

基于灰色预测 - 决策树模型下车险保费浮动的研究

肖思佳, 林萍芝, 张靖文, 王浩华*

海南大学理学院, 海南 海口

Email: huazi8112@hainanu.edu.cn

收稿日期: 2021年6月23日; 录用日期: 2021年7月15日; 发布日期: 2021年7月27日

摘要

汽车保险的续保问题对于保险市场具有举足轻重的作用, 如何有效的预测市场各个因素对续保率的影响并未十分清楚。通过分析保费浮动优惠等因素对续保率的影响, 本文建立影响力指数的等级比较模型, 通过计算其性质系数的灰色关联度, 考察具有不同性质车辆的续保率。计算出有关索赔系数和使用性质系数, 建立数学模型, 给出了较为合理的保费浮动方案。最后, 建立决策树模型, 根据统计数据对不同车辆进行分类同时, 通过对三个因素进行加权, 得到了评价企业质量的决策树累积模型, 分析各因素对续保率的影响, 从而提供可行性建议。

关键词

等级比较模型, 续保率, 灰色预测法, 保费浮动系数, 决策树

Study on the Floating Premium of Car Insurance Based on Grey Prediction Method and Decision Tree

Sijia Xiao, Pingzhi Lin, Jingwen Zhang, Haohua Wang*

School of Sciences, Hainan University, Haikou Hainan

Email: huazi8112@hainanu.edu.cn

Received: Jun. 23rd, 2021; accepted: Jul. 15th, 2021; published: Jul. 27th, 2021

*通讯作者。

文章引用: 肖思佳, 林萍芝, 张靖文, 王浩华. 基于灰色预测 - 决策树模型下车险保费浮动的研究[J]. 应用数学进展, 2021, 10(7): 2579-2591. DOI: [10.12677/aam.2021.107268](https://doi.org/10.12677/aam.2021.107268)

Abstract

This paper intends to predict the floating premium of car insurance in an insurance company, use the renewal rate to measure the floating demand of premium, collect the data of relevant departments, and establish the grade comparison model of influence index by analyzing the influence of each factor on the renewal rate. Then the grey correlation degree is used to predict the influence of vehicles with different use properties on the renewal rate, and the use property coefficient is calculated. According to the model of “premium floating coefficient = claim coefficient A × use property coefficient B”, a more reasonable premium floating scheme is put forward. Finally, the decision tree is established, combined with the statistical data, the vehicles with different usage properties are weighted, and the three factors that evaluate the quality of the company are assigned weights, and the cumulative model of the decision tree is obtained. The results show that the scheme evaluation is effective.

Keywords

Grade Comparison Model, Renewal Rate, Grey Prediction Method, Premium Floating Coefficient, Decision Tree

Copyright © 2021 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

续保率一定程度上反映了保险公司业务稳定程度和对客户的保留能力，其成为汽车保险行业评估保险公司市场份额的决定性因素。实际上，续保率由保费浮动等度量，给出一套较为合理的保费浮动方案已成为保险公司提高续保率的重要依据。目前为止，已经对续保率影响因素从不同角度有了研究。陈曼发现在众多不同种类车辆中家庭自用车的续保率最低[1] [2] [3]。刘亮利用 K-means 聚类算法将客户划分为某几类，并用多元回归分析得到每一类客户的主要影响因素进行了对比[4]。王梦晨选择广义线性模型中的 logistic 模型为工具从多个层面分析了车险业务中影响续保率的因素[5]。本文在此基础上，收集相关部门数据，综合考虑了各方面因素的影响，建立了影响力指数的等级比较模型，利用灰色预测法进行灰色关联度预测。进而根据“保费浮动系数 = 索赔系数 A × 使用性质系数 B”这一模型，提出了较为合理的保费浮动方案。最后建立决策树模型，判断方案可行性。

2. 建立等级比较模型

首先，从原始表格中提取出“承包车辆的使用性质”，“承包车辆的出现次数”、“承包车辆的年龄”和“承包车辆的新车购买价”这四个具有代表性的市场因素，使用 7 个月份的数据(来源：中国统计局)，建立新的数据表格，进行基础数据处理后，建立图 1。

选取前四个月的续保率数据来进行研究，这样就避免了研究全部 7 个月份时由于前四个月与后三个月的数据变化较大而造成的误差。提取“承包车辆的出现次数”、“承包车辆的新车购买价”、“承包车辆的使用性质”和“承包车辆的年龄”这四个代表性因素的 9、10、11、12 四个月份的数据，根据方差公式

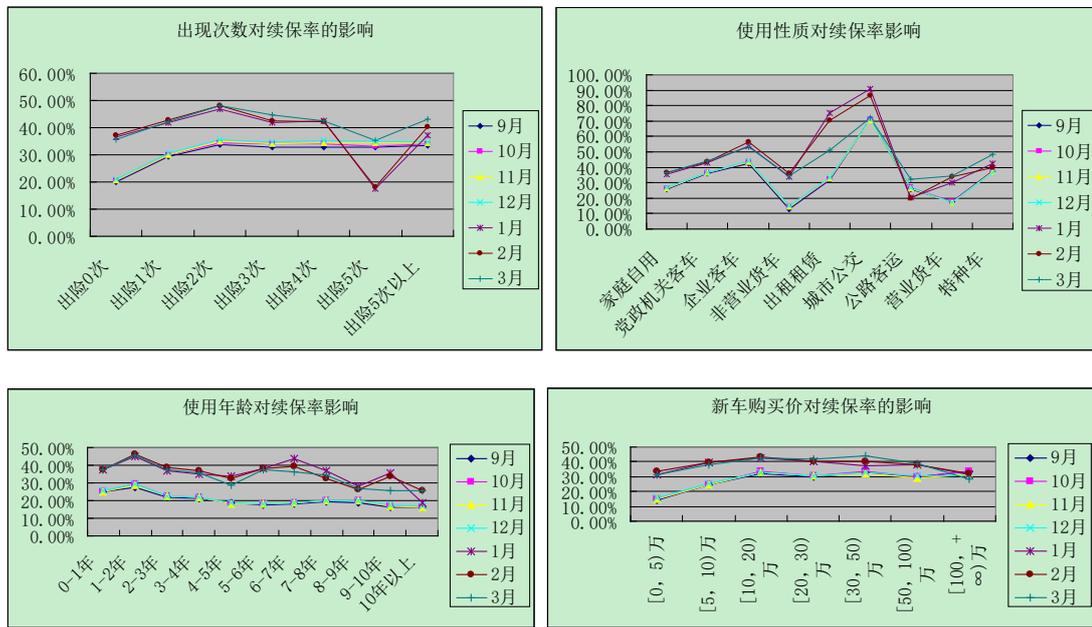


Figure 1. The effect of different factors on renewal rate in seven months
图 1. 7 个月份中不同因素对续保率的影响图

$$\mu = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2,$$

得到各个市场因素 9、10、11、12 月份的四个方差，进而得到影响力评估指数[4] [5] [6]

$$V = \sum_{i=9}^{12} \mu_i。$$

将保险业市场因素对该企业续保率的影响力评估指数分为 A、B、C、D、E 五个等级比较模型，得到表 1。

Table 1. Grade comparison model of influence evaluation index
表 1. 影响力评估指数的等级比较模型

等级	A	B	C	D	E
影响力	很强	较强	一般	较弱	较弱
评估指数范围	≥0.2	0.2~0.015	0.015~0.01	0.01~0.0005	≤0.0005

将四个因素对续保率的影响力评估指标数据带入等级比较模型，获得每个因素的代表等级，得到表 2。

Table 2. Table of factors influence evaluation index and grade
表 2. 各个因素影响力评估指数与影响力等级对应表

	V_c	V_g	V'_c	V'_g	V_N	V_s
评估指数	0.010313622	0.016174167	0.000124356	0.000801448	0.013664124	0.223500188
等级	一般	较强	很弱	较弱	一般	很强

3. 续保率的灰色预测模型

从 2016 年 9 月至 2017 年 3 月七个月的到期车辆续保率的原始数据, 得表 3。

Table 3. Renewal rate of family car

表 3. 家庭自用车续保率

月份	9	10	11	12	1	2	3
续保率	25.81%	26.53%	26.48%	27.05%	35.68%	36.81%	36.37%

通过计算(见附录), 可知 GM(1,1)模型的参数 $a = -0.0802$, $\mu = 0.2244$, 即其微分方程为[7] [8]:

$$\frac{dX^{(1)}}{dt} + aX^{(1)} = \mu, X^{(1)} = 0.2244。$$

GM(1,1)灰色微分方程时间响应序列为

$$\hat{x}^{(1)}(k+1) = \left(x^{(0)}(1) - \frac{\mu}{a} \right) e^{-ak} + \frac{\mu}{a} = 3.0561e^{0.0802k} - 2.798。$$

其中 $X^{(1)}$ 的模拟值为

$$\begin{aligned} \hat{X}^{(1)} &= \{ \hat{x}^{(1)}(1), \hat{x}^{(1)}(2), \hat{x}^{(1)}(3), \hat{x}^{(1)}(4), \hat{x}^{(1)}(5), \hat{x}^{(1)}(6), \hat{x}^{(1)}(7) \} \\ &= \{ 0.2581, 0.5133, 0.7897, 1.0892, 1.4183, 1.7654, 2.1464 \} \end{aligned}$$

由 $\hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k)$, 还原出 $X^{(0)}$ 的模拟值得:

$$\begin{aligned} \hat{X}^{(0)} &= \{ \hat{x}^{(0)}(1), \hat{x}^{(0)}(2), \hat{x}^{(0)}(3), \hat{x}^{(0)}(4), \hat{x}^{(0)}(5), \hat{x}^{(0)}(6), \hat{x}^{(0)}(7) \} \\ &= \{ 0.2581, 0.2552, 0.2765, 0.2995, 0.3245, 0.3516, 0.3810 \} \end{aligned}$$

进而计算 X 与 \hat{X} 的灰色关联度 \hat{S} 以及其相对误差 ε

$$\begin{aligned} |S| &= \left| \sum_{k=2}^6 (x(k) - x(1)) + \frac{1}{2}(x(7) - x(1)) \right| \\ &= \left| (0.2653 - 0.2581) + (0.2648 - 0.2581) + (0.2705 - 0.2581) \right. \\ &\quad \left. + (0.3568 - 0.2581) + (0.3681 - 0.2581) + \frac{1}{2} \times (0.3637 - 0.2581) \right| \\ &= 0.6574 \end{aligned}$$

$$|\hat{S}| = \left| \sum_{k=2}^6 (\hat{x}(k) - \hat{x}(1)) + \frac{1}{2}(\hat{x}(7) - \hat{x}(1)) \right| = 0.7084$$

$$|\hat{S} - S| = \left| \sum_{k=2}^6 [(x(k) - x(1)) - (\hat{x}(k) - \hat{x}(1))] + \frac{1}{2}[(x(7) - x(1)) - (\hat{x}(7) - \hat{x}(1))] \right| = 0.0510$$

$$\varepsilon = \frac{1 + |S| + |\hat{S}|}{1 + |S| + |\hat{S}| + |\hat{S} - S|} = \frac{1 + 0.6574 + 0.7084}{1 + 0.6574 + 0.7084 + 0.0510} = 0.9789 > 0.9。$$

根据得到的结果对比精度检验等级参照表可知(附录), 一级精度的家庭自用这一使用性质对续保率有很显著的影响。

同样的方法，分别对以下几种车应用灰色预测法进行精度检验，得到表 4。

Table 4. Table of grey relational degree of using properties and renewal rate

表 4. 使用性质与续保率灰色关联度列表

	家庭自用	党政机关客车	企业客车	非营业货车	出租租赁	营业货车	特种车
平均相对误差	0.0027	0.0014	0.0045	0.0149	0.1141	0.0056	0.0032
灰色关联度	0.0620	0.0344	0.0516	0.1835	0.2423	0.1229	0.0511
残差平方和	0.9789	0.9876	0.9672	0.9623	0.9059	0.9815	0.9585

结果表明，灰色关联度分别是党政机关客车为、企业客车为、非营业货车为、出租租赁为、营业货车为、特种车为，都大于 0.9，所以它们的续保率的精度都为一级，这也即是证明车辆使用性质这一因素会明显影响公司的续保率。

4. 保费浮动方案

首先把商车险和交强险分开，分别统计两个险别中各类车的使用的投保辆数、赔付赔付款总额、出险车辆数、浮动前保费总额、总的车辆数。其中在这两种险别中，特种车很少发生事故，所占的比例小，可以不考虑。统计结果得到表 5。

Table 5. Statistics table of commercial and compulsory insurance

表 5. 保户投保商车险和交强险的投保车辆数统计表

	家庭自用	企业非营业用车	出租、租赁车	营业挂货车	城市公交	特种车	党政机关、事业团体
商车险车辆数	1877	519	69	745	2	49	153
百分比/%	56.14	15.52	2.06	22.29	0.06	1.47	4.58
交强险车辆数	2884	296	166	908	1	79	147
百分比/%	58.09	5.96	3.34	18.29	0.02	1.59	2.96

其中百分比是该种车辆数占商车险或交强险总投保车辆的比例，在两个险别中城市公交与特种车投保车辆都非常少，故不予以考虑。若投保客户的投保额小，但客户发生事故的次数多，那总赔付款与出险次数成正比，公司的盈利与出险次数成反比。下面根据车辆使用性质的不同，统计出两个险别的其他参数数据。得到表 6 和表 7。

Table 6. The data of commercial insurance

表 6. 商车险数据统计表

	家庭自用车	企业非营业用车	出租、租赁车	营业挂货车	党政机关、事业团体
投保车辆数	1877	519	69	745	153
赔付款总额(万)	470.75	75.61	0.26	249.76	18.07
出险车辆数(辆)	916	84	27	266	39
出险车辆总保额(万)	35559.17	7055.62	405.00	11771.48	1440.69
浮动前保费总额(万)	975.35	190.08	19.79	735.38	42.79

Continued

A	0.49	0.17	0.39	0.36	0.25
B	1.32%	1.07%	0.06%	2.12%	1.25%
C	48.26%	39.78%	1.30%	33.94%	42.24%

Table 7. The data of compulsory insurance

表 7. 交强险数据统计表

	家庭自用车	企业非营业用车	出租、租赁车	营业货车	党政机关、事业团体
投保车辆数	2884	296	166	908	147
赔付款总额(万)	110.26	4.52	25.65	174.62	13.73
出险车辆数(辆)	538	49	71	244	21
出险车辆总额(万)	6563.60	597.80	866.20	2976.80	256.20
浮动前保费总额(万)	291.03	23.74	29.52	246.31	14.57
A	0.81	0.83	0.57	0.73	0.85
B	1.68%	0.76%	2.96%	5.87%	5.36%
C	37.89%	19.05%	86.88%	70.90%	94.23%

由以上数据可以计算出：

$$\text{出险车辆数比例} A = \frac{\text{出险的车辆数}}{\text{投保的车辆数}}$$

$$\text{赔付款占出险车辆浮动前保费总额比例} B = \frac{\text{赔付款总额}}{\text{出险车辆浮动前保费总额}}$$

$$\text{赔付款占投保车辆浮动前保费总额比例} C = \frac{\text{赔付款总额}}{\text{投保车辆浮动前保费总额}}$$

根据以上三种比例可得到三种方案($i = 1$ 表示商车险, $i = 2$ 表示交强险):

方案一: 浮动后保额 = $(1 - A_i)$ 浮动前保额

方案二: 浮动后保额 = $(1 - B_i)$ 浮动前保额

方案三: 浮动后保额 = $(1 - C_i)$ 浮动前保额

代入数据结果如表 8。

Table 8. The programme of commercial and compulsory insurance

表 8. 商车险和交强险方案统计表

使用性质系数 B	家庭自用车	企业非营业用车	出租、租赁车	营业货车	党政机关、事业团体	
商车险	方案一	0.51	0.83	0.61	0.64	0.75
	方案二	0.99	0.99	0.99	0.987	0.99
	方案三	0.52	0.60	0.99	0.66	0.58
交强险	方案一	0.75	0.95	0.57	0.68	0.85
	方案二	0.98	0.99	0.97	0.94	0.95
	方案三	0.62	0.81	0.13	0.29	0.06

方案一根据出险车辆数的比例提出的,只考虑了所投保的车辆是否会发生事故,却没考虑其他因素,例如公司的保户个数、总保费、赔付款总额。该方案考虑的并不全面,所以不能广泛推广,作为辅助参考一下。

方案二的提出是根据出险车辆赔付款占浮动前保费总额比例,充分考虑了各个因素,因而具有普适性,推荐采用此方案。

方案三以保险公司的利益为基本,考虑了赔付款对投保车辆浮动前保费总额的占比,但是没有考虑到出险的车辆数。如果一个公司在这一年利益不错的情况下,但是如果处理客户出险事故的次数太多的话,就会造成人力资源在该业务倾斜从而被浪费,由此可以看出这个方案有一定的局限性。

因此方案二更可行。

根据索赔系数 A 与使用性质系数 B,结合模型[9][10],得到最终的保费浮动系数方案表 9。

Table 9. The final programme of floating premium

表 9. 最终的保费浮动方案

保费浮动系数	家庭自用车	企业非营业用车	出租、租赁车	营业货车	党政机关、事业团体	
商车险	一级客户	0.891	0.891	0.891	0.8883	0.891
	二级客户	0.99	0.99	0.99	0.987	0.99
	三级客户	1.188	1.188	1.188	1.1844	1.188
	四级客户	1.584	1.584	1.584	1.5792	1.584
	五级客户	2.475	2.475	2.475	2.4675	2.475
交强险	一级客户	0.882	0.891	0.873	0.846	0.855
	二级客户	0.98	0.99	0.97	0.94	0.95
	三级客户	1.078	1.089	1.067	1.034	1.045

5. 建立决策树

决策树是一种树形结构,上面的每一个节点就代表对其属性上的一个判断,树上的每个分支就是一个判断结果的输出。用决策树方法来评价一个项目的风险,易于理解,结果更加直观,更容易解释方案是否可行[11][12]。根据[13][14]

$$\text{出险车辆数比例} A = \frac{\text{出险的车辆数}}{\text{投保的车辆数}}$$

$$\text{赔付款占投保车辆浮动前保费总额比例} C = \frac{\text{赔付款总额}}{\text{投保车辆浮动前保费总额}}$$

$$\text{未决比例} D = \frac{\text{未决件数}}{\text{立案件数}},$$

对使用性质不同的车辆进行以下赋权值:

	家庭自用车	企业非营业用车	出租、租赁车	营业货车	党政机关、事业团体
权 p_i	7	3	1	4	1

同时,再对纯收入、风险投保以及拖延理赔赋权值:

	纯收入	风险投保	拖延理赔
权 p_j	7	5	3

经过分析得到决策树模型如图 2 所示。

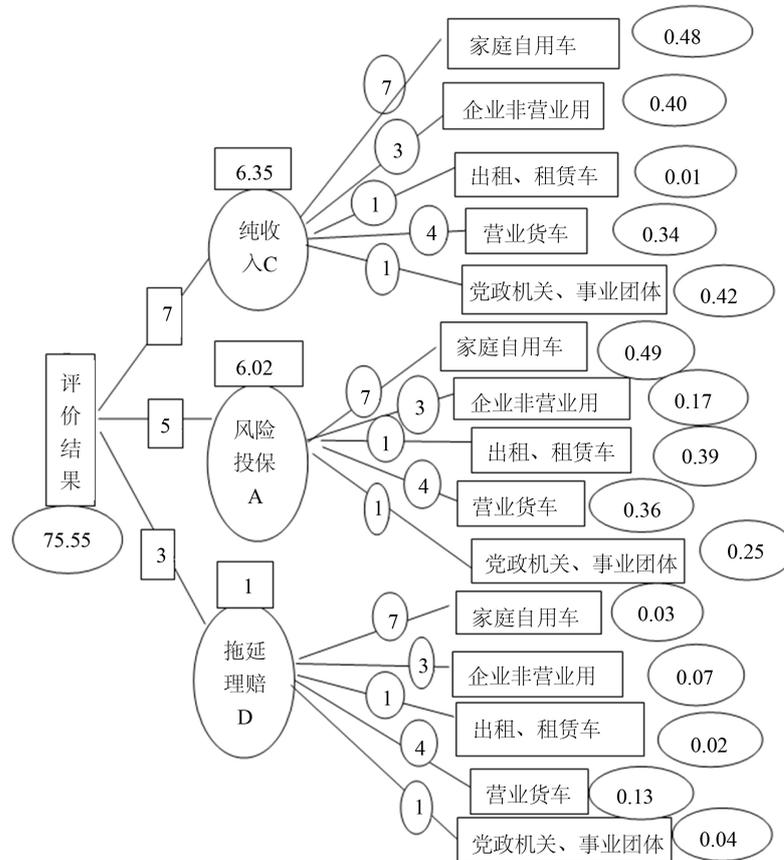


Figure 2. The analysis chart of decision tree
图 2. 决策树分析图

已知方案的期望收益就是把未来收益各种可能结果出现的概率对了解到的估计值做加权平均。分公司得分 $S = \sum p_j (\sum p_i A_i + 5 \sum p_i C_i + \sum p_i D_i) = 75.55$ ，其中 A、C、D 的上极限为 1，将上极限值带入公式得到 270，可以得到总公司对分公司的风险评估分数区间为(0, 270)，根据公司评价表(见附录)可知，参考数据中的汽车保险公司一般[15]。

6. 结束语

随着科技的发展汽车使用率只增不减，越来越多的人关注汽车保险，其在保险行业的地位与日俱增，合理的评估续保率，对于保险公司有具有重大意义。通过选取重要影响因素，本文建立影响力指数的等级比较模型，说明了各因素对续保率的影响。而后利用灰色预测法进行灰色关联度预测，计算出使用性质系数 B，并得到三种使用性质系数 B 的方案，经分析第二种方案更加合理。最后，通过三个因子赋权构建决策树模型，数据结果表明，该方案具有合理性，值得推广。

基金项目

海南省自然科学基金(120RC451), 国家自然科学基金(11761025, 11961018, 11901114), 广东省教育厅青年创新人才类(2017KQNCX081), 广州市科技创新一般项目(201904010010), 中山大学广东省计算科学重点实验室开放课题基金资助(2018001), 海南省研究生创新科研课题项目(Hys2020-108)。

参考文献

- [1] 倪琪, 刘骅飞, 田雪颖. 车险续保率影响因素模型[J]. 企业研究, 2011(5): 112-113.
- [2] 陈曼. 影响车险客户续保因素刍议[J]. 上海保险, 2012(8): 34-36.
- [3] 刘亮. 一种组合模型在车险续保率中的应用[J]. 数字技术与应用, 2019, 37(7): 106+108.
- [4] 薛恒. 商业车险费率改革对保险公司的影响及应对[J]. 保险职业学院学报, 2016, 30(5): 23-25.
- [5] 王梦晨. A 公司车险业务续保率影响因素研究[D]: [硕士学位论文]. 长沙: 湖南大学, 2017.
- [6] 杨孝良, 周猛, 曾波. 灰色预测模型背景值构造的新方法[J]. 统计与决策, 2018, 34(19): 14-18.
- [7] 李翠凤. 灰色系统建模理论及应用[D]: [硕士学位论文]. 杭州: 浙江工商大学, 2006.
- [8] 张永建, 余挺. 车险续保模型的构建及实例[N]. 中国保险报, 2014-11-26(002).
- [9] 张永霞, 孟生旺. 我国商业车险奖惩系统研究[J]. 保险研究, 2016(10): 3-15.
- [10] 郑丹丹. 中小财险公司应对车险费率市场化改革对策[D]: [硕士学位论文]. 沈阳: 辽宁大学, 2017.
- [11] 张梅. 决策树在风险决策中的应用[J]. 高师理科学刊, 2013, 33(2): 41-43.
- [12] 朱建平. 经济预测与决策[M]. 厦门: 厦门大学出版社, 2012.
- [13] 王彘. 车险赔付风险影响因素决策树分析[J]. 价值工程, 2010, 29(5): 248-249.
- [14] 叶慧珍. 中小型产险公司车险索赔风险影响因素研究——以 M 保险公司为例[D]: [硕士学位论文]. 杭州: 浙江工商大学, 2016.
- [15] 曾媛媛. 基于决策树和支持向量机在公司财务风险评估中的应用[J]. 科技经济市场, 2013(9): 17-19.

附录

1: 对家庭自用车对续保率的影响进行灰色预测

家庭自用车续保率表

数据 \ 月份	9	10	11	12	1	2	3
年到期车辆续保率	25.81%	26.53%	26.48%	27.05%	35.68%	36.81%	36.37%

1) 建立原始序列

由于原始数据序列 $X^{(0)}$ 为非负序列, 则

$$X^{(0)} = \{X^{(0)}(1), X^{(0)}(2), \dots, X^{(0)}(n)\}$$

其中, $x^{(0)}(k) \geq 0, k=1, 2, \dots, n$

由表格可知, 时间序列 $X^{(0)}$ 有 7 个观察值, 即 2010 年 9 月至 2011 年 3 月, 用 1 到 7 代表, 则

$$\begin{aligned} X^{(0)} &= \{X^{(0)}(1), X^{(0)}(2), \dots, X^{(0)}(7)\} \\ &= \{25.81\%, 26.53\%, 26.48\%, 27.05\%, 35.68\%, 36.81\%, 36.37\%\} \end{aligned}$$

2) 对 $X^{(0)}$ 作 1-AGO

D 为 $X^{(0)}$ 的通过累加生成新数据序列为 $X^{(1)}$

$$\begin{aligned} X^{(1)} &= \{X^{(1)}(1), X^{(1)}(2), \dots, X^{(1)}(7)\} \\ &= \{25.81\%, 52.34\%, 78.82\%, 105.87\%, 141.55\%, 178.36\%, 214.73\%\} \end{aligned}$$

则 GM(1, 1) 模型相应的微分方程为:

$$\frac{dX^{(1)}}{dt} + aX^{(1)} = \mu$$

其中 a, μ 是模型的参数。

3) 对 $X^{(1)}$ 作紧邻均值生成, 令

$$Z^{(1)}(k) = 0.5x^{(1)}(k) + 0.5x^{(1)}(k-1), k=0, 1, 2, \dots, 7$$

$$\begin{aligned} Z^{(1)} &= \{z^{(1)}(1), z^{(1)}(2), z^{(1)}(3), z^{(1)}(4), z^{(1)}(5), z(6), z(7)\} \\ &= \{25.81\%, 39.08\%, 65.58\%, 92.34\%, 123.71\%, 159.95\%, 196.54\%\} \end{aligned}$$

于是,

$$B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ -z^{(1)}(4) & 1 \\ -z^{(1)}(5) & 1 \\ -z^{(1)}(6) & 1 \\ -z^{(1)}(7) & 1 \end{bmatrix} = \begin{bmatrix} -39.08\% & 1 \\ -65.58\% & 1 \\ -92.34\% & 1 \\ -123.71\% & 1 \\ -159.95\% & 1 \\ -196.54\% & 1 \end{bmatrix}, Y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ x^{(0)}(4) \\ x^{(0)}(5) \\ x^{(0)}(6) \\ x^{(0)}(7) \end{bmatrix} = \begin{bmatrix} 26.53\% \\ 26.48\% \\ 27.05\% \\ 35.68\% \\ 36.81\% \\ 36.37\% \end{bmatrix}$$

$$\begin{aligned}
 B^T B &= \begin{bmatrix} -39.08\% & -65.58\% & -92.34\% & -123.71\% & -159.95\% & -196.54\% \\ 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \times \begin{bmatrix} -39.08\% & 1 \\ -65.58\% & 1 \\ -92.34\% & 1 \\ -123.71\% & 1 \\ -159.95\% & 1 \\ -196.54\% & 1 \end{bmatrix} \\
 &= \begin{bmatrix} 9.387 & -6.772 \\ -6.772 & 6 \end{bmatrix} \\
 (B^T B)^{-1} &= \begin{bmatrix} 9.387 & -6.772 \\ -6.772 & 6 \end{bmatrix}^{-1} = \frac{1}{9.387 \times 6 - 6.772^2} \begin{bmatrix} 6 & 6.772 \\ 6.772 & 9.387 \end{bmatrix} = \begin{bmatrix} 0.5736 & 0.6473 \\ 0.6473 & 0.8972 \end{bmatrix}
 \end{aligned}$$

设 $\hat{\alpha}$ 为待估参数向量, $\hat{\alpha} = \begin{pmatrix} a \\ \mu \end{pmatrix}$, 可利用最小二乘法求解。解得:

$$\begin{aligned}
 \hat{\alpha} &= (B^T B)^{-1} B^T Y_n \\
 &= \begin{bmatrix} 0.5736 & 0.6473 \\ 0.6473 & 0.8972 \end{bmatrix} \times \begin{bmatrix} -0.3908 & -0.6558 & -0.9234 & -1.2371 & -1.5995 & -1.9654 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \times \begin{bmatrix} 0.2653 \\ 0.2648 \\ 0.2705 \\ 0.3568 \\ 0.3681 \\ 0.3637 \end{bmatrix} \\
 &= \begin{bmatrix} 0.4231 & 0.2711 & 0.1176 & -0.623 & -0.2702 & -0.4801 \\ 0.6442 & 0.4727 & 0.2995 & 0.0964 & -0.1382 & -0.375 \end{bmatrix} \times \begin{bmatrix} 0.2653 \\ 0.2648 \\ 0.2705 \\ 0.3568 \\ 0.3681 \\ 0.3637 \end{bmatrix} \\
 &= \begin{bmatrix} -0.0802 \\ 0.2244 \end{bmatrix}
 \end{aligned}$$

4) 确定模型

因此, GM(1,1)模型的参数 $a = -0.0802, \mu = 0.2244$, 则其微分方程为:

$$\frac{dX^{(1)}}{dt} - 0.0802, X^{(1)} = 0.2244$$

GM(1,1)灰色微分方程的时间响应序列为

$$\begin{aligned}
 \hat{x}^{(1)}(k+1) &= \left(x^{(0)}(1) - \frac{\mu}{a} \right) e^{-ak} + \frac{\mu}{a} \\
 &= 3.0561e^{0.0802k} - 2.798
 \end{aligned}$$

5) 求 $X^{(1)}$ 的模拟值

$$\begin{aligned}
 \hat{X}^{(1)} &= \{ \hat{x}^{(1)}(1), \hat{x}^{(1)}(2), \hat{x}^{(1)}(3), \hat{x}^{(1)}(4), \hat{x}^{(1)}(5), \hat{x}^{(1)}(6), \hat{x}^{(1)}(7) \} \\
 &= \{ 0.2581, 0.5133, 0.7897, 1.0892, 1.4183, 1.7654, 2.1464 \}
 \end{aligned}$$

还原出 $X^{(0)}$ 的模拟值, 由

$$\hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k)$$

得

$$\begin{aligned} \hat{X}^{(0)} &= \{\hat{x}^{(0)}(1), \hat{x}^{(0)}(2), \hat{x}^{(0)}(3), \hat{x}^{(0)}(4), \hat{x}^{(0)}(5), \hat{x}^{(0)}(6), \hat{x}^{(0)}(7)\} \\ &= \{0.2581, 0.2552, 0.2765, 0.2995, 0.3245, 0.3516, 0.3810\} \end{aligned}$$

6) 误差检验

Table A1. Table of residual test

附表 A1. 残差检验表

序号	实际数据	模拟数据	残差
	$x^{(0)}(k)$	$\hat{x}^{(0)}(k)$	$\varepsilon(k) = x^{(0)}(k) - \hat{x}^{(0)}(k)$
2	0.2653	0.2552	0.0101
3	0.2648	0.2765	-0.0117
4	0.2705	0.2995	-0.0290
5	0.3568	0.3245	0.0323
6	0.3681	0.3516	0.0165
7	0.3637	0.3810	-0.0173

残差平方和

$$\begin{aligned} s &= \varepsilon^T \varepsilon = [\varepsilon(2) \ \varepsilon(3) \ \varepsilon(4) \ \varepsilon(5) \ \varepsilon(6) \ \varepsilon(7)] * \begin{bmatrix} \varepsilon(2) \\ \varepsilon(3) \\ \varepsilon(4) \\ \varepsilon(5) \\ \varepsilon(6) \\ \varepsilon(7) \end{bmatrix} \\ &= [0.0101 \ -0.0117 \ -0.0290 \ 0.0323 \ 0.0165 \ -0.0173] \times \begin{bmatrix} 0.0101 \\ -0.117 \\ -0.0290 \\ 0.0323 \\ 0.0165 \\ -0.0173 \end{bmatrix} = 0.0027 \end{aligned}$$

平均相对误差

$$\Delta = \frac{1}{6} \sum_{k=1}^6 \Delta_k = \frac{1}{6} (3.82\% + 4.40\% + 10.73\% + 9.04\% + 4.47\% + 4.75\%) = 6.20\%$$

计算 X 与 \hat{X} 的灰色关联度

$$\begin{aligned}
|S| &= \left| \sum_{k=2}^6 (x(k) - x(1)) + \frac{1}{2}(x(7) - x(1)) \right| \\
&= \left| (0.2653 - 0.2581) + (0.2648 - 0.2581) + (0.2705 - 0.2581) \right. \\
&\quad \left. + (0.3568 - 0.2581) + (0.3681 - 0.2581) + \frac{1}{2} \times (0.3637 - 0.2581) \right| \\
&= 0.6574 \\
|\hat{S}| &= \left| \sum_{k=2}^6 (\hat{x}(k) - \hat{x}(1)) + \frac{1}{2}(\hat{x}(7) - \hat{x}(1)) \right| = 0.7084 \\
|\hat{S} - S| &= \left| \sum_{k=2}^6 [(x(k) - x(1)) - (\hat{x}(k) - \hat{x}(1))] + \frac{1}{2}[(x(7) - x(1)) - (\hat{x}(7) - \hat{x}(1))] \right| = 0.0510 \\
\varepsilon &= \frac{1 + |S| + |\hat{S}|}{1 + |S| + |\hat{S}| + |\hat{S} - S|} = \frac{1 + 0.6574 + 0.7084}{1 + 0.6574 + 0.7084 + 0.0510} = 0.9789 > 0.9
\end{aligned}$$

Table A2. Referenced table of the precision test ranks**附表 A2.** 精度检验等级参照表

精度等级	指标临界性	相对误差	关联度	均方差比值	小误差概率
一级		0.01	0.90	0.35	0.95
二级		0.05	0.80	0.50	0.80
三级		0.10	0.70	0.65	0.70
四级		0.20	0.60	0.80	0.60

Table A3. Table of company evaluation**附表 A3.** 公司评价表

评价	优秀	较好	一般	较差	极差
得分	0-27	28~54	55~81	82~108	109 以上