

基于 ARMA 与 BP 神经网络模型的产品 质量安全风险预测

周荣喜 蔡小龙 崔清德 徐步祥

(北京化工大学 经济管理学院, 北京 100029)

摘 要: 主要通过对产品伤害人数的预测来表征产品质量安全风险。依据产品伤害人数的时间序列数据, 建立较为稳定的 ARMA 模型和 BP 神经网络模型对产品伤害人数做出了科学的预测; 并对两种预测方法进行了实证结果比较与分析, 为管控产品质量安全风险提供方法支持。

关键词: 产品质量安全; ARMA 模型; BP 神经网络; 风险预测

中图分类号: F272.1

引 言

产品质量安全事件的发生会极大地危害消费者权益及人身安全, 有效地管控产品质量安全风险具有重要意义。目前, 相关研究大多以食品为对象, 构建食品质量安全风险预警及评价体系^[1-3]。国外具有代表性的有欧盟食品与饲料快速预警系统、欧盟非食用消费品快速通报系统^[4-7]等, 国内也有类似的产品质量安全风险预测模型及预警系统, 如朱凌志等^[8]应用模糊综合评估法构建了乳业食品质量安全风险预测模型, 李仲超^[9]基于数理统计控制图理论建立了蜜饯质量安全风险预警系统。但目前这些系统大多属于定性预警, 基于定量分析产品质量安全风险预测的研究成果较少。

自回归移动平均模型(ARMA 模型)在短期线性预测中可以不断地调整模型来提高预测精度; 而 BP 神经网络模型在给定样本数据后可根据训练好的网络来达到高精度的预测效果, 两种方法对于不同时段上趋势值的预测有着很好的适应性。鉴于此, 本文试图用 ARMA 模型和 BP 神经网络模型对产品质量安全风险进行定量的预测, 预测结果可为监管部门及产品生产者管控产品质量安全风险提供更具体直观的方法参考。

1 数据来源

由于我国产品伤害监测系统建设正处于起步阶段, 还未配置相应的信息系统来存储和发布相关数据, 故本文以美国电子危害监督系统(简称 NEISS)中的数据为依据, 通过查阅系统发布的 2003 年至 2014 年充气娱乐相关受伤的人数报告(受伤人数在一定程度上可刻画充气娱乐产品质量安全风险的大小)得到数据如表 1 所示。其中 2003 年至 2011 年数据用以建模, 2012 年至 2014 年数据用以模型验证。

表 1 2003 ~ 2014 年受伤人数

Table 1 The number of injured people from 2003 to 2014

年份	观测值	年份	观测值	年份	观测值	年份	观测值
2003	172	2006	209	2009	297	2012	521
2004	163	2007	246	2010	367	2013	608
2005	202	2008	319	2011	477	2014	582

2 基于 ARMA 模型的产品质量安全 风险预测

2.1 数据预处理

为了消除异方差的影响, 对表 1 中的数据取对数后的变化趋势如图 1 所示。

从图 1 可看出受伤人数的对数值有上升中震荡的趋势, 需要对其做平稳性 ADF 检验。在显著性水平为 5% 的情形下, P 值为 0.9954, ADF 检验值为

收稿日期: 2015-06-26

基金项目: 国家科技支撑计划(2013BAK04B02)

第一作者: 男, 1972 年生, 教授

E-mail: zhourx@mail.buct.edu.cn

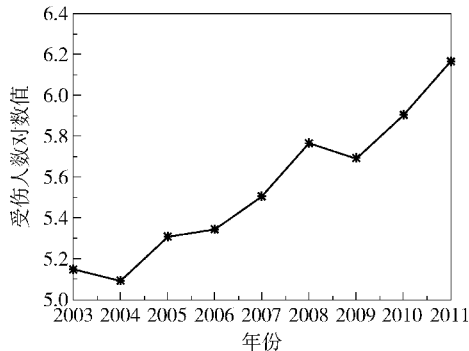


图1 受伤人数取对数后的变化趋势

Fig. 1 Logarithmic plot of the number of injured people

2. 8288, 临界值为 -1.9587 , 说明该受伤人数的对数时间序列不平稳, 需要对数据做差分运算。差分后的 ARMA 模型即为自回归积分移动平均模型 (ARIMA), 后者只是前者的变形, 但更常用^[10]。一阶差分后进行检验, P 值为 0.0349, ADF 检验值为 -2.1469 , 临界值为 -1.9594 , 说明该时间序列数据平稳, 可用于建模。

2.2 模型的识别

ARMA 模型的识别重点是考察其自相关函数 (ACF) 和偏自相关函数 (PACF) 的图形特征, 即利用给定序列的样本均值、自相关函数、偏自相关函数判定序列所属的模型类型。自相关系数和偏相关系数的计算所公式为

$$\gamma = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum (X_i - \bar{X})^2 \sum (Y_i - \bar{Y})^2}} \quad (1)$$

$$\gamma_{12,3} = \frac{r_{12} - r_{13}r_{23}}{\sqrt{(1-r_{13}^2)(1-r_{23}^2)}} \quad (2)$$

根据自相关与偏自相关图形特征可选定相应的模型, 表 2 列举了 $AR(p)$ 、 $MA(q)$ 、 $ARMA(p, q)$ 模型的自相关、偏自相关函数图形特征。

表 2 $AR(p)$ 、 $MA(q)$ 、 $ARMA(p, q)$ 模型的自相关、偏自相关函数图形特征

Table 2 Self-correlation and partial self-correlation function $AR(p)$, $MA(q)$ and $ARMA(p, q)$ models

模型类型	ACF	PACF
$AR(p)$	拖尾	p 步后截尾
$MA(q)$	q 步后截尾	拖尾
$ARMA(p, q)$	拖尾	拖尾

图 2 和图 3 分别是表 1 中数据的对数一阶差分序列后的 ACF 函数和 PACF 函数的图形。

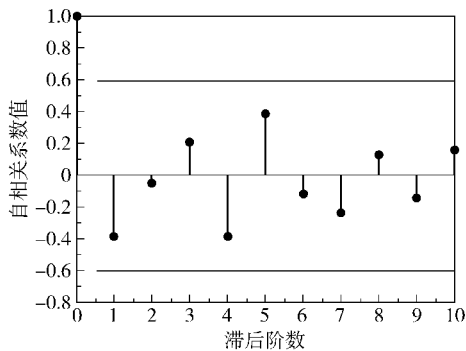


图2 受伤人数的对数后的 ACF 图形

Fig. 2 ACF of the logarithmic data for injured people

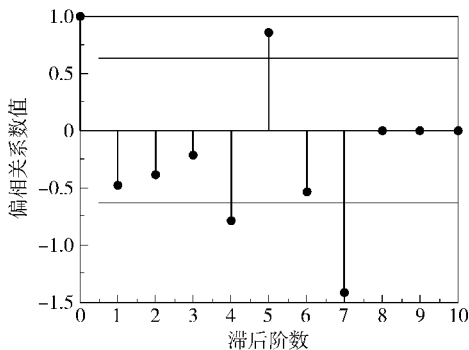


图3 受伤人数对数后的 PACF

Fig. 3 PACF of the logarithmic data for injured people

由图 2 可看出 ACF 的图形特征是拖尾, 而图 3 中 PACF 的图形特征是截尾。但仅依据图形做出判断还不够准确, 需要进一步识别。

2.3 模型定阶

自相关和偏自相关函数的图形特征只是有助于对模型的基本识别, 模型阶数的确定仍需进一步检验。通常模型定阶采用最小最终预报误差准则 (FPE) 和赤池信息量准则 (AIC), 其对应准则函数分别为

$$f_{FPE}(d) = \frac{N+d}{N-d} \hat{\sigma}^2 \quad (3)$$

$$f_{AIC}(d) = N \ln \hat{\sigma}^2 + 2d \quad (4)$$

其中, N 为观测时序样本个数, $\hat{\sigma}^2$ 为模型残差方差, d 为模型参数, 即模型的阶数。下面将结合具体数例进行分析。

以受伤人数对数一阶差分序列为例, 分别选取不同的 p, q 模型参数, 得出估计模型的 f_{FPE} 值和 f_{AIC} 值。不断改变模型的阶数, 综合考虑模型的 f_{FPE} 值和 f_{AIC} 值, 尽可能使两者的实验值达到极小, 这样对应的模型被认为是最佳模型。由表 3 综合得出 ARMA(1,2) 模型 f_{FPE} 值为 0.0161, f_{AIC} 值为

-3.9721,是所有模型中值最小的,故 ARMA(1,2) 为充气娱乐相关的受伤人数的最佳模型。

表 3 ARMA 模型的 f_{FPE} 值和 f_{AIC} 值

Table 3 Values of f_{FPE} and f_{AIC} in the ARMA model

(p,q)	f_{FPE}	f_{AIC}
(1,0)	0.0360	-3.3017
(0,1)	0.0303	-3.4764
(1,1)	0.0414	-3.0911
(2,0)	0.0285	-3.4798
(0,2)	0.0342	-3.2985
(2,1)	0.0208	-3.7179
(1,2)	0.0161	-3.9721
(2,2)	0.0694	-2.1819
(3,0)	0.0276	-3.3993
(0,3)	0.0204	-3.7357
(3,1)	0.0182	-3.7005
(1,3)	0.0733	-2.1270

根据时间序列理论,采用精确的极大似然估计方法求得模型参数,可得伤害人数值最佳预测模型 ARMA(1,2)为

$$Y_t=0.9674Y_{t-1}-1.611e_{t-1}+0.8124e_{t-2}+e_t \tag{5}$$

其中, Y_t 为第 t 期的值, e_t 为第 t 期的残差项。

2.4 模型预测效果验证及结果分析

利用 2012 年至 2014 年的数据进行验证,得到相应的预测值,相关结果见表 4。

表 4 原始值与预测值验证效果比较

Table 4 Comparison between original values and predictive values

年份	原始值	预测值	误差	相对误差/%
2012	521	536	15	2.9
2013	608	673	65	10.7
2014	582	639	57	9.8

由表 4 可知, ARMA(1,2)模型的最大、最小误差人数分别为 65 人和 15 人,平均相对误差为 7.8%,平均相对精度为 92.2%,说明 ARMA 模型的预测效果较好,选取 ARMA 模型是可行且合理的。

接下来,运用式(5)预测得到 2015 年至 2017 年充气娱乐相关受伤人数分别为 594 人、638 人和 640 人。预测结果表明,未来阶段充气娱乐产品安全质量风险将逐步上升,需加强风险管理,采取相应的风

险防范措施。

由于建立该 ARMA 模型的数据是取对数并做差分后才形成的平稳时间序列,且只是由有限的数
据构成,所以该模型反应的是短期的变化趋势,下
一步的研究重点应关注长期时间序列的预测模型。

3 基于 BP 神经网络模型的产品质量安全
安全风险预测

BP 神经网络是一种按误差反向传播算法训练
的多层前馈网络,是目前应用最广泛的神经网络模
型之一^[11]。下面将其用于预测产品质量安全风险。

3.1 BP 神经网络风险预测模型

根据 BP 神经网络的学习过程,首先初始化 BP
神经网络,创建一个新的向前神经网络为 Net_2 =
newff(P,T,[5,1],{'tansig','purelin'},'train-
lm'),并设置训练参数,迭代收敛曲线显示周期
trainparam.show = 50,训练速度 trainparam.lr =
0.05,训练次数 trainparam.epochs = 5000 等。然后
调用 Matlab 软件工具箱中的 TRAINGDM 算法训练
BP 网络,调用格式为[Net_2,tr] = train(Net_2,P,
T),其中,Net_2 为函数返回值,P 为输入信号,T 为
网络的目标。最后运用工具箱中 sim 函数对训练好
的网络进行预测,调用格式为 y = sim(Net_2,p_
test),其中 p_test 为待仿真的数据。

本文的样本训练数据为表 1 中 2003 年至 2014
年充气娱乐相关受伤人数,以第 t 年、 $t+1$ 年和 $t+2$
年的数据作为输入项,第 $t+3$ 年的数据作为输出
项,网络在自动学习样本的隐含规律后,以求导的方
式调节各层权值,权值节点导数大于 0,则说明权值
偏高,需适当减少,反之适当增大。重复上述过程,
便可使训练效果达到最佳。最后,将预测数据代入
已经完成函数学习的神经网络中,神经网络根据已
学习过的非线性映射关系便可得到结果,实现未来
年度充气娱乐相关受伤人数的预测。

3.2 模型预测效果验证及结果分析

在验证阶段,系统返回的训练输出值与目标输
出值拟合的情况如图 4,误差效果如图 5。由图可
知,此网络训练结果的最大误差为 64 人,最小误差
为 0;经计算得出,平均相对误差为 6.95%,平均相
对精度为 93.05%,达到了精度要求,训练效果较
佳。

在预测阶段,通过保存的神经网络重复迭代计
算后,可以预测得到 2015 年至 2017 年充气娱乐相

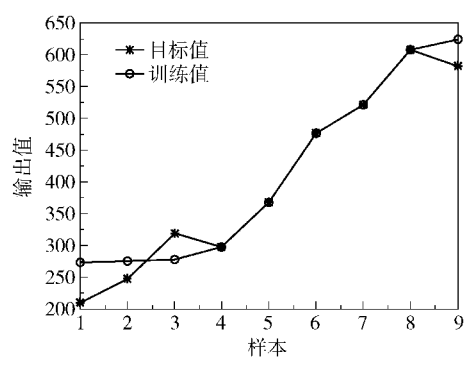


图4 BP训练值与目标值拟合图

Fig. 4 Comparison between training values and objective values

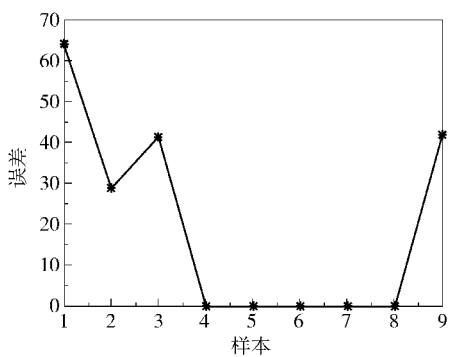


图5 误差效果图

Fig. 5 Estimation of the error between predictive values and objective values

关受伤人数分别为 588 人、635 人和 641 人。该预测结果同样表明,未来阶段充气娱乐产品质量安全风险仍在持续扩大,需适时调整风险管理措施,做好风险预警工作。

4 两种预测模型方法比较

ARMA 模型和 BP 神经网络是不同形式的预测模型,通过对产品质量安全风险预测实证研究,可以对二者的优缺点、预测误差和预测精度等方面做出比较。结果如表 5 所示。

表5 两种预测模型结果的比较

Table 5 Comparison of the two prediction model

项目	BP 神经网络	ARMA 模型
所属领域	人工智能	统计
模型优点	预测精度高、学习能力强	适用性广泛、拟合效果较好
模型缺点	数据预处理复杂	考虑面窄、局限性
平均相对误差/%	6.95	7.8
平均相对精度/%	93.05	92.2

作为人工智能领域的预测方法,BP 神经网络以其特有的层次网络结构,实现非线性函数学习功能,根据训练好的网络预测得到结果,在解决许多复杂机理问题时更能体现其强大之处,如果选取的样本合理,样本规律与整体规律统一,BP 神经网络将能实现高精度的预测;而 ARMA 模型对平稳时间序列进行建模,且根据风险评分本身的特性进行预测,拟合效果较好,适用性更广泛。由表 5 看出,此次预测结果中,ARMA 模型的预测精度略低,误差略大,这是由于数据差分预处理使得其精确度没有 BP 神经网络高。虽然 ARMA 模型只适合短期预测,但它能够揭示预测对象在时间序列上的先后影响关系。总之,二者各有优缺点,从预测效果上看,BP 神经网络模型略优于 ARMA 模型。

5 结束语

本文将 ARMA 模型与 BP 神经网络模型应用到充气娱乐产品质量安全风险预测中,充分利用 ARMA 模型在短期线性内可不断调整的特点以及 BP 神经网络模型的模拟生物机理寻找事物潜在规律的特点,预测得到 2015 年至 2017 年充气娱乐相关受伤的人数。从预测结果上看,两模型的预测数据很接近,预测精度相对较高,证明了该模型的可行性和合理性。

参考文献:

[1] 顾小林, 张大为, 张可, 等. 基于关联规则挖掘的食品安全信息预警模型[J]. 软科学, 2011, 25(11): 136-141.
Gu X L, Zhang D W, Zhang K, et al. The information pre-warning model of food safety based on association rules mining [J]. Soft Science, 2011, 25(11): 136-141. (in Chinese)
[2] Forristal P M, Wilke D L, McCarty L S. Improving the quality of risk assessments in Canada using a principle-based approach [J]. Regulatory Toxicology and Pharmacology, 2008, 50: 336-344.
[3] 毛丽君. 我国食品安全预警体系的构建—以欧盟对我国食品通报为视角[J]. 江西社会科学, 2013(11): 58-61.
Mao L J. The construction of food safety early warning system in China—In perspective of food bulletin which the European Union gave China [J]. Jiangxi Social Sciences, 2013(11): 58-61. (in Chinese)
[4] Houghton J R, Rowe G, Frewer L J, et al. The quality of

- food risk management in Europe: perspectives and priorities[J]. Food Policy, 2008, 33: 13–26.
- [5] 焦阳, 郭力生, 凌文涛. 欧盟食品安全的保障——食品、饲料快速预警系统[J]. 中国标准化, 2006(3): 20–21/29.
- Jiao Y, Guo L S, Ling W T. EU food safety guarantee-food, feed rapid alert[J]. China Standardization, 2006(3): 20–21/29. (in Chinese)
- [6] Wijtzes T, van't Riet K, Huis in't Veid J H J, et al. A decision support system for the prediction of microbial food safety and food quality[J]. International Journal of Food Microbiology, 1998, 42: 79–90.
- [7] Rotondo A, Young P, Geraghty J. Quality risk prediction at a non-sampling station machine in a multi-product, multi-stage, parallel processing manufacturing system subjected to sequence disorder and multiple stream effects[J]. Annals of Operations Research, 2013, 209: 255–277.
- [8] 朱凌志, 周宇. 运输调度过程中乳业食品质量安全风险预测模型研究[J]. 现代计算机, 2015(2): 24–28.
- Zhu L Z, Zhou Y. Research on dairy food quality and safety risk prediction model in transportation scheduling process[J]. Modern Computer, 2015(2): 24–28. (in Chinese)
- [9] 李仲超. 食品质量安全风险预警系统的建立[J]. 管理观察, 2014(31): 189–192.
- Li Z C. The warning system establishment of food quality and safety risk[J]. Management Observer, 2014(31): 189–192. (in Chinese)
- [10] 宋子房. 公路短时车流量预测模型研究[J]. 科学决策, 2014(4): 83–94.
- Song Z F. Study on short-term traffic volume prediction model[J]. Scientific Decision Making, 2014(4): 83–94. (in Chinese)
- [11] 周荣喜, 徐步祥, 邱苑华, 等. ANP-BP 铁路系统安全风险评估模型及应用[J]. 科技进步与对策, 2014, 31(11): 90–95.
- Zhou R X, Xu B X, Qiu W H, et al. The application of railway security risk assessment model based on ANP-BP[J]. Science & Technology Progress and Policy, 2014, 31(11): 90–95. (in Chinese)

Safety risk prediction of product quality based on the auto-regressive moving average prediction (ARMA) model and the back propagation (BP) neural networks model

ZHOU RongXi CAI XiaoLong CUI QingDe XU BuXiang

(School of Economics and Management, Beijing University of Chemical Technology, Beijing 100029, China)

Abstract: This paper monitors the safety risk of product quality by forecasting the number of injured people. First, based on the time series data of the number of people injured by the product, the auto-regressive moving average (ARMA) prediction model and the back propagation (BP) neural network model are employed to forecast the the number of people injured by the product. The predictions of the two models are compared and analyzed. The results offer a method for the safety risk management of product quality.

Key words: product quality and safety risk; ARMA model; BP neural network; risk prediction