the trend of premature by using ARIMA prediction modeling. Results: Predicted the rising trend of premature and the rising periodic feature by building ARIMA (2,1,1) (0,1,1). Conclusion: ARIMA (2,1,1) (0,1,1) modeling fitted the time variation trend of premature well during 2009 to 2013 in MCH hospital of Hunan province, and predicted the development trend of the coming year. It is proved that ARIMA modeling can provide satisfied predictions of the premature is to be thought as the ideal research method.

Key words: Hospital ; premature; ARIMA modeling; Prediction

早产是指在满 28 孕周至 37 孕周之间 (196~258 天) 的分娩。此时分娩出的新生儿称早产儿, 为各器官未成熟的新生儿。早产儿即使存活, 亦多有神经智力发育缺陷^[1]。近年来由于早产儿治疗学及监护手段的进步, 其生存率明显提高, 伤残率下降。国外学者建议将早产定义事件上限提前到妊娠 20 周。因此, 防止早产是降低围生儿死亡率和提高新生儿素质的主要措施之一。本文根据统计预测的原理, 运用 ARIMA 模型预测医院的早产儿情况, 一方面预测分析医院的早产儿情况发展趋势, 为医院预防早产儿情况, 政策制定提供科学依据; 另一方面探索 ARIMA 模型在医院的早产儿情况综合预测中的科学性。

1 资料来源、原理与方法

1.1 资料来源

作者简介: 莫佳琪, (1987.11), 湖南益阳人, 女, 护师, 学士, 研究方向: 临床护理。
qinaideliuyi@163.com, 15173305440.

作者联系地址: 长沙市湘春路 53 号 湖南省妇幼保健院产房, 湖南长沙, 410008

本文以湖南省妇幼保健院 2005 年 1 月—2013 年 12 月医院的早产儿情况统计数据为研究对象。（见表 1）

1.2 研究原理

ARIMA 模型全称综合自回归移动平均（Auto Regressive Intergrated Moving Average）模型，简称为 ARIMA（p,d,q）模型，是由 Box 和 Jenkins 于 20 世纪 70 年代初提出的著名时间序列预测模型，又称为 Box-Jenkins 模型。其中 AR 是自回归，p 为自回归阶数；MA 为移动平均，q 为移动平均阶数；d 为时间序列成为平稳时间序列时所做的…差分次数^[2]。ARIMA（p,d,q）模型的实质就是差分运算与 ARIMA（p,d,q）模型的组合，即 ARIMA（p,d,q）模型经 d 次差分后，便为 ARIMA（p,d,q）^[2]。ARIMA 模型的建模过程包括 3 个阶段：模型的识别阶段、模型的估计和诊断阶段、模型的预测应用阶段。

1.3 研究方法

本文在 SPSS19.0 软件上完成分析建模过程，并根据湖南省妇幼保健院 2005 年 1 月—2013 年 12 月医院的早产儿情况统计数据，运用 ARIMA 模型预测医院未来两年的早产儿情况发展趋势。

2 结果

2.1 在医院的早产儿情况统计中 ARIMA 模型识别

2.1.1 序列的平稳性分析

序列的平稳化处理目的是为满足一些模型对数据的基本要求而实施的一种数据处理方法，使处理后的序列成为平稳序列。

表 1 2005 年—2013 年湖南省妇幼保健院的早产儿情况统计表

年份	1 月	2 月	3 月	4 月	5 月	6 月	7 月	8 月	9 月	10 月	11 月	12 月
2005	40	22	21	29	17	33	23	24	17	29	24	30
2006	38	25	32	30	37	26	33	40	38	44	38	44
2007	46	53	64	53	57	47	66	56	73	55	66	72
2008	61	58	57	41	54	57	73	85	51	71	70	63
2009	105	89	48	50	65	75	67	73	76	67	68	93
2010	79	81	87	69	79	87	94	104	109	100	107	130
2011	128	102	112	91	84	107	97	117	105	103	101	132
2012	114	123	121	98	126	104	107	92	127	124	132	166
2013	130	145	149	111	135	123	127	108	142	139	148	168

从湖南省妇幼保健院 2005 年—2013 年每月的早产儿情况统计序列图 1 分析可知：原始数据的序列很不稳定，具有明显的向上趋势，而且原始数据在 1 月、3 月、9 月和 12 月呈现高峰，具有一定的季节性和周期性。为了消除原始数据的不平稳性，采用对湖南省妇幼保健院 2005 年—2013 年每月早产儿情况统计取自然对数形式，经过一阶差分和一阶季节差分后，使序列平稳（如图 2）。

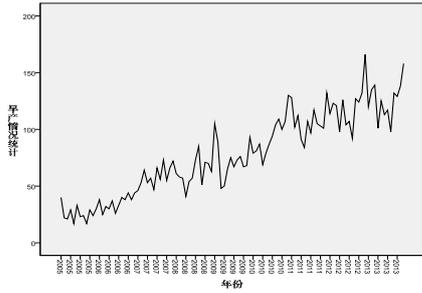


图1 湖南省妇幼保健院早产情况序列图

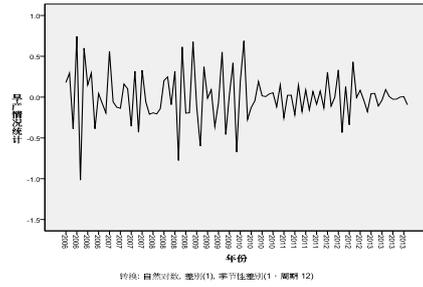


图2 自然对数、一阶差分、一阶季节差分的序列图

2.1.2 模型的识别

根据原始数据的平稳处理过程，可初步确定 ARIMA 模型的的差分 $d=1$ ，季节差分 $D=1$ 。从湖南省妇幼保健院每月早产儿情况统计序列图 3 可知：（1）自相关函数图的 1 阶显著不为 0，在 1 阶以后函数值明显趋于 0，呈现拖尾性，因此可将 q 取为 1，而第 12 阶的函数值显著不为 0，可将 Q 取为 1。（2）偏自相关图的前二阶均显著不为 0，之后趋于 0，呈现拖尾性，因此可将 p 取 2。（3）因此可以初步得到模型的形式 ARIMA (2,1,1) (0,1,1) 和 ARIMA (2,1,1) (1,1,1)。

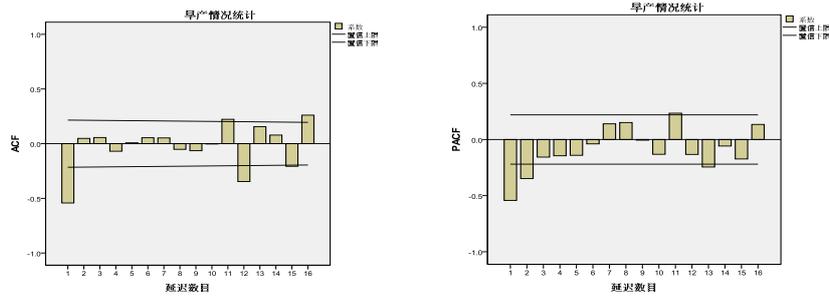


图3 自然对数、一阶差分、一阶季节差分自相关函数、偏相关函数图

2.2 在医院的早产儿情况统计中 ARIMA 模型估计和模型诊断

模型的参数估计见表 2，从表 2 中可知：（1）ARIMA (2,1,1) (0,1,1) 模型的所有参数都具有显著性差异，而 ARIMA (2,1,1) (1,1,1) 模型的参数有一部分没有显著性差异。（2）模型 ARIMA (2,1,1) (0,1,1) 的拟合效果最好，该模型的决定系数为 0.860，决定系数最高，正态化的 BIC 值为 5.408，BIC 值越小的模型，拟合效果越好。（3）由于进行了差分运算，因此常数项应该为 0，模型的最终形式为： $\nabla \nabla_{12} X_t = (1 - 0.673B) (1 - 0.863B^{12}) / (1 - 0.211B - 0.397B^2) \times a_t$ 。其中 B 为滞后算子， a_t 为随机干扰项。

表 2 自然对数化后模型的估计表

参数	ARIMA (2, 1, 1) (0, 1, 1)			ARIMA (2, 1, 1) (1, 1, 1)		
	估计值	t 值	P 值	估计值	t 值	P 值
常数	-0.003	-0.674	0.004	-0.014	-0.498	0.069
AR1	-0.211	-0.986	0.027	0.253	-1.082	0.008
AR2	-0.397	0.341	0.020	-0.142	-0.301	0.089
MA1	0.673	3.796	0.000	0.702	3.572	0.000

MA2	-	-	-	0.772	0.592	0.073
SMA1	0.863	2.792	0.010	0.872	2.214	0.032
平稳的 R 方		0.567			0.532	
R 方		0.860			0.754	
正态化的 BIC		5.408			5.371	
LB 统计量的 P 值		0.438			0.568	

根据 ARIMA (2,1,1) (0,1,1) 模型残差自相关函数和偏自相关函数图 4 可知：残差自相关函数和偏自相关函数都近似 0 阶截尾，基本都在虚线内，说明残差是一个近似的白噪声序列，序列中的相关性信息都被提取完全，所建立的模型是适合的。

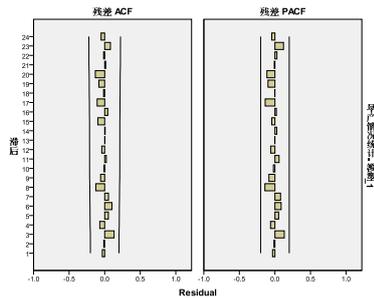


图 4 模型残差自相关函数和偏自相关函数图

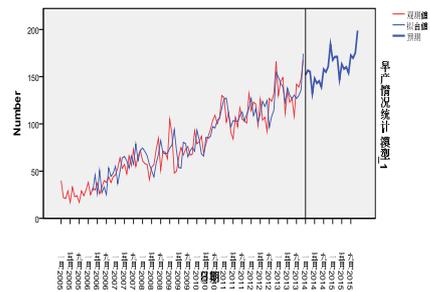


图 5 模型拟合值和观测值时序图

从 ARIMA (2,1,1) (0,1,1) 模型拟合值和观测值时序图 5 可以看出：图中红色线为实际值，蓝色线为预测值，拟合值和观测值几乎重合，模型的拟合效果较好，这也证实了前面的参数估计和函数图的分析。

2.3 在医院的早产儿情况统计中 ARIMA 模型预测

从 ARIMA (2,1,1) (0,1,1) 模型预测表 3 可知：（1）2014 年 1 月—2015 年 12 月在医院的早产儿情况中有向上的趋势性、季节性；（2）预测数据在 2 月、9 月、11 月和 12 月早产儿情况统计中呈现高峰。

表 3 ARIMA (2,1,1) (0,1,1) 模型预测表

模型	一月 2014	二月 2014	三月 2014	四月 2014	五月 2014	六月 2014	七月 2014	八月 2014	九月 2014	十月 2014	十一月 2014	十二月 2014
早产情况统计模型	预测 152	157	155	132	148	143	145	139	158	155	161	184
	UCL 180	187	190	172	191	190	196	192	215	215	223	250
	LCL 124	126	120	92	105	96	95	85	101	95	98	119
模型	一月 2015	二月 2015	三月 2015	四月 2015	五月 2015	六月 2015	七月 2015	八月 2015	九月 2015	十月 2015	十一月 2015	十二月 2015
早产情况统计模型	预测 168	171	171	147	163	158	160	154	173	170	175	199
	UCL 240	247	252	232	253	252	258	255	278	279	288	315
	LCL 95	95	90	61	74	65	63	53	67	61	63	83

3 结论

3.1 在医院的早产儿情况统计中 ARIMA 模型结果讨论。

本文根据湖南省妇幼保健院 2005 年—2013 年产房在医院早产儿情况的统计,运用时间序列 ARIMA 模型分析法,经过对原始数据序列的平稳化处理,模型的估计、诊断和预测,构建了 ARIMA (2,1,1) (0,1,1) 模型。该模型预测了湖南省妇幼保健院未来两年医院早产儿情况有较大的上升趋势,且呈现较强的周期性。这将为医院对预防早产儿情况的发生提供科学的依据。

3.2 在医院的早产儿情况统计中 ARIMA 模型研究意义。

(1) 本文运用 ARIMA 模型预测了医院未来两年的早产儿情况发展趋势,为医院相关政策的制订提供了科学依据,为产房管理提供可靠的数据参考。

(2) 根据模型预测分析结果,结合医院的实际发展情况,本文提出了几点建议:第一,医院未来两年的早产儿情况有向上的发展趋势,医院要针对早产儿发生的原因,加强宣传力度,从引起的病因、临床表现和预防手段方面积极宣传,定期对产妇进行产前检查,加强孕期指导,加强对高危妊娠的管理,积极治疗妊娠合并症,积极预防胎膜早破,预防亚临床感染,从而最终达到降低产妇早产情况发生的目的。第二,提高医院医生、助产士的业务能力。加强先兆早产的处理能力,尽可能延长妊娠周数。第三,提高助产士对早产临床的处理能力,尽力提高早产儿的存活率,减少并发症,预防新生儿呼吸窘迫综合症,临床后缩短第二产程等。第四,医院需要制定详细的预防早产情况发生的制度、处理方法和工作原则等相关文件,做到医院管理的制度化,使医生、助产士面对早产情况处理时有章可循,进一步完善医院管理的规范性。

4 参考文献

- [1]谢幸. 妇产科学[M].北京:人民卫生出版社(第8版),2013,3.
- [2]邓维斌,唐兴艳,胡大权,周玉敏. SPSS19 统计分析实用教程 [M]. 北京:电子工业出版社,2012:264-266.
- [3]袁世新. 应用移动平均趋势剔除法预测产妇分娩量[J]. 中国医院统计,2008,15(2):126-128.
- [4]严伟. 产妇分娩数季节变动的测定分析[J]. 中国医院统计,2007,14(4):341-343.
- [5]罗树生,安琳. 1999—2008 年中国住院分娩与孕产妇死因变化趋势分析[J]. 中国妇幼卫生杂志,2010,1(2):64-67.
- [6]刘洁,高茵茵等. 灰色系统 GM (1,1) 模型在孕产妇死亡率预测研究中的应用[J]. 实用预防医学,2011,18(2):217-219.
- [7]刘洁,何钦成. GM (1,1) 模型在 5 岁以下儿童死亡率预测研究中的应用[J]. 实用预防医学,2012,19(1):27-30.
- [8]郭海强,曲波等. 灰色系统 GM (1,1) 模型在我国梅毒发病预测研究中的应用[J]. 实用预防医学,2010,17(12):2397-2400.