

基于外部环境因素影响的高速公路 交通事故聚类分析*

周小勇¹, 宁建标², 张兵^{2,3}, 邹玉霞⁴

(1.江西交通咨询公司, 江西 南昌 330008; 2.华东交通大学 交通运输与物流学院, 江西 南昌 330013;

3.江西省交通科学研究院, 江西 南昌 330200; 4.江西省交通规划勘察设计院, 江西 南昌 330003)

摘要: 根据大广(大庆—广州)高速公路 K2952+350—K3044+450 段 2007—2012 年交通事故统计资料, 考虑事故严重程度, 对死亡事故和受伤事故赋予不同的权重系数 θ 换算成相同严重程度的不同事故次数, 综合考虑人、车、路和环境对交通事故的影响, 结合车况、路况及天气状况等外界影响因素, 采用模糊聚类法和系统聚类法对机动车司机不同违章驾驶行为的关联性进行分析。结果显示, 当截取水平 $\lambda \in (0.955\ 2, 0.973\ 6]$ 、聚类个数 $k=5$ 时, 两种聚类方法的聚类结果相同, 12 种常见违章驾驶行为聚为五类, 类间距与各违章驾驶行为引发的直接经济损失之差成正比, 符合实际数据统计情况。最后标定了各种违章驾驶行为在外界环境因素综合作用下引发交通事故的严重程度, 为进一步规范交通安全管理提供理论依据。

关键词: 交通安全; 交通事故; 高速公路; 驾驶行为; 外部环境因素; 聚类分析

中图分类号: U491.3

文献标志码: A

文章编号: 1671-2668(2017)04-0036-04

高速公路交通事故的发生与道路环境有关, 但主要还是人为因素造成的, 据统计, 90.3% 的交通事故和人为因素(违章驾驶行为、不良驾驶状态、生理、心理等)有关。交通事故作为和平时人员非正常死亡和财产意外损失的重要原因, 备受社会关注。很多学者对此作了深入研究: 马壮林等对高速公路隧道内交通事故严重程度的影响因素进行了分析, 从时间、环境和交通动态三方面选取 9 个变量研究其对交通事故严重程度的影响; 李金龙等考虑高速公路独特的行车环境, 综合人、车、路和环境的影响, 运用层次分析法对高速公路交通事故主要成因进行了分析; 文斌等考虑外部天气因素和道路状况对高速公路交通事故的影响, 得出了雨天事故率、雨天事故危险率与路面抗滑力之间的关系; 戴琪等分析了雾天高速公路交通事故的特征和成因, 并提出了相应交通管理和处置措施。大部分学者以交通事故发生率作为统计量, 结合驾驶行为、道路环境、车辆状况等分析交通事故特征及成因。该文以交通事故发生次数为统计量, 考虑事故的严重程度, 结合人、车、路和环境等因素, 采用模糊聚类法和系统聚类法对司机违章驾驶行为进行聚类分析, 为规范司机驾驶行为和降低道路交通事故提供理论依据。

1 外部环境影响系数及伤亡权值的确定

考虑人、车、路和环境综合作用下的交通事故, 研究不同违章驾驶行为在外界条件促进下造成交通事故的严重程度。

1.1 外界影响因素系数

(1) 天气状况。将天气状况等自然因素作为外界环境影响因素, 考虑雨天、雾天、冰冻等恶劣天气的影响。天气影响系数 η_1 为:

$$\eta_1 = \begin{cases} 1 & \text{(雨天、雾天、冰冻等恶劣天气)} \\ 0 & \text{(晴天、阴天等良好天气)} \end{cases}$$

(2) 道路状况。高速公路由于很少存在路面坑洼、路基破坏后仍允许通行等情况, 将道路线形作为外界道路主要影响因素。道路线形影响系数 η_2 为:

$$\eta_2 = \begin{cases} 1 & \text{(连续下坡、急弯、陡坡、弯陡坡等} \\ & \text{不良道路线形)} \\ 0 & \text{(平直、一般弯、一般坡等较良好的} \\ & \text{道路线形)} \end{cases}$$

(3) 车辆状况。将“驾驶机件不符合技术标准的机动车”及“驾驶安全设施不全的机动车”作为机动车司机的一项违章驾驶行为, 车辆状况只考虑车辆类型。车辆状况影响因素 η_3 为:

* 基金项目: 江西省交通运输厅科技项目(2015C0012; 2014R0011); 江西省重点研发计划项目(20161BBG70080)

$$\eta_3 = \begin{cases} 1 & \text{(挂车、牵引车、大型客车、重型货车等大型车辆)} \\ 0 & \text{(小型客车、轻型货车、中型客货车等中小型车辆)} \end{cases}$$

1.2 伤亡权值

高速公路发生交通事故的后果都较严重,而通常统计的事故数据资料中事故次数及事故率都无法体现事故的严重程度,如果只是简单地将事故次数进行累加,则会掩盖虽具有相同事故率但事故严重程度完全不同的事实。因此,在处理交通事故时必须考虑交通事故严重程度。在此通过赋予受伤及死亡事故一定的权重将实际事故次数转换成同种严重程度的事故次数。伤亡权值 θ_l 为:

$$\theta_l = \begin{cases} 2.0 & (l=1, \text{死亡事故}) \\ 1.5 & (l=2, \text{重伤事故}) \\ 1.2 & (l=3, \text{轻伤事故}) \end{cases}$$

2 数据处理

数据来源于大广(大庆—广州)高速公路 K2952+350—K3044+450 段 2007—2012 年的交通事故统计资料。将机动车司机违章驾驶行为作为样本 $a_i (i=1,2,\cdots,12)$,将车辆类型、道路线形和天气状况三类外界影响因素作为评价指标 $b_k (k=1,2,3)$,在 a_i 和 b_k 综合影响下的交通事故起数 x_{ik} 为:

$$x_{ik} = \sum_{l=1, r=1}^{l=3, r=3} \theta_l n_r + N_i \eta_k \tag{1}$$

$(i=1,2,\cdots,12; k=1,2,3)$

式中: n_r 为在 a_i 和 b_k 综合影响下交通事故的伤亡情况, $r=1$ 时代表死亡人数, $r=2$ 时代表受重伤人数, $r=3$ 时代表受轻伤人数; N_i 为调查统计期间第 i 种驾驶行为引发的交通事故起数。

2.1 原始数据统计分析

对原始交通事故统计数据进行处理,得到各种违章驾驶行为在外界因素影响下发生同等严重程度事故的起数(见表 1)及各违章驾驶行为引发交通事故导致的直接经济损失(见表 2)。

2.2 数据标准化

按式(2)对表 1 所示数据进行偏差转换,得到 x'_{ik} 。对 x'_{ik} 应用极差化法[见式(3)]进行处理,得到标准化数据 x''_{ik} (见表 3), $x''_{ik} \in [0,1]$ 。

$$\begin{cases} x'_{ik} = \frac{x_{ik} - \bar{x}_k}{\sqrt{S_{kk}}}, \bar{x}_k = \frac{1}{n} \sum_{i=1}^n x_{ik} \\ S_{kk} = \frac{1}{n-1} \sum_{i=1}^n (x_{ik} - \bar{x}_k)^2 \end{cases} \tag{2}$$

表 1 违章驾驶行为在外界因素影响下发生同等严重程度事故的起数 起

违章驾驶行为 a_i	外界因素影响下的事故数		
	天气状况	道路线形	车辆类型
超速行驶 a_1	405	225	138
疲劳驾驶 a_2	60	194	287
通过施工路段未减速行驶 a_3	4	14	20
违章变更车道 a_4	11	30	88
未保持安全行车距离 a_5	153	19	238
驾驶违章车辆 a_6	45	88	201
违章超车 a_7	15	9	14
违章驾驶 a_8	80	59	6
超载 a_9	54	42	16
无证驾驶 a_{10}	24	4	16
酒后驾驶 a_{11}	34	34	11
其他违章驾驶行为 a_{12}	53	48	22

表 2 违章驾驶行为引发交通事故导致的直接经济损失 万元

违章驾驶行为 a_i	经济损失	违章驾驶行为 a_i	经济损失
a_1	738	a_7	27
a_2	900	a_8	114
a_3	157	a_9	90
a_4	362	a_{10}	30
a_5	1 010	a_{11}	40
a_6	1 224	a_{12}	113

表 3 标准化数据

违章驾驶行为 a_i	外界因素影响下的事故数		
	天气状况	道路线形	车辆类型
a_1	1.000 0	1.000 0	0.469 8
a_2	0.139 6	0.859 7	1.000 0
a_3	0.000 0	0.045 2	0.049 8
a_4	0.017 4	0.117 6	0.291 8
a_5	0.371 6	0.067 9	0.825 6
a_6	0.102 2	0.380 1	0.694 0
a_7	0.027 4	0.022 6	0.028 5
a_8	0.189 5	0.248 8	0.000 0
a_9	0.124 7	0.171 9	0.035 6
a_{10}	0.049 9	0.000 0	0.035 6
a_{11}	0.074 8	0.135 7	0.017 8
a_{12}	0.122 2	0.199 1	0.056 9

$$x''_{ik} = \frac{x'_{ik} - x'_{ij\min}}{x'_{ik\max} - x'_{ik\min}} \tag{3}$$

3 聚类分析

采用系统聚类法和模糊聚类法对在外界环境因素作用下的机动车司机违章驾驶行为进行聚类。

3.1 系统聚类法

N 个样本,开始各自作为一类,即 N 类,并规定样本与样本之间的距离和类与类之间的距离,然后将距离最近的合并成一个新的类,计算新类与其他类的距离;重复进行两个最近类的合并,每次减少一类,直至所有样本合并成一大类。类与类之间距离的远近(相似度)由 Euclidean 距离计算:

$$d_{ij}^2 = \sum_{k=1}^3 (x_{ik} - x_{jk})^2$$

在常用的 6 种系统聚类法中,离差平方和法的类间距和两类样本数有较大关系,两个大类倾向于有较大距离,不宜合并,这更符合聚类的实际要求,是一种较好的系统聚类法。鉴于此,采用离差平方和法对在外界环境影响因素作用下的机动车司机违章驾驶行为进行聚类,运用 R-Project 生成聚类结果(见图 1)。

3.2 模糊聚类法

模糊聚类法是将模糊数学中的有关概念引入聚类分析,通过建立模糊相似关系构造等价矩阵,利用模糊等价关系对客观事物进行分类。

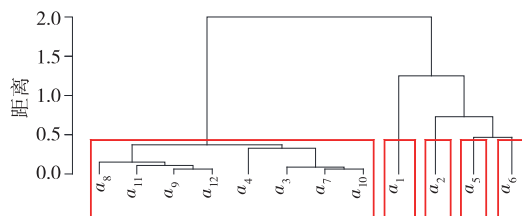


图 1 离差平方和法聚类结果

其分析步骤如下:1) 计算相似性统计量 r_{ij} 。对于每项待分类对象 a_i ,均有相同的 k 维变量 b_k ,将分类对象看作样本,则样本相似度 r_{ij} 可由夹角余弦法、相似系数法等方法得出。2) 构造等价矩阵 R 。模糊矩阵 $R = (r_{ij})_{n \times k}$ 只具有自反性和对称性,不具备传递性。用平方法得到传递闭包,即论域上的等价矩阵。 $R^2 = R \circ R = \bigvee_{k=1}^m (r_{ij} \bigwedge_{i,j=1}^m r_{ij})$,依次类推,当 $R^{2k} = R^k \circ R^k = R^k$ 时, R^k 即为等价矩阵。3) 求解布尔矩阵 R_λ 。根据 R^k 选定不同截取水平 λ 值 ($\lambda \in [0,1]$),由 λ 值得出动态聚类结果。

运用 MATLAB 软件编程对数据进行处理,运用相似系数法计算模糊相似矩阵 $R = (r_{ij})_{n \times k}$,得出等价矩阵 R^k 及 λ 截矩阵 $R_\lambda = (\lambda r_{ij})_{n \times k}$ 。其中:

$$\lambda r_{ij} = \begin{cases} 1 & (r_{ij} \geq \lambda) \\ 0 & (r_{ij} < \lambda) \end{cases} \quad (4)$$

通过程序运算得出等价矩阵 R^4 :

$$R^4 = \begin{pmatrix} 1.00 & 0.92 & 0.92 & 0.93 & 0.97 & 0.93 & 0.96 & 0.97 & 0.97 & 0.90 & 0.97 & 0.97 \\ 0.92 & 1.00 & 0.99 & 0.92 & 0.92 & 0.92 & 0.88 & 0.92 & 0.92 & 0.80 & 0.88 & 0.88 \\ 0.92 & 0.99 & 1.00 & 0.92 & 0.89 & 0.92 & 0.88 & 0.89 & 0.89 & 0.80 & 0.88 & 0.88 \\ 0.93 & 0.92 & 0.92 & 1.00 & 0.93 & 0.99 & 0.93 & 0.93 & 0.93 & 0.90 & 0.93 & 0.93 \\ 0.97 & 0.92 & 0.89 & 0.93 & 1.00 & 0.93 & 0.96 & 0.99 & 0.99 & 0.90 & 0.99 & 0.99 \\ 0.93 & 0.92 & 0.92 & 0.99 & 0.93 & 1.00 & 0.90 & 0.93 & 0.93 & 0.88 & 0.93 & 0.93 \\ 0.96 & 0.88 & 0.88 & 0.93 & 0.96 & 0.90 & 1.00 & 0.96 & 0.96 & 0.90 & 0.96 & 0.96 \\ 0.97 & 0.92 & 0.89 & 0.93 & 0.99 & 0.93 & 0.96 & 1.00 & 0.99 & 0.90 & 0.99 & 0.99 \\ 0.97 & 0.92 & 0.89 & 0.93 & 0.99 & 0.93 & 0.96 & 0.99 & 1.00 & 0.90 & 0.99 & 0.99 \\ 0.90 & 0.80 & 0.80 & 0.90 & 0.90 & 0.88 & 0.90 & 0.90 & 0.90 & 1.00 & 0.90 & 0.90 \\ 0.97 & 0.88 & 0.88 & 0.93 & 0.99 & 0.93 & 0.96 & 0.99 & 0.99 & 0.90 & 1.00 & 0.99 \\ 0.97 & 0.88 & 0.88 & 0.93 & 0.99 & 0.93 & 0.96 & 0.99 & 0.99 & 0.90 & 0.99 & 1.00 \end{pmatrix}$$

根据等价矩阵,选取不同 λ 值,由式(4)得出动态聚类结果(见表 4 和图 2)。其中: $\lambda \in [0, 0.8016]$ 忽略了外部环境因素对不同违章驾驶行为造成交通事故影响的差异性,分类太过粗糙; $\lambda \in (0.9736, 1]$ 则过分强调外部环境因素对不同违章驾驶行为造成

交通事故影响的差异性,而忽略了其中的关联性,导致没有起到分类的作用。

3.3 结果分析

根据两种聚类方法的聚类结果,在聚类个数 $k=5$ 时其聚类结果相同。

表 4 模糊聚类法聚类结果

λ 取值范围	聚类结果	聚类数/类
$[0,0.801\ 6]$	$\{a_1,a_2,a_3,a_4,a_5,a_6,a_7,a_8,a_9,a_{10},a_{11},a_{12}\}$	1
$(0.801\ 6,0.955\ 2]$	$\{a_1\},\{a_2,a_3,a_4,a_5,a_6,a_7,a_8,a_9,a_{10},a_{11},a_{12}\}$	2
$(0.955\ 2,0.973\ 6]$	$\{a_1\},\{a_2\},\{a_5\},\{a_6\},\{a_3,a_4,a_7,a_8,a_9,a_{10},a_{11},a_{12}\}$	5
$(0.973\ 6,1]$	$\{a_1\},\{a_2\},\{a_3\},\{a_4\},\{a_5\},\{a_6\},\{a_7\},\{a_8\},\{a_9\},\{a_{10}\},\{a_{11}\},\{a_{12}\}$	12

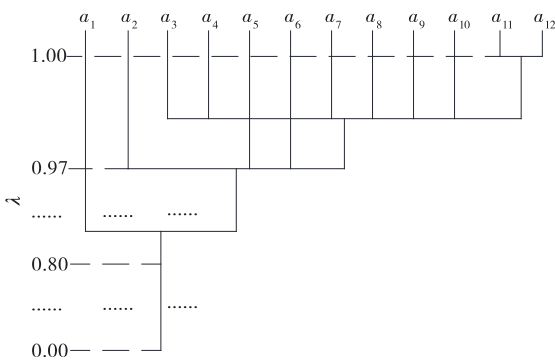


图 2 模糊聚类法聚类结果

(1) 在以驾驶人违章驾驶行为为统计样本、外部影响因素为指标变量时,可将大广高速公路交通事故起因分为五大类。同时根据图 1 和表 2,类间距与各类违章驾驶行为引发交通事故的直接经济损失之差成正比,符合实际数据统计情况。

(2) 根据聚类结果及交通事故造成的经济损失,大广高速公路在天气状况、道路线形、车辆状况等外界环境因素影响下,机动车司机违章驾驶行为引发交通事故带来的损失程度大小为 $\{a_6\}>\{a_5\}>\{a_2\}>\{a_1\}>\{a_3,a_4,a_7,a_8,a_9,a_{10},a_{11},a_{12}\}$ 。

4 结论

(1) 通过聚类分析,得出归并为一类的引发交通事故的驾驶行为与其对应的外部影响因素具有一定联系。在外界环境无法改变的情况下,可从人着手,教育司机文明驾驶,提高司机驾驶素质。

(2) 为减少高速公路交通事故,必须重点整治驾驶违章车辆和未保持安全行车距离等违章驾驶行

为,其次要整治疲劳驾驶和超速驾驶等违章驾驶行为,还应预防违章变更车道、违章超车等驾驶行为。

参考文献:

[1] Klauter S G,Dingus T A,Neale V L,et al. Comparing real-world behaviors of drivers with high versus low rates of crashes and near crashes[R].National Highway Traffic Safety Administration,2009.

[2] 马壮林,邵春福,李霞.高速公路隧道交通事故严重程度的影响因素分析[J].北京交通大学学报,2009,33(6).

[3] 李金龙,孙晚华.高速公路交通事故成因分析及对策研究[J].中国安全科学学报,2005,15(1).

[4] 文斌,曹东伟.高速公路路面抗滑力与交通事故的统计分析[J].公路交通科技,2006,23(8).

[5] 戴齐,高建平,刘敏.雾天高速公路交通事故成因分析及管理措施[J].西部交通科技,2010,31(1).

[6] 张长生,马荣国.山区高速公路交通事故分析及多发路段鉴别[J].长安大学学报:自然科学版,2010,30(6).

[7] 盛骤,谢式千,潘承毅.概率论与数理统计[M].第 4 版.北京:高等教育出版社,2008.

[8] 田兵.系统聚类法及其应用研究[J].阴山学刊,2014,28(2).

[9] 李倩星.R 语言实战:编程基础、统计分析与数据挖掘宝典[M].北京:电子工业出版社,2016.

[10] 黄卫华,方康玲.模糊控制系统及应用[M].北京:电子工业出版社,2012.

[11] 陈明,郑彩云,张铮.Matlab 函数与实例速查手册[M].北京:人民邮电出版社,2014.

收稿日期:2017-02-22

(上接第 35 页)

[5] 姜爱华,张驰.城镇化进程中的“城市病”及其治理路径探析[J].中州学刊,2012(6).

[6] 冉晓宁.网约车合法化再起争端[N].北京青年报,2016-10-04.

[7] 林园,刘云.鼓励网约车新业态创新规范发展[N].羊城晚报,2016-03-14.

[8] 李晓伟.网约车合法后还有三大问题待解[EB/OL].中国青年网,2016-07-29.

[9] 牛学辉.网络使用行为规制路径研究[J].管理世界,2011(8).

[10] 冀勇庆.Uber 中国遭遇战[J].中国企业家,2014(15).

收稿日期:2016-12-30