

#1051

2011 年第四届“互动出版杯”数学中国 数学建模网络挑战赛

题 目 你的爱车入保险了吗？

关 键 词 保费浮动率 决策树 续保率 灰色预测模型

摘 要：

本文运用灰色预测法和最大期望原则下用决策树解决了汽车保险保费浮动和保险公司业绩考核问题在问题。

问题 1 中，首先通过灰色预测法对不同使用性质的车辆的续保率的影响进行灰色关联度预测，证明使用性质的不同对续保率的影响较大。

然后，统计数据，根据车辆出险比例、赔付款占浮动前保费总额的比例、赔付款占出险车辆浮动前保费的比例，计算出保费浮动系数，提出三种保费浮动方案，经分析，推荐使用保费浮动系数方案如下：

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	0.99	0.99	0.03	0.987	0.99
交强险	0.98	0.99	0.98	0.963	0.996

对于问题 2，首先根据要求建立决策树，结合问题 1 统计的数据，利用不同使用性质的车辆保户数占总保户数的百分比进行赋权值，得到决策似累加模型：

$$S = \sum p_j (\sum p_i A_i + \sum p_i C_i + \sum p_i D_i)$$

其次计算该公司的得分为 75.78 分，根据公司评价表对其评价为一般。

然后通过对模型的深入分析对该公司今后的风险控制提出相关建议。

最后，对模型进行推广与评价。

参赛队号 1051

所选题目 C

参赛密码 _____
(由组委会填写)

#1051

英文摘要（选填）

This paper USES grey forecasting method and decision tree solve auto insurance premium floating and insurance companies in the performance evaluation problem. Question 1, first by grey forecasting method to use the vehicles different influence on attachment rate gray associationr-prediction, proof of use of the different nature of the influence of attachment rate is bigger.

Then, the statistics data, according to the proportion of vehicles be or get out of danger, PeiFuKuan up before PeiFuKuan proportion of the total premium of be or get out of danger, the proportion of vehicles, up before premium calculated premium floating coefficient, this article proposes three premium floating scheme, classics analysis, recommend using premium floating coefficient plan is as follows:

Premium floating coefficient family since the non-operating transport enterprises transport, car rental and leasing business truck in party and government organs, business groups

Business 0.987 0.99 0.99 0.99 0.03 insurers
0.996 0.963 0.98 0.99 0.98 vehicle

To question 2, first of all, according to the request to establish decision tree, combined with problem 1 statistical data, using different vehicles will enable the percentage of the total number of insured for weighting, get decision like accumulate model:

Secondly computed the company's score for 75.78 points, according to the company for the evaluation PingJiaBiao for general.

Then through thorough analysis of the model of the company's further risk control related Suggestions.

Finally, promotion and evaluation model.

#1051

一、问题重述

问题背景：

近年来，国内汽车销售市场非常火爆，销售量屡创新高。汽车保险，简称车险，是指对机动车辆由于自然灾害或意外事故所造成的人身伤亡或财产损失负赔偿责任的一种商业保险。汽车保险是财产保险中的主要险种。我国首个由国家法律规定实行的强制保险制度交强险，从2006年7月1日起正式实施。交强险，全称机动车交通事故责任强制保险。交强险的基本定义是：交强险是由保险公司对被保险机动车发生道路交通事故造成受害人（不包括本车人员和被保险人）的人身伤亡、财产损失，在责任限额内予以赔偿的强制性责任保险。除交强险外，各个保险公司还有很多种类的自己的商业车险产品。

在我国保险业，汽车保险地位难以撼动。相对财产保险公司而言，有得车险得天下之说。连续多年，汽车保险稳居国内产险业第一大险种。

第二阶段问题：

问题 1：

汽车保险公司为了降低车辆出险率，鼓励保户续保，发展潜在保户，通常都会对满足一定要求的保户或者投保人给与一定比例的保费浮动优惠，就是通常所说的保费折扣。请根据附件中的参考数据，以及第一阶段中对于影响续保率因素的分析，给出一套较为合理的保费浮动方案。

问题 2：

一些大型的保险公司要在全中国很多地区设立分公司。总公司每年要对分公司的业绩情况进行考核，考核结果直接影响分公司领导班子的去留。传统的考核方法就是计算分公司的保费收入和理赔支出的差额。一些分公司为了提高自己的考核成绩，会使用受理一些风险较大的投保或者故意拖延理赔的处理时间等方法。因此，很多保险公司开始考虑引入风险评估机制来对分公司进行考核，潜在风险较低的分公司会得到较高的考核成绩，请建立合理的模型对参考数据中的汽车保险公司进行潜在风险的评估，并通过对模型的深入分析对该公司今后的风险控制提出建议。

二、模型假设与符号说明

模型假设：

1. 假设题中所给数据真实、可靠、有效；
2. 假设总体经济基本稳定，保险新政策不会给保险业带来太大影响；
3. 假设影响保险业（续保、浮动）只考虑题给的因素，忽略保险公司其他影响因素；
4. 假设将车险仅分为商业险和交强险两类；

符号说明：

a ：发展灰数

μ ：内生控制灰数

$X^{(0)}$ ：原始数据列

$X^{(1)}$ ：累加生成列

B ：数据矩阵

Y ：常数矩阵

$\hat{X}^{(0)}$ ：相应的模拟误差序列

$\varepsilon^{(0)}$ ：残差序列

Δ_k ： k 点相当对误差

#1051

Δ : 相对误差序列

$\bar{\Delta}$: 平均模拟相对误差

$\eta(k)$: 关联系数

ρ : 分辨率

r : 关联度

k : 1, 2, 3, 4, 5, 6, 7, 8, 9

\bar{x} : 均值

$\bar{\varepsilon}$: 残差均值

s : 残差方差

c : 均方差比值

A : 车辆出险次数比例

B : 赔付款总额占出险车辆浮动前保费总额比例

C : 赔付款总额占投保车辆浮动前保费总额比例

D : 未决件数与立案件数的比例

三、问题分析

问题 1 分析:

由题意可知, 汽车保险公司为了鼓励保户续保和降低出险车辆在所有投保车辆中所占的比例, 即车辆的出险率, 因此给予保户一定的保费浮动优惠, 在第一阶段问题 1 的分析和求解可得车险续保率涉及到了承保车辆的使用性质, 及承保车辆的出险次数等因素。其因素之间的发展趋势不确定, 因素数据离散。因此, 可以用灰色预测法对影响车险续保率的因素进行研究。并建立其灰色预测模型。

首先根据第一阶段问题 1 的求解, 用灰色预测法评估车的使用性质对续保率的影响。

灰色预测通过鉴别系统因素之间发展趋势的相异程度, 即进行关联分析, 并对原始数据进行生成处理来寻找系统变动的规律, 生成有较强规律性的数据序列, 然后建立相应的微分方程模型, 进而说明影响续保率的因素。

其中, 承保车辆的使用性质包含: 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、特种车等等。

其次结合保费浮动相关数据进行分析, 保户的保额越小, 出险次数越多, 总赔付款越多, 公司的盈利就越小。针对某一种使用性质的车辆, 统计出浮动前保费总额、赔付款总额、投保车辆数、各出险次数的车辆数, 计算车辆出险比率、赔付款占浮动前保费总额的比率、赔付款占出险车辆浮动前保费的比率, 提出保费浮动方案。

问题 2 分析:

在问题 1 的基础上进一步对数据进行处理, 找出影响分公司的业绩考核因素, 可知影响因素有: 公司每年的总收入、风险投保、故障拖延。引用决策树法来考核分公司的业绩情况, 根据题目所给的数据和主观意愿对影响因素赋值概率, 再对每个因素进行细分。由于每年的考核因素又由车的使用性质(家庭自用, 营业货车, 企业非营业用车, 党政机关、事业团体, 出租、租赁)决定, 对数据进行处理, 可以得到不同使用性质的车辆所占的保费总额、风险投保、故障拖延的比例, 由每个车的使用性占的比例数据和量化的数据的和, 可求每个因素的的比重情况。

最后根据结果可以评估哪个因素对公司的投资存在潜在的风险大, 哪个因素对公司潜在风险小, 哪个分公司领导班子的去留情况, 并结合现实社会实际情况对公司今后的风险控制提出建议。

#1051

四、模型的建立、求解以及检验

模型一：

I、灰色预测模型的建立、求解及检验

首先对承保车辆的使用性质不同形式进行分析求解。

在第一阶段的第一问中，得到以下模型：

根据题意，由续保率的数据可知，家庭自用、企业非营业货车、出租租赁车、营业货车、党政机关、事业团体车这几种的车辆在当年到期车辆数和当年到期车辆续保率都是较多的，当年到期的目标客户车辆数和目标客户续保率也较多。而其它几种则较少，因此，我们对当年到期车辆数较多的承保车辆使用性质的几种进行灰色预测。

首先，对家庭自用车对续保率的影响进行灰色预测。

由2010年9月至2011年3月这七个月的当年到期车辆续保率的原始数据组成的新数据。入下表所示：

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%

灰色系统理论的微分方程成为Gm模型，G表示gray（灰色），m表示model（模型），Gm（1，1）表示1阶的、1个变量的微分方程模型。

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

有表格可知，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。

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

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。

对 $X^{(1)}$ 作紧邻均值生成，令

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

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

于是，

#1051

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

$$B^T B = \begin{bmatrix} -39.08\% & -65.58\% & -92.34\% & -123.71\% & -159.95\% & -196.54\% \\ 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}^* \begin{bmatrix} -39.08\% & 1 \\ -65.58\% & 1 \\ -92.34\% & 1 \\ -123.71\% & 1 \\ -159.95\% & 1 \\ -196.54\% & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 9.387 & -6.772 \\ -6.772 & 6 \end{bmatrix}$$

$$(B^T B)^{-1} = \begin{bmatrix} 9.387 & -6.772 \\ -6.772 & 6 \end{bmatrix}^{-1} = \frac{1}{9.387 \times 6 - 6.772^2} \begin{bmatrix} 6 & 6.772 \\ 6.772 & 9.387 \end{bmatrix} = \begin{bmatrix} 0.5736 & 0.6473 \\ 0.6473 & 0.8972 \end{bmatrix}$$

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:

$$\hat{\alpha} = (B^T B)^{-1} B^T Y_n$$

$$= \begin{bmatrix} 0.5736 & 0.6473 \\ 0.6473 & 0.8972 \end{bmatrix}^* \begin{bmatrix} -0.3908 & -0.6558 & -0.9234 & -1.2371 & -1.5995 & -1.9654 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}^* \begin{bmatrix} 0.2653 \\ 0.2648 \\ 0.2705 \\ 0.3568 \\ 0.3681 \\ 0.3637 \end{bmatrix}$$

$$= \begin{bmatrix} 0.4231 & 0.2711 & 0.1176 & -0.623 & -0.2702 & -0.4801 \\ 0.6442 & 0.4727 & 0.2995 & 0.0964 & -0.1382 & -0.375 \end{bmatrix}^* \begin{bmatrix} 0.2653 \\ 0.2648 \\ 0.2705 \\ 0.3568 \\ 0.3681 \\ 0.3637 \end{bmatrix}$$

$$= \begin{bmatrix} -0.0802 \\ 0.2244 \end{bmatrix}$$

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) 灰微分方程的时间响应序列为

$$\hat{x}^{(1)}(k+1) = (x^{(0)}(1) - \frac{\mu}{a})e^{-ak} + \frac{\mu}{a}$$

#1051

$$= 3.0561e^{0.0802k} - 2.798$$

则 $X^{(1)}$ 的模拟值为：

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

还原出 $X^{(0)}$ 的模拟值，由

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

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

误差检验

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

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

平均相对误差

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

计算 X 与 \hat{X} 的灰色关联度

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

#1051

$$\begin{aligned}
&= \left| \begin{aligned} &(0.2653 - 0.2581) + (0.2648 - 0.2581) + (0.2705 - 0.2581) \\ &+ (0.3568 - 0.2581) + (0.3681 - 0.2581) + \frac{1}{2} * (0.3637 - 0.2581) \end{aligned} \right| \\
&= 0.6574 \\
|\hat{S}| &= \left| \sum_{k=2}^6 (\hat{x}(k) - \hat{x}(1)) + \frac{1}{2} (\hat{x}(7) - \hat{x}(1)) \right| \\
&= 0.7084 \\
|\hat{S} - S| &= \left| \sum_{k=2}^6 [(x(k) - x(1)) - (\hat{x}(k) - \hat{x}(1))] + \frac{1}{2} [(x(7) - x(1)) - (\hat{x}(7) - \hat{x}(1))] \right| \\
&= 0.0510 \\
\varepsilon &= \frac{1 + |S| + |\hat{S}|}{1 + |S| + |\hat{S}| + |\hat{S} - S|} = \frac{1 + 0.6574 + 0.7084}{1 + 0.6574 + 0.7084 + 0.0510} \\
&= 0.9789 > 0.9 \text{ 参照精度检验等级参照表（见附录 1）。}
\end{aligned}$$

因此可知，家庭自用影响续保率的精度为一级，可以用。说明了承保车辆的使用性质中的家庭自用对续保率的影响显著。

同理，应用灰色预测法分别对党政机关客车、企业客车、非营业货车、出租租赁、营业货车、特种车等几种车进行续保率的精度检验。由于这几种车险续保率精度检验方法雷同，我们就不一一叙述，利用续保率的数据，根据公式，编辑程序先确定模型的参数 $a, \mu(b)$ （程序见附录 2.1、2.2），再确立其模型方程，算出模拟数据，计算其车辆使用性质统计的实际数据与模拟数据的相对误差、平均相对误差、残差平方和、灰色关联度。

结果如下表 1：

	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

由于党政机关客车、企业客车、非营业货车、出租租赁、营业货车、特种车的灰色关联度：0.9876 > 0.9, 0.9672 > 0.9, 0.9623 > 0.9, 0.9059 > 0.9, 0.9815 > 0.9, 0.9585 > 0.9 都大于 0.9，故它们的续保率的精度为一级，可以用。其对续保率的影响显著。

II、浮动方案的给出

先根据数据把商车险和交强险分开，分别统计两个险别中车的使用性（家庭自用；企业非营业用车；出租、租赁；营业货车；城市公交；党政机关、事业团体）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。

不管是在商车险还是在交强险中，对数据进行处理可知：特种车在发生事故机遇好少，因此对她的影响力不大，所占的比例小，可以不考虑。

第一步①统计保户投保商车险的投保车辆数得到下表 2

（百分比是该种车辆数占商车险总投保车辆的比例）

#1051

	家庭自用	企业非营业用车	出租、租赁车	营业货车	城市公交	党政机关、事业团体
投保车辆数	1877	519	104	1181	2	283
百分比	88.46%	22.88%	3.05%	34.59%	0.06%	8.29%

②在商车险中，非营业货车、与城市公交投保车辆非常少，不予以考虑。

统计交强险的投保车辆数得到下表 3：

（百分比是该种车辆数占交强险总投保车辆的比例）

	家庭自用	企业非营业用车	出租、租赁车	营业货车	城市公交	党政机关、事业团体
投保车辆数	1143	780	35	436	1	130
百分比	58.57%	10.47%	3.34%	9.28%	0.02%	2.96%

在交强险中，城市公交投保车辆非常少，不予以考虑。

第二步

保户的保额越小，出险次数越多，总赔付款越多，公司的盈利就越小。针对不同使用性质的车辆，统计出浮动前保费总额、赔付总额、投保车辆数、出险的车辆数。

商车险的有关统计数据如下表 4：

	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	1877	519	104	1181	283
赔付款总额(万)	402.67	72.13	571	169.62	7.91
出险车辆数(辆)	917	139	42	405	58
出险车辆总保额(万)	35610.64	7055.62	588	13467.28	1672.49
浮动前保费总额(万)	60318.01	23611.46	1005.00	29801.23	4312.13
A	0.49	0.27	0.40	0.34	0.20
B	0.01	0.01	0.97	0.01	0.005
C	0.007	0.003	0.568	0.006	0.002

交强险的有关统计数据如下表 5：

	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	1143	780	35	436	130
赔付款总额(万)	52.83	3.67	3.25	62.72	1.63
出险车辆数(辆)	285	42	15	139	19
出险车辆总保额(万)	3477	512.4	183	1695.8	439.2
浮动前保费总额(万)	13883.6	3172	427	5307	1586
A	0.25	0.05	0.43	0.32	0.15
B	0.02	0.01	0.02	0.04	0.004
C	0.004	0.001	0.008	0.012	0.001

由以上数据可以计算出：

#1051

- 出险车辆数比例 $A = \frac{\text{出险的车辆数}}{\text{投保的车辆数}}$
- 赔付款占出险车辆浮动前保费总额比例 $B = \frac{\text{赔付款总额}}{\text{出险车辆浮动前保费总额}}$
- 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

即得到下表 6：

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	方案一	0.51	0.73	0.60	0.66	0.51
	方案二	0.99	0.99	0.03	0.987	0.99
	方案三	0.993	0.997	0.432	0.994	0.993
交 强 险	方案一	0.75	0.95	0.57	0.68	0.85
	方案二	0.98	0.99	0.98	0.963	0.996
	方案三	0.996	0.999	0.992	0.988	0.999

III、方案比较

方案一是根据出险车辆数比例提出的，只考虑所投保的车辆是否会发生事故，而未考虑公司的保户个数、总保费、赔付款总额。该方案考虑不全面，不能广泛推广，但可以作为一种辅助参考。

方案二根据赔付款占出险车辆浮动前保费总额比例提出，充分考虑到出险的车辆赔付款和浮动前保费总额，适合于大部分保险公司，推荐采用此方案。

方案三由赔付款占投保车辆浮动前保费总额比例得到，从公司的利益出发，考虑了出险的车辆赔付款和投保车辆浮动前保费总额，但没涉及出险的车辆数。若一公司全年收入利益可观，但处理出险业务较多，会造成人力资源浪费，可见该方案一般，有一定的局限性。

模型二

I、决策树的介绍

(1) 算法说明：

决策树提供了一种展示类似在什么条件下会得到什么值这类规则的方法。决策树的基本组成部分：决策节点、分支和叶子。决策树学习是以实例为基础的归纳学习算法。它着眼于从一组无次序、无规则的事例中推理除决策树表示形式的分类规则。它采用自顶向下的递归方式，在决策树的内部节点进行属性值的比较并根据不同属性判断从该节点向下的分支，然后进行剪枝，最后在决策树的叶节点得到结论。所以从根到叶节点就对应着一条合取规则，整棵树就对应着一组析取表达式规则。基于决策树的分类有很多实现算法。ID3 和 C4.5 是较早提出并普遍使用的决策树算法。常用的算法有 CHAID、

#1051

CART、Quest 和 C5.0。

(2) 适用条件：

用来处理分类问题，产生容易理解的规则。

(3) 优点：

- 1) 可以生成可以理解的规则；
- 2) 计算量相对来说不是很大；
- 3) 可以处理多种数据类型；
- 4) 决策树可以清晰的显示哪些字段较重要；
- 5) 可以处理有缺失值的数据；
- 6) 不容易受极值的影响。

(4) 缺点：

- 1) 对连续性的字段比较难预测。
- 2) 有时间顺序的数据，要很多预处理工作。
- 3) 当类别太多时，错误可能就会增加较快。
- 4) 一般的算法分类的时候，只是根据一个字段来分类。
- 5) 不是全局最优。

II、决策树与模型的建立

(一)决策

决策树是用二叉树形图来表示处理逻辑的一种工具。可以直观、清晰地表达加工的逻辑要求。特别适合于判断因素比较少、逻辑组合关系不复杂的情况。

决策树提供了一种展示类似在什么条件下会得到什么值这类规则的方法。

根据数据分析可到下面的决策树：

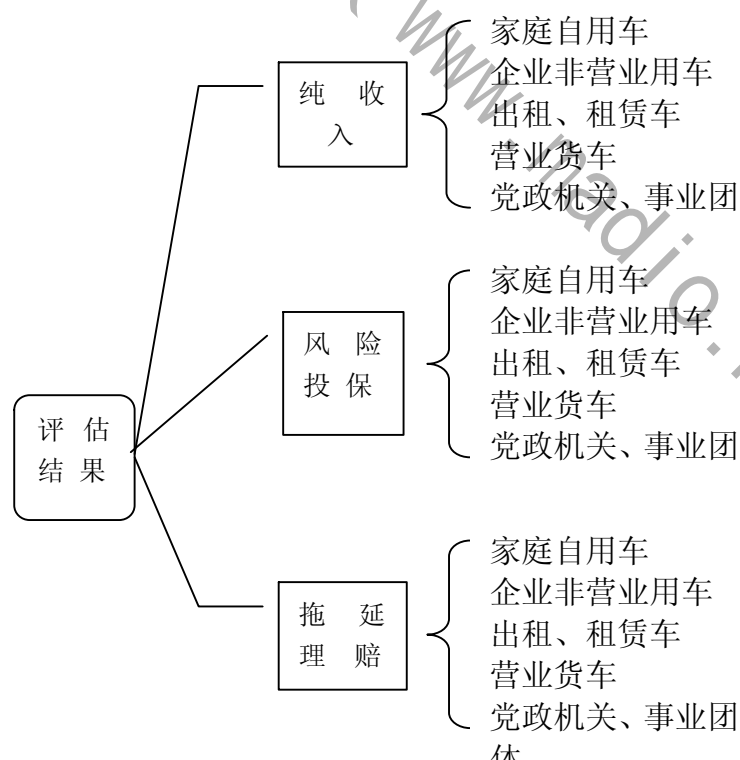


图 A 决策树

经过对问题的分析可知汽车保险公司的风险评估引用决策树法对分公司业绩情况进行考核。在决策树中，纯收入可用赔付款占投保车辆浮动前保费总额比例 C 来描述，

#1051

风险投保可用出险车辆数比例 A 描述，拖延理赔可用未决比例 D 描述。其中 C 、 A 、 D 分别是决策树中的状态结点，

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

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

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

同问题 1 一样用 Excel 表对数据统计，把投保车辆总数；车的使用性（家庭自用；企业非营业用车；出租、租赁车；营业货车；党政机关、事业团体车）的出险车辆数，得到数据进行比例。可得以下比例数据见表 7：

	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
A	0.40	0.24	0.43	0.35	0.22
C	0.42	0.36	0.12	0.27	0.17
D	0.03	0.07	0.02	0.13	0.04

(二)赋权

为了计算风险小最大期望收益原则，由问题 1 所统计的数据表 2 和表 3，再根据主观意愿给使用性质不同的车辆赋权值，家庭自用车所占的比率很大，赋权 7；出租、租赁车和营业货车所占比率大体上相持平，赋权 3；企业非营业用车所占比率稍大，赋权 4；而党政机关、事业团体所占比率极低，赋权 1，得到表 8。

其中 i = 家庭自用车，出租、租赁车，营业货车，党政机关，事业团体。

	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	4	3	3	1

同时，分别给纯收入、风险投保、拖延理赔赋权。

评估一个公司的好坏首先考虑它的纯收入，纯收入是基础的指标，赋权 7；保险公司受理一些风险较大的投保，虽暂时得到高利益，可能会造成以后的损失，可见风险投保是一个重要的指标，赋权 5；一些分公司为了提高自己的考核成绩，会故意拖延理赔的处理时间，但只能在很小的程度上提高考核成绩，赋权 3，总结后得到表 9。

其中 j = 纯收入，风险投保，拖延理赔。

	纯收入	风险投保	拖延理赔
权 p_j	7	5	3

(三)得出结论

一个方案下的期望收益就是各种可能状况发生的概率与对应的收益的加权平均。

分公司得分 $S = \sum p_j (\sum p_i A_i + \sum p_i C_i + \sum p_i D_i) = 75.78$ ，其中 A 、 C 、 D 最大值为 1，但永不为 1，将 1 带入公式得到 270，即总公司对分公司的风险评估分数在 0 到 270 之间。据公司评价表（见附录 3）可知，参考数据中的汽车保险公司一般。

(四)提出建议

保险风险来自于保险经营的方方面面、内部环境和外部环境。

该公司的大部分保户是为家庭自用车上的保险，而家庭自用车的出险率较高，造成

#1051

一定的风险。应大量发展出租车用户和党政机关、事业团体用户，降低公司的总出险率。

除了对发展用户的控制，还应完善企业内部控制，减少保险风险源的发生。保险公司为保障自身健康、稳定发展，在推出新的险种以及费率的厘定方面，要充分考虑新业务的各种风险源，从源头严格控制风险；在业务经营过程中，要严格执行内部控制制度，规范业务操作，避免或降低经营风险发生。

可以设施控制措施有效性评估：

1. 完全有效：控制措施对于降低风险水平是充分、有效，且能持续发挥作用的，不存在明显的缺陷，也没有改进的必要。

2. 基本有效：控制措施对于降低风险水平能力起到重要的作用，也能持续发挥作用，但存在少量明显的不足，需要改进。

3. 效果有限：控制措施存在一些明显的缺陷且需要改进，因此在降低风险水平上能起到的作用有限。不过已经制定并开始实施解决控制措施缺陷的行动计划。

4. 效果不明显：控制措施存在明显的缺陷有待改进，虽然制定了解决这些缺陷的行动计划，但还没有开始实施或处于实施的早期阶段。因此对于风险控制的效果不明显。

5. 无效：控制措施存在重大的缺陷以至于对风险控制完全无效，而且目前还没有相应的解决方案和行动计划。

五、模型的评价、推广

5.1 模型的评价

优点：

- (1) 建立的模型能与实际紧密联系，结合实际情况对所提出的问题进行求解，使模型更贴近实际。
- (2) 模型的建立中有成熟的理论基础和利用专业的MATLAB软件和Excel软件计算求解，方法简单实用，结果明确；可信度较高。
- (3) 对所涉及的重要参数进行了关联分析，通过利用数学工具，严格地对模型求解。
- (4) 模型中运用灰色预测的思想，对题中给出的各个因素都进行了量化评价与比较，结果可行性好。

缺点：

- (1) 模型综合考虑了很多因素，做了许多假设，具有一定的局限性。
- (2) 车的续保率因素较多，使得该模型不能将其全部准确的反映出来。
- (3) 计算和系统误差的存在，影响模型的准确性。

5.2 模型的推广

(1) 灰色预测模型适用于其它各个因素间联系的比照型的问题，如，人口模型、财政收入模型、能源、模型等等；

(2) 模型2还可以进一步深化解决更多因素，推广到经济计划和管理、能源政策和分配、公司的决策方案中去。

六、参考文献

- [1] 钱小军 主编《数量方法》高等教育出版社
- [2] 孙祥、徐流美主编《MATLAB基础教程》清华大学出版社
- [3] 李春刚 著《灰色预测法在商业连锁企业的应用》 经济师 2001
- [4] 姜启源 谢金星 叶俊 编《数学模型》高等教育出版社 2003
- [5] 吴建国 主编《数学建模案例精编》中国水利水电出版社 2005

#1051

七、附录

附录 1:

精度检验等级参照表

指标临界性 精度等级	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

附录 2.1:

```
function lel(x0,a,b)
%x0代表原始数据
%检验,计算误差
x1(1)=x0(1);x00(1)=x0(1);%x00代表模拟数据
for k=1:6
    x1(k+1)=(x0(1)-b/a)*exp(-a*k)+b/a;%求x1的模拟值
    x00(k+1)=x1(k+1)-x1(k);%还原x0的模拟值
    e(k+1)=x0(k+1)-x00(k+1);%残差
    m(k)=abs(e(k+1)/x0(k+1));%相对误差
end
x1;
x00;
e;
m
```

附录 2.2:

```
function sheng(x0,x)
%计算a和b
z(1)=x(1);
for i=2:length(x)
    z(i)=0.5*x(i)+0.5*x(i-1);
    B(i-1)=z(i);
end
z=vpa(z,6);
m=length(B);
B=B';
B=[-B,ones(m,1)];%计算矩阵B
B=vpa(B,6);
for i=2:length(x0)
    Y(i-1)=x0(i);%计算矩阵Y
end
Y=Y';
a0=vpa(inv(B'*B)*B'*Y,7);
a=a0(1)
```

#1051

b=a0(2)

```
function ff4()
```

```
%计算按照车辆使用性质数据的累加矩阵
```

```
a=[ 0.2581    0.2653    0.2648    0.2705    0.3568    0.3681    0.3637
    0.3601    0.3645    0.3641    0.3636    0.4279    0.4373    0.4397
    0.4266    0.4350    0.4357    0.4389    0.5357    0.5629    0.5328
    0.1274    0.1434    0.1456    0.1503    0.3411    0.3600    0.3459
    0.3172    0.3229    0.3256    0.3328    0.7508    0.7026    0.5097
    0.7083    0.7072    0.7067    0.7080    0.9091    0.8636    0.7222
    0.2598    0.2664    0.2636    0.2704    0.2059    0.2000    0.3200
    0.1759    0.1781    0.1732    0.1742    0.3026    0.3364    0.3455
    0.3746    0.3736    0.3703    0.3653    0.4272    0.4000    0.4840 ];
```

```
s=0;
```

```
[m,n]=size(a);
```

```
for i=1:m
```

```
    for j=1:n
```

```
        b(i,j)=a(i,j)+s;
```

```
        s=a(i,j)+s;
```

```
    end
```

```
    s=0;
```

```
end
```

```
b
```

附录 3:

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