

高速公路长大坡路段交通事故安全评价与分析

王丽丽，张生瑞

(长安大学公路学院，西安，710064)

摘要：本文运用主成分—聚类分析法，结合 SPSS 统计分析软件，进行了高速公路长大坡路段的交通事故安全评价，理论与实际相结合，针对具体实例，找出了主要的事故形态及事故多发时段，并结合安全评价的结果，进行了高速公路长大坡路段事故多发的原因分析，有利于高速公路的安全治理。该评价方法考虑到现有的评价指标的局限性，在现有评价指标的基础上，通过主成分分析法构造出新的综合评价指标，并对主成分分析结果进行了聚类分析。交通事故安全综合评价指标，是针对现有交通事故评价指标体系所存在的不足，提出的一种新的交通事故评价分析方法，是对现有交通事故评价指标体系的补充与完善。

关键词：主成分分析；聚类分析；长大坡；交通事故；安全评价

Research on Traffic Safety Evaluation at Long and Steep Road based on Principal Component analysis and Cluster analysis

Wang Li-li, Zhang Sheng-rui

(Highway College, Chang'an University, Xi'an 710064, China)

Abstract: By applying the principal component and cluster analysis method and the statistic software SPSS, this paper single out principal component and cluster analysis in traffic safety evaluation at long and steep road. It not only makes up for the mistakes made by artificial factors in AHP and vague comprehensive valuating methods but also overcomes the shortcomings in principal component analysis without considering relative importance between index. Traffic accident safety comprehensive evaluation index, is aim at the shortage of existing traffic accident evaluation index system, a kind of new traffic accident evaluation for putting forward analytical method, is an evaluate index system to the existing traffic accident of complement with perfect.

Keywords: Principal Component analysis; Cluster analysis; Long and Steep Road; Traffic accident; Safety evaluation

随着我国公路事业的不断发展，高速公路大规模的修建。高速公路由于道路线形限制或造价方面的影响，出现纵坡或急弯等的道路在所难免，有些地段由于地形限制会出现长大坡，甚至长大陡坡，影响车辆的正常行驶。在目前的高速公路交通系统中，长大坡路段的交通事故率相对较高，常常成为“事故多发路段”。道路的纵坡，无论是上坡或下坡，对交通事故的形成都有直接影响，这一点在山区公路中尤为明显^[1]。因此，高速公路长大坡路段的交通事故安全评价成为安全治理的一项重要环节。

1 现有的交通事故安全评价指标

在高速公路交通事故安全评价方面，国内外做了大量研究工作，有些评价方法已应用到实践中，取得了很好的效果。国内外用于统计分析交通事故状况所用最广泛的四项统计指标是交通事故次数、死亡人数、受伤人数以及直接经济损失。这四项绝对指标是认识事故的起点，又是构造其他相对评价指标的基础，在事故分析中具有重要意义。这四项指标基本涵盖了交通事故所造成各种损害的主要

作者简介：张生瑞（1963—），男，长安大学教授，工学博士、博士后，现为长安大学工程设计研究院院长。主要研究领域：交通规划、交通流理论、交通运输可持续发展、交通安全与环保、综合交通运输、ITS 等。

王丽丽（1983—），女，山东济宁人，长安大学在读硕士研究生，公路学院交通运输规划与管理专业，主要研究领域：交通规划、交通管理、交通安全与环保等。

要方面，表征的交通事故情况直接且直观，其含义十分明确。这四项统计数字是目前可以获得的积累时间最长，使用范围最广，统计口径最统一，数字来源最准确的交通事故记录方式，其所应用的各种交通事故统计分析方法的数据源，几乎无一不是建立在这四项统计数字之上的。^[2]高速公路的交通事故也普遍使用这四项进行事故统计。

现行交通安全评价指标是建立在交通事故四项统计数字的基础上的，实际应用结果表明，在对交通事故的四项统计数字进行对比时，往往会出现各项数字的对比结果并不一致的现象。

本文将在现有各种交通事故评价指标的基础上，应用主成分分析法，构造新的交通事故评价指标。该指标来源于交通事故四项统计数字，携带了四项统计数字所具有的主要信息，且保留了四项统计数字之间的相互关系、内在联系和变化规律。^[2]

2 基于主成分—聚类分析法的交通事故安全评价

2.1 主成分—聚类分析法的原理

主成分分析是一种把原来多个指标化为一组新的互不相关（或相互独立）的综合指标的一种统计方法，用以达到数据简化，揭示变量之间关系和进行统计解释的目的，为进一步分析总体的性质和数据的统计特性提供重要信息^[2]。聚类分析是根据事物本身的特性研究个体分类的数学方法，本文采用聚类分析中的 Q 型聚类，即对事件（Cases）进行聚类或是说对观测变量进行聚类。

主成分—聚类分析法，即主成分分析法与聚类分析法相结合的方法，该方法利用主成分分析法排除冗余指标的干扰，构造出交通事故综合评价指标，并利用聚类分析法将主成分分析结果进行聚类^[3]。

2.2 交通事故安全评价的步骤

运用主成分—聚类分析法的安全评价步骤如下，其中前 5 步都为主成分分析的相关步骤^[3~4]：
设有 n 个样本，每个样本观测 p 个指标。将原始数据写成矩阵形式如下：

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix}$$

(1) 数据标准化处理

用 Z—score 法对数据进行标准化处理，计算公式如下：

$$X_{ij}^* = (X_{ij} - \bar{X}_j) / S_j \quad (1)$$

其中： $i=1,2,3,\dots,n$ ， n 为样本数； $j=1,2,\dots,p$ ， p 为样本变量数目。

(2) 求指标数据的相关系数矩阵 R

$$R = (r_{ij})_{p \times p} \quad (2)$$

其中： r_{ij} 为指标 j 与指标 k 的相关系数。

(3) 求相关系数矩阵 R 的特征值与特征向量

特征根记为 $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \dots \geq \lambda_p \geq 0$ ，以及相应的单位特征向量：

$$\alpha_i = (\alpha_{1i}, \alpha_{2i}, \dots, \alpha_{pi}), i = 1, \dots, p$$

(4) 确定主成分因子的个数 m :

本文取因子的累计贡献率来确定 m , 前 m 个主成分因子的累计贡献率的计算方法:

$$G(m) = \sum_{i=1}^m \lambda_i / \sum_{i=1}^p \lambda_i \quad (3)$$

一般方差的累计贡献率应在 80%以上。贡献率越大，则对应的主成分反映 X 的能力越强，反之越弱。本文取 85%的累计贡献率作为界限。用较少的不相关的综合指标反映 X 的特征，得出主成分方程组。

贡献率以及上述的标准化数据、相关系数矩阵、特征根、主成分因子矩阵可以利用统计分析软件 SPSS (Statistics Package for Social Science) 求得。SPSS 软件强大的计算功能，可以简便快捷的得出想要的结论。

(5) 计算主成分因子得分:

令 $X = (X_{ij}^*)_{n \times p}$, 将标准化后的指标变量转化为主成分:

$$F_j = \alpha_{j1} X_1 + \alpha_{j2} X_2 + \dots + \alpha_{jp} X_p \quad (j=1, 2, \dots, m) \quad (4)$$

其中: F_1 称为第一主成分, ……, F_m 称为第 m 个主成分; X_i 是标准化后的数据。

然后根据这些主成分因子的得分，并将贡献率作为赋予权重的依据，得出各样本的综合得分，即各样本的综合评价指标值:

$$F_i = a_1 F_1 + \dots + a_m F_m \quad (i=1, \dots, n) \quad (5)$$

(6) 聚类分析

根据上述得出的主分量组成的指标体系，利用 SPSS 软件将所得到的主成分聚类，得出聚类树图及每类的最终综合评价值。对于同一类中不同指标，结合道路交通评价的特点，本文取这些不同指标的平均值作为类的最终值。

(7) 结果分析

依据主成分—聚类分析法的结果，分析产生事故的直接原因或间接原因。

3 交通事故安全评价实例应用

在此根据某高速公路的长大坡路段事故数据，采用主成分—聚类分析法，结合 SPSS 统计分析软件，对高速公路长大坡路段事故形态以及事故时段进行安全评价。

3.1 事故形态评价分析

(1) 特征值: $\lambda_1 = 3.022$, $\lambda_2 = 0.779$, $\lambda_3 = 0.160$, $\lambda_4 = 0.038$

贡献率: $r=2$ 时，主成分的累计贡献率为 95.030%，大于 85%。

λ_1 、 λ_2 对应的特征向量分别为:

$$C_1 = (0.50104, 0.52290, 0.50507, 0.47055)$$

$$C_2 = (0.44527, -0.44527, -0.52005, 0.57783)$$

(2) 评价统计分析

将主成分的贡献率作为各主成分的权重，得出的各事故形态的综合评价指标值。该高速公路的长大坡路段的交通事故数据及综合评价指标值见表 1。

表 1 长大坡路段交通事故形态数据及综合评价指标值一览表

序号	事故形态	事故次数	死亡人数	受伤人数	经济损失	综合评价指标值(F)
1	追尾	11	2	7	53665	2.556
2	方向失灵	6	0	1	35985	-0.036
3	刹车失灵	3	0	0	17085	-0.772
4	爆胎	12	0	1	51245	0.750
5	车辆着火	2	0	0	7990	-1.021
6	操作不当	19	1	2	82610	3.355
7	路滑	3	0	0	4360	-1.000
8	侧滑	2	0	0	17355	-0.852
9	躲避大车	3	0	0	62175	0.039
10	轮胎脱落	1	0	0	5000	-1.160
11	其他	2	0	0	5875	-1.059

根据表 1 中的综合评价指标值得出的交通事故形态评价指标图，如图 1 所示。

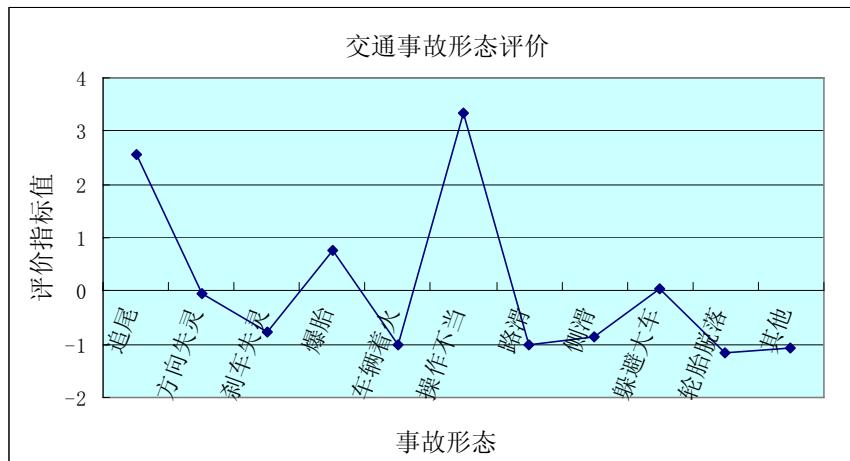


图 1 该高速公路长大坡路段交通事故形态评价指标图

(3) 聚类分析

根据事故形态状况的数据，利用 SPSS 软件，以欧氏距离作为衡量各原始指标差异大小，采用最短距离法对各事故形态进行系统聚类，并经过对图形的整理，得出该高速公路长大坡事故形态的聚类分析图，见图 2。

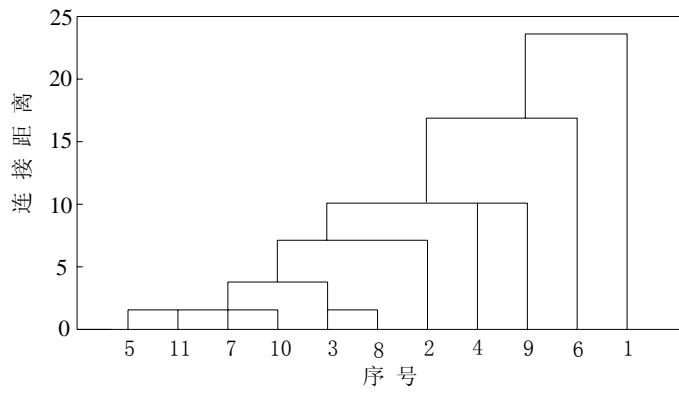


图2 欧氏距离最短距离法的聚类分析图

聚为4类的聚类分析结果见表2中。

表2 该高速公路的交通事故形态聚类结果

类序	类别	综合得分	结果分析
1类	操作不当	3.355	最严重
2类	追尾	2.557	很严重
3类	躲避大车	0.039	中等
4类	方向失灵, 刹车失灵, 爆胎, 车辆着火, 轮胎脱落, 路滑, 侧滑, 其他	-0.644	相对较好

(4) 评价结果原因分析

从分析及聚类结果可以看出, 该高速公路的长大坡路段上操作不当引起的交通事故最为严重, 追尾次之。这两种的事故形态的严重程度均与长大坡路段的道路条件有很大的关系, 特别是长大坡的坡度和坡长方面。

车辆上坡时, 由于坡度的阻力使车速降低, 坡道越陡, 车速下降越快; 坡道越长, 车速降低越多。为了维护汽车的爬坡能力, 要不断增加牵引力, 如果牵引力不足, 制动不及时, 操作失误, 就会造成车辆溜滑, 从而引起交通事故的发生。车辆下坡时, 驾驶员为节油常采用熄火滑行的操作方法, 一旦遇到紧急情况来不及采取应急措施, 这种由于操作不当而引起的交通事故在高速公路中屡见发生。另外, 由于重力作用, 车辆下长大坡时行驶速度过高, 制动非安全区过长, 频繁使用制动致使制动产生热衰减, 遇到紧急情况不能及时停车, 常造成追尾事故。^[5]

此外, 超载也是引起这两种主要事故形态的重要原因之一。由于车辆超载, 使得车辆沿坡道方向的平行分重力加大, 从而使得上述分析的问题加重, 促使长大坡路段事故的发生频率增大。

3.2 事故时段评价分析

对于事故时段的评价步骤不再复述, 事故时段数据及其评价指标值见表3, 根据表3中的综合评价指标值得出的交通事故时段评价指标图见图3。

表3 该高速公路长大坡路段交通事故时段数据及综合评价指标值一览表

序号	时段统计	事故次数	死亡人数	受伤人数	经济损失	综合评价指标值 (F)
1	0~1	2	0	1	8945	-0.386
2	1~2	1	0	0	3010	-0.834
3	2~3	0	0	0	0	-1.12
4	3~4	3	0	0	21065	0.19
5	4~5	1	0	0	50000	0.934

6	5~6	5	1	2	34265	1.502
7	6~7	2	0	1	18615	-0.022
8	7~8	8	0	0	28335	1.326
9	8~9	3	1	2	12205	0.33
10	9~10	4	0	0	16605	0.192
11	10~11	2	1	2	4800	-0.124
12	11~12	2	0	0	3450	-0.645
13	12~13	4	0	0	21665	0.385
14	13~14	3	0	0	8630	-0.278
15	14~15	1	0	0	2290	-0.861
16	15~16	9	0	0	43300	2.062
17	16~17	3	0	0	26025	0.379
18	17~18	1	0	0	2500	-0.854
19	18~19	0	0	0	0	-1.12
20	19~20	4	0	0	18580	0.269
21	20~21	2	0	0	6935	-0.514
22	21~22	1	0	0	9050	-0.607
23	22~23	4	0	3	15235	0.301
24	23~24	2	0	0	7230	-0.503

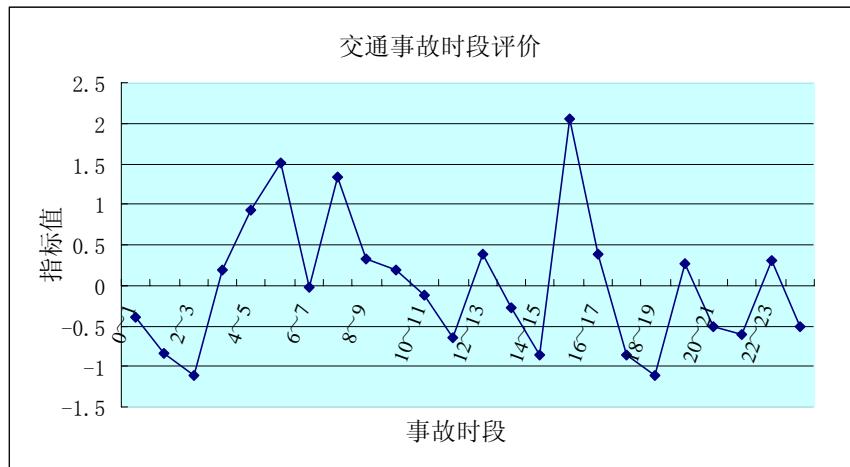


图 3 交通事故时段评价指标图

下面仅对聚类分析及结果分析结果进行介绍。

(1) 聚类分析

根据事故时段状况的数据，经过图形整理，得出的欧氏平方距离重心法聚类分析图见图4，表4为聚为4类的聚类分析结果。

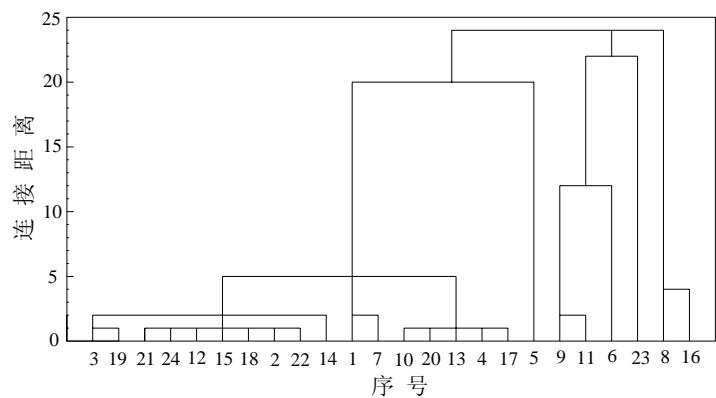


图4 欧氏平方距离重心法的聚类分析图

表4 该高速公路的事故时段聚类结果

类序	类别	综合得分	结果分析
1	7~8, 15~16	1.694	最严重
2	5~6, 8~9, 10~11	0.568	较严重
3	22~23	0.301	严重
4	0~5, 6~7, 9~10, 11~15, 16~22, 23~24	-0.230	相对较好

(2) 评价结果原因分析

从上述的分析及聚类结果看出，一天 24 小时内，事故高峰期出现在 5~6 时、7~8 时及 15~16 时。前两个时段为凌晨时段，由于夜间驾驶员的连续驾驶，到这个时段已经产生疲劳，再加上这两个时段交通流量较小，驾驶员的注意力不够集中，从而容易诱发交通事故。15~16 时为高速公路的高峰时段，交通流量较大，引发的交通事故也相对较多。

4 结语

本文通过运用主成分——聚类分析方法，结合 SPSS 统计分析软件，对长大坡路段的交通安全进行了评价，利用方差贡献率作为综合评价指标的权重，弥补了层次分析法、模糊综合评价法等方法人为因素较多的缺陷，同时又克服了主成分分析法未考虑指标间相对重要程度的不足^[6]。从交通事故安全评价结果中可以找出了事故原因的所在，有利于交通安全工作人员针对一定条件下的事故成因及事故时段的分析结果，采取相应的道路改善和交通管理措施，对交通事故进行治理和防范。

参考文献

- [1] 张永生等. 山区长大下坡对道路交通安全的影响及其工程措施研究[J]. 交通标准化, 2006, (8): 160.
- [2] 吴克海等. 基于主成分法的道路交通事故综合评价方法研究[J]. 交通标准化, 2005, 147 (11): 143—144.
- [3] 王俊人. 道路交通安全综合评价方法及对策研究[D]. 中北大学硕士学位论文, 2006.
- [4] 余建英, 何旭宏. 数据统计分析与 SPSS 应用[M]. 北京: 人民邮电出版社, 2003.
- [5] 郭克清等. 公路安全保障工程实用手册[M]. 北京: 人民交通出版社, 2007.
- [6] 张慧颖. 基于加权关联距离的公路网综合评价方法研究[J]. 中南公路工程, 2006, 31 (5): 104.