

文章编号: 1671- 251X(2010) 12- 0027- 03

基于一元线性回归的中国煤炭行业年死亡人数与 年煤炭产量的预测研究

王文才, 乔旺, 王瑞智, 李刚

(内蒙古科技大学矿业工程学院, 内蒙古 包头 014010)

摘要: 为了预测我国煤炭行业未来安全形势的发展状况, 建立了我国煤炭行业年死亡人数与年煤炭产量的一元线性回归预测模型, 采用该模型对我国未来 3 年煤炭行业年死亡人数与年煤炭产量的发展趋势进行了预测。预测结果表明, 该模型具有一定的可行性, 为预测我国煤炭行业的安全形势提供了可靠的理论依据。

关键词: 煤炭行业; 死亡人数; 煤炭产量; 一元线性回归; 预测

中图分类号: TD79 **文献标识码:** A

Prediction Research of Annual Death Numbers and Annual Coal Production in China's Coal Industry Based on Simple Linear Regression

WANG Wen cai, QIAO Wang, WANG Rui zhi, LI Gang

(School of Mining Engineering of Inner Mongolia University of Science and Technology,
Baotou 014010, China)

收稿日期: 2010- 08- 23

作者简介: 王文才(1963-), 男, 内蒙古伊金霍洛旗人, 教授, 博士, 硕士研究生导师, 现主要从事安全工程及矿业技术经济的教学与研究工作。E-mail: wencai88888@yahoo.com.cn

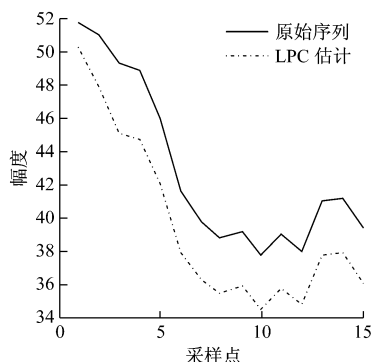


图 3 3 阶 LPC 估计值与原始信号比较曲线

得出 $y_{16} = 38.89$, $y_{17} = 38.25$, $y_{18} = 39.54$, 由式 (11) 得出平均误差率为 4.3%, 说明瓦斯涌出量预测值比较可靠。

4 结语

时间序列预测法是工程统计中的常用方法, 利用它可建立煤矿瓦斯涌出量与深度的函数关系, 而对矿井未开采的区域进行预测, 具有很强的实际

操作性; 为了增强该方法的通用性, 需要开展广泛的实验, 更多地获得煤矿瓦斯涌出量在各种环境条件下的实测数据, 建立一套完整详细的瓦斯涌出量数据库, 进而选定条件由数据库中的实测数据对瓦斯涌出量进行仿真预测; 在建立时间序列时, 要考虑数据的准确性, 同时还要考虑煤层赋存条件及地质构造的影响, 以提高预测的准确性。

参考文献:

- [1] 施式亮, 宋译, 何利文, 等. 矿井掘进工作面瓦斯涌出混沌特性判别研究[J]. 煤炭学报, 2006, 31(6): 58-62.
- [2] RABINER L R. 语音数字信号处理[M]. 北京: 电子工业出版社, 1993.
- [3] QUATITERI T F. 离散时间语音信号处理——原理与应用[M]. 北京: 电子工业出版社, 2004.
- [4] 杨位钦, 顾岚. 时间序列分析与动态数据建模[M]. 北京: 北京工业学院出版社, 1986.
- [5] 宿晶亮, 栗苹, 刘宁. AR 模型在直升机声学环境模拟中的应用[J]. 探测与控制学报, 2001, 23(2): 45-48.

Abstract: In order to predict development status of future safety situation of China's coal industry, a prediction model of simple linear regression for annual death numbers and annual coal production in China's coal industry was established, and the model was used to predict annual death numbers and annual coal production of China's coal industry in three years. The prediction result showed that the model has certain feasibility, which can provide reliable theoretical basis for predicting safety situation of China's coal industry.

Key words: coal industry, death numbers, coal production, simple linear regression, prediction

0 引言

随着我国煤炭年产量的连年增长,煤炭行业死亡的人数也是一个触目惊心的数字。因此,对我国煤炭行业死亡人数进行科学的分析和预测,对于煤炭行业的安全有着重要意义^[1-2]。

一个企业或部门的安全状况,受其生产性质、规模、人员素质、物质条件、环境状况以及管理水平等一系列因素的影响,它们之间的函数关系由于影响因素众多、关系复杂而难以确定。然而大量统计资料表明,一个企业或部门的安全状况与影响它的各种因素却是一个密切联系着的整体。这个整体具有相对稳定性和持续性,即时间序列平稳性。这就为抛开对逐个因素的分析,就其整体利用惯性原理对企业或部门的安全状况进行预测提供了可能。一元线性回归分析具有预测结果比较接近实际、易于表示数据的离散性并给出预测区间等优点,在工矿企业伤亡事故趋势预测中已得到广泛应用^[3]。基于此,笔者建立了一种一元线性回归预测模型,用以预测我国未来 3 年煤业行业年死亡人数与年煤炭产量的发展趋势,预测结果较准确、可靠。

1 一元线性回归模型的建立

1.1 一元线性回归模型

设一元线性回归方程^[4-5]为 $y = a + bx$, 其中 x 、 y 分别为自变量和因变量, a 、 b 均为模型参数, 分别表示直线的纵截距和斜率。

设有 n 对 x 与 y 的数值, 令

$$W(a, b) = \sum (y - a - bx)^2 \quad (1)$$

当 $W(a, b)$ 值最小时, 采用最小二乘法可求得

$$a = \frac{\sum x \sum xy - \sum x^2 \sum y}{(\sum x)^2 - n \sum x^2} \quad (2)$$

$$b = \frac{\sum x \sum y - n \sum xy}{(\sum x)^2 - n \sum x^2} \quad (3)$$

1.2 一元线性回归模型精度检验

在回归分析中, 还应研究计算得到的回归直线是否符合实际数据变化的趋势。因此引入相关系数 r 的概念^[6-8], 其计算公式为

$$r = \frac{L_{xy}}{\sqrt{L_{xx} L_{yy}}} \quad (4)$$

$$\text{式中: } L_{xy} = \sum xy - \frac{1}{n} \sum x \sum y; L_{xx} = \sum x^2 - \frac{1}{n} (\sum x)^2; L_{yy} = \sum y^2 - \frac{1}{n} (\sum y)^2。$$

相关系数 r 取不同的数值时, 分别表示实际数据和回归直线之间的不同符合情况:

(1) $r = 0$ 时, 表示回归直线不符合实际数据的变化情况。

(2) $0 < |r| < 1$ 时, 表示回归直线在一定程度上符合实际数据的变化趋势。 $|r|$ 越大, 说明回归直线与实际数据变化趋势的符合程度越大; $|r|$ 越小, 则符合程度越小。

(3) $|r| = 1$ 时, 表示回归直线完全符合实际数据的变化情况。

2 全国煤炭行业死亡人数与煤炭产量预测

2.1 模型建立

我国煤炭行业 2001—2009 年的年死亡人数与年煤炭产量如表 1 所示^[9], 根据表 1 绘制出的年死亡人数与年煤炭产量散点图分别如图 1 和图 2 所示。

表 1 我国煤炭行业 2001—2009 年的年死亡人数与年煤炭产量

年份序号 x	年份	死亡人数 y_1 / 人	年产量 y_2 / 亿 t
1	2001	5 670	10.89
2	2002	6 995	14.20
3	2003	6 702	17.36
4	2004	6 027	19.56
5	2005	5 986	21.13
6	2006	4 746	23.30
7	2007	3 786	25.23
8	2008	3 292	27.16
9	2009	2 700	29.10

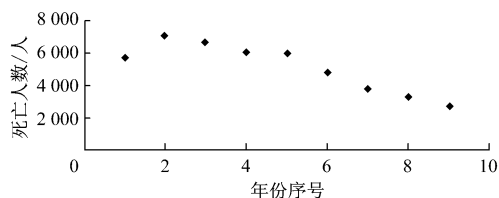


图1 年死亡人数散点图

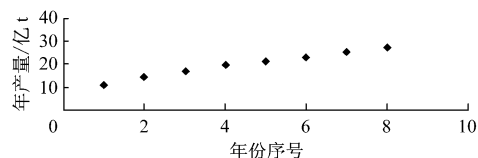


图2 年煤炭产量散点图

根据表1,由式(2)、式(3)得到我国煤炭行业年死亡人数的一元线性回归模型参数 a 、 b 的值分别为7 608.94和-501.7,则我国煤炭行业年死亡人数的一元线性回归模型为

$$y_1 = 7\,608.94 - 501.7x \quad (5)$$

根据表1,由式(2)、式(3)得到我国煤炭行业年煤炭产量的一元线性回归模型参数 a 、 b 的值分别为13.36和1.65,则我国煤炭行业年煤炭产量的一元线性回归模型为

$$y_2 = 13.36 + 1.65x \quad (6)$$

2.2 预测精度检验

由式(4)得到我国煤炭行业年死亡人数和年煤炭产量的一元线性模型检验相关系数分别为 $r_1 = 0.837$ 、 $r_2 = 0.926$ 。

相关系数 r_1 和 r_2 均接近于1,说明实际数据变化趋势与式(5)、式(6)的预测趋势符合程度比较大。由此可知,采用式(5)、式(6)预测的我国煤炭行业未来年死亡人数与年煤炭产量具有较大的参考价值。

2.3 未来3年我国煤矿死亡人数的预测

将 $x = 10$ 、11和12代入式(5),可以预测出2010年、2011年和2012年全国煤炭行业年死亡人数分别为2 592人、2 090人和1 589人,即未来3年我国煤炭行业年死亡人数呈整体下降趋势。将 $x = 10$ 、11和12代入式(6),可以预测出2010年、2011年和2012年我国煤炭行业年煤炭产量分别为29.76亿t、31.46亿t和33.16亿t,即未来3年我国煤炭总产量呈增长趋势。

2.4 预测结果分析

预测结果:未来3年内我国煤炭年产量呈递增趋势,而我国煤矿的年死亡人数呈递减趋势。从预测结果可看出,在煤炭年产量逐年增加的前提下,我国煤炭年死亡人数却逐年减少,说明我国煤炭行业

随着开采技术的提高,安全技术和措施也在不断地提高和完善,其安全效益已开始逐步显现。

3 结语

(1)就目前的煤炭开采技术、安全措施和市场环境,可以采用一元线性回归模型预测煤炭行业的年死亡人数和年煤炭产量数据,从而为进一步做好我国煤炭行业的开发和安全生产工作提供可靠的理论依据。

(2)根据安全科学理论可知,一定水平的科学技术、安全措施和行业环境对应一定水平的行业年死亡人数。由于我国煤炭科技发展水平、煤矿管理水平、人为因素、煤矿的地质及开采因素等众多因素的限制,我国煤炭行业年死亡人数可能降低到某一数值后会变化缓慢甚至停滞在这一数值上,还有可能会再次回升。因此,必须不断更新技术,实施先进的安全技术,建立健全的安全规章制度,才能保证我国煤炭行业在煤炭产量按目前增长率增长的前提下,其年死亡人数一直保持逐年下降或恒定趋势。

(3)上述预测结果成立的前提是我国煤炭行业煤炭产量按目前增长率增长。如果我国未来煤炭行业煤炭产量增长率与目前煤炭产量增长率相差较大时,需重新建立预测模型。

参考文献:

- [1] 金龙哲,宋存义.安全科学原理[M].北京:化学工业出版社,2004.
- [2] 付丽华,王晶.一元线性回归分析在煤矿企业中的应用[J].煤矿现代化,2006,72(3):87-88.
- [3] 韩中庚,宋明武,邵广纪.数学建模竞赛[M].北京:科学出版社,2007.
- [4] 郭志军.应用Excel对一元线性回归模型的分析[J].宁波职业技术学院学报,2009(5):57-60.
- [5] 邵碧雄,叶左局.利用Excel软件进行多元回归与多项式回归分析[J].广东奶业,2006(2):11-14.
- [6] 王义宏.基于多元线性回归分析的县域经济发展程度的评价——以江阴为例[J].生产力研究,2009(18):27-29.
- [7] 高芳,崔勇.多元线性回归分析在房地产市场中的应用[J].河南机电高等专科学校学报,2009(3):41-43.
- [8] 葛培运.主成分回归分析在经济学中的应用[J].科技信息,2009(27):209-210.
- [9] 蔡昉.2009年全国煤炭安全生产总结汇报[R].北京:社会科学文献出版社,2010.