贝叶斯矩法在交通事故预警中的应用

王 宁,邹志云
(华中科技大学,湖北武汉 430074)

摘 要:该文简要分析了我国交通事故的严重态势和交通事故发生的原因,提出了交通事故预警的计算模型-BMOM方法,并结合实例进行了分析。

关键词:交通事故;BMOM方法;交通预警
中图分类号:U491.3 **文献标识码**:A **文章编号**:1009-7716(2006)05-0191-03

1 概述

随着社会经济的发展和人民生活水平的提高,机动车的拥有量急剧上升,随之而来的交通事故问题也摆在了人们的眼前。如何减少交通事故成了社会的一个话题,有关专家也积极探讨减少交通事故的方法。BMOM方法不仅能确切地预测交通事故,还可预警交通事故,从而在一定程度上有助于减少交通事故的发生。

2 我国交通事故的严峻态势与原因简析

有关部门统计数字表明,2002年我国交通事故次数为773137次,死亡人数109381人次,受伤人数562248人,直接经济损失33.24亿元。我国的交通事故人数已经连续10余年居世界第一。

产生交通事故的原因有很多,主要有人、车、道路和交通管理四个方面的原因。

2.1 人的原因

人是产生交通事故的直接原因,可以说90%(按1999年交通事故的主要原因统计,人的因素占93%)以上的交通事故的发生都或多或少含有人的因素,因此人是交通事故分析的主要对象。其中驾驶员对交通事故的发生有着很大的作用。不同性别、年龄、体质的驾驶员,其生理、心理、感知、分析、判断和反应均不可能完全相同。而感知迟钝、判断不准、操作失误在事故中占绝大多数,其中感知迟缓一般约占60%左右,这多半是由于生理、心理因素所决定的。

2.2 车的原因

在大量事故统计资料中,由于车辆的各种故障而造成交通事故的虽不太多,但从预防考虑,“车”仍是一个重要因素。如转向系统、刹车系统的稳定性、灵活性都有影响,同时由于车辆的性能、新旧、维修的好坏、加速及减速度的大小也有一定的作用。车辆带“病”行驶、保养不当也会增加交通事故发生的几率,而且后者的影响更多一些。

2.3 道路的原因

道路是车辆的载体、行驶的基础,道路的等级质量和线性标准等对交通事故的发生有着重要的影响。平曲线、竖曲线、坡度和路面质量对交通事故也有很重要的影响。

2.4 交通管理的原因

交通管理方面的原因主要表现为管理技术和手段落后、交通法规不健全等。交通管理水平高,交通事故必然要减少,反之则增多。

3 BMOM模型及参数估计

通常的回归模型的表达式如下:

$$Y_{n\times 1}=X_{n\times k}\beta_{\kappa\times 1}+u_{n\times 1} \quad (1)$$

式中,X 为解释变量矩阵,Y 为因变量,β 为未知参数,u 为误差项。

假定1

$X'E(u\mid D)=0$,其中D为全部已知数据Y和X的条件下,对式(1)两边取条件期望,有$Y=XE(\beta\mid D)+E(\beta\mid D)$,两边同乘以$(X'X)^{-1}X'$,就得到:

$$E(\beta\mid D)=(X'X)^{-1}X'Y\equiv\hat{\beta} \quad (2)$$

$$E(u\mid D)=Y-XE(\beta\mid D)=[1-(X'X)^{-1}X']Y\equiv\hat{u} \quad (3)$$

此处$\hat{u}$即为残差,且

$$X'\hat{u}=X'Y-X'XE(\beta\mid D)=X'Y-X'Y=0$$

同时,$u-\hat{u}=(Y-X\beta)-(Y-X\hat{\beta})=X(\hat{\beta}-\beta)$ (4)

$$\hat{\beta}-\beta=(X'X)^{-1}X'X(\hat{\beta}-\beta)=(X'X)^{-1}X'(u-\hat{u}) \quad (5)$$

$$u-\hat{u}=X(X'X)^{-1}X'(u-\hat{u})$$

记 $P_x=X(X'X)^{-1}X'$ $P_x(u-\hat{u})$

由此可得:

$$Var(u\mid D)=E[(u-\hat{u})(u-\hat{u})'\mid D]$$

收稿日期:2006-06-06
作者简介:王宁(1981-),女,河北石家庄人,研究生,从事交通规划与管理研究工作。

$$=E[P_x(u-\hat{u})(u-\hat{u})'P_x \mid D]$$
$$=P_x Var(u \mid D)P_x \qquad (6)$$

$$Var(\beta \mid D)=E[(\beta-\hat{\beta})(\beta-\hat{\beta})' \mid D]$$
$$=E[(X'X)^{-1}X'(u-\hat{u})(u-\hat{u})'X(X'X)^{-1} \mid D]$$
$$=(X'X)^{-1}X'E[(u-\hat{u})(u-\hat{u})' \mid D]X(X'X)^{-1}$$
$$=(X'X)^{-1}X'Var(u \mid D)X(X'X)^{-1} \qquad (7)$$

由式(1)可知，β 是由 k 个未知参数构成的向量，若 β 已知，u 则确定，因此 u 有 k 个自由度，可以假定 $Var(\beta \mid D)$为一个秩为 k 的矩阵乘以一个正的常数。

假定 2

$Var(u \mid D,\sigma^2)=\sigma^2 P_x$，其中 $\sigma^2=u'u/n$，这里 σ^2 是未知的，由 σ^2 的定义和式(4)可得：

$$E(\sigma^2 \mid D)=\frac{1}{n}E(uu' \mid D)=\frac{1}{n}E[\hat{u}+X(\hat{\beta}-\beta)]'[\hat{u}+X(\hat{\beta}-\beta) \mid D]$$
$$=\frac{1}{n}E(\hat{u}'\hat{u}+(\beta-\hat{\beta})'X'X(\beta-\hat{\beta}) \mid D)$$
$$=\frac{1}{n}[\hat{u}\,\hat{u}+trX'E(\beta-\hat{\beta})'(\beta-\hat{\beta}) \mid D]$$
$$=\frac{1}{n}[\hat{u}'\hat{u}+kE(\sigma^2 \mid D]$$

解得　$$s^2=E(\sigma^2 \mid D)=\frac{1}{n-k}\hat{u}'\hat{u} \qquad (8)$$

$E(\sigma^2 \mid D)=s^2P_x$，代入式(7)有 $Var(\beta \mid D)=s^2(X'X)^{-1}$，至此 β 的期望和方差协方差阵均得到，利用极大熵原则，有 $\beta \mid D \sim N[\hat{\beta},s^2(X'X)^{-1}]$。

根据 X 与 Y 建立模型后，设已知 X_f，要对 Y_f 进行预测，于是 $Y_f=X_f\beta+u_f$，假定 $E(u_f \mid D')=0$、$E(u_f u'_f \mid D',\sigma^2)=\sigma^2 I$、$E(u'_f X \mid D')=0$（其中 D' 为 X、Y 和 X_f 已知数据），可得：

$$E(Y_f \mid D')=X_f E(\beta \mid D')=X_f\hat{\beta}$$
$$Var(Y_f \mid D')=Var(X_f\beta \mid D')+Var(u_f \mid D')$$
$$=s^2[X_f(X'X)^{-1}X'_f+I]$$

同样，可得 Y_f 的极大熵分布为：

$$Y_f \sim N\{X_f\hat{\beta},s^2[I+X_f(X'X)^{-1}X'_f]\}$$

4　BMOM 方法在交通事故预警中的应用

4.1　模型的建立

设影响某地区交通事故数量的因素分别为 $X_1,X_2\cdots X_{k-1}$，可以利用第 T 年的 k 个影响因素指标对第 T+1 年的交通事故数量 Y_{T+1} 做出预测，于是对应的 X、Y、β、X_f 分别表示如下：

$$X_{(T-1)\times k}=\begin{bmatrix} 1 & x_{11} & x_{12} & \wedge & x_{1k-1} \\ 1 & x_{21} & x_{22} & \wedge & x_{2k-1} \\ 1 & & \wedge & & \\ 1 & x_{T-1,1} & x_{T-1,2} & \wedge & x_{T-1,k-1} \end{bmatrix}$$

$$Y=\begin{bmatrix} y_2 \\ y_3 \\ M \\ y_T \end{bmatrix} \qquad \beta=\begin{bmatrix} \beta_0 \\ \beta_1 \\ M \\ \beta_{k-1} \end{bmatrix}$$

满足模型 $Y=X\beta+u$，在给出第 T 年的各个的指标值 $X_f=(1,x_{T1},\wedge,x_{TK-1})$，预测值 Y_f 即为第 T+1 年的交通事故数量预测值 Y_{T+1}。由于 $Y_f \sim N\{X_f\hat{\beta},s^2[1+X_f(X'X)^{-1}X'_f]\}$，因此 Y_{T+1} 大于 C 的概率即需要警报的概率为：

$$P1=\{\hat{Y}_{T+1} \geqslant C\}=1-\Phi[\frac{C-X_f\hat{\beta}}{s\sqrt{1+X_f(X'X)^{-1}X'_f}}]$$

4.2　自变量的选取及检验

通过上述讨论，可以得出，一个地区交通事故的发生有人的原因，车的原因、道路的原因和交通管理的原因，因此，可以选取一个地区的驾驶员的行车遵章率 X_1、行人交通安全意识 X_2、机动车保有量 X_3、道路面积 X_4、道路长度 X_5、主次干道的平均车速 X_6 等 6 个指标作为自变量。可以视不同的城市增减，只要适合城市的交通特点就可以。其中，X_1、X_2 都是人的因素，可以加权平均得到一个自变量，同理 X_4、X_5 也可以合并成一个自变量。

一个自变量的选取是否合理，应该经过检验，具体的检验方法如下。

设有模型如下：

$$Y=X_1\beta_1+X_2\beta_2+u$$

这里的 X_i 为 $n\times k_i$，秩为 k_i 的矩阵，X_2 为待考察的变量。X_2 是否值得引入的问题等价于考察 β_2 是否为零，即 $\beta_2=0$ 是否成立。考虑预测分布，引入损失函数，用 $\delta(y)$估计 β_2 时相应的损失为：

$$E(L \mid D,\delta(y))=[(Y_f-\hat{Y}_f)'(Y_f-\hat{Y}_f) \mid D,\delta(y)]$$

其中 Y_f 为真实值，$\hat{Y}_f$ 为预测值，比较 $\delta(y)=0$ 和 $\delta(y)=\hat{\beta}_2$ 的平均损失。

$$E[L \mid D,\delta(y)=0]$$
$$=E[(Y_f-\hat{Y}_f)'(Y_f-\hat{Y}_f) \mid D,\delta(y)=0]$$
$$=E[(\beta_1-\hat{\beta}_1)'X_1'X_1(\beta_1-\hat{\beta}_1) \mid D,\delta(y)=0]+E[u'u \mid D,\delta(y)=0]$$
$$=trE[X_1'X_1(\beta_1-\hat{\beta}_1)'(\beta_1-\hat{\beta}_1) \mid D,\delta(y)=0]$$

$+trE[u'u \mid D, \delta(y)=0]$

$= k_1 s_1^2 + n s_1^2$

$=(k_1+n)s_1^2$

同理，$\delta(y)=\beta_2$ 的平均损失为：

$$E[L \mid D, \delta(y)=\hat{\beta}_2] = (n+k_1+k_2)s_2^2+L_a$$

式中，L_a 相当于引入 X_2 后对模型负责化的惩罚，a 表示 L_a 的值与模型的假定有关。两种平均损失比为：

$$R=\frac{s_1^2}{s_2^2}\left(\frac{n+k_1}{n+k_1+k_2+L_a/s_2^2}\right)$$

$$=\frac{s_1^2/s_2^2}{(n+k_1+k_2+L_a/s_2^2)/(n+k_1)}$$

$$=\frac{F}{(n+k_1+k_2+L_a/s_2^2)/(n+k_1)}$$

式中，若R=1，表示两种损失相等，其中 F 为传统回归分析中模型比较时的 F 统计量，当 $F>\frac{n+k_1+k_2+L_a/s_2^2}{n+k_1}$ 时，表示可以引入变量 X_2，反之不引入 X_2 为好。

通过以上的分析，可以得到如下结论：

(1)交通事故的发生是有自身原因的，这些原因的统计数据可以作为自变量 X_f，而交通事故数量是因变量 Y_f，交通事故是上述原因的函数。

(2)可以用 BMOM 方法来建立二者的关系

(3)建立二者关系以后，就可以得到交通事故数量 Y_f 服从正态分布的结论，这样就可以得到交通事故数量小于一个预期值 C 的概率。

5　应用实例

利用全国 1970～1993 年的有关资料，利用 BMOM 方法进行预测，对变量进行筛选建立模型 $y_t=90.9224+0.0920x_{t-1}$，其中 y_t 为第 t 年得交通事故总数（单位为千起），x_{t-1} 为 t-1 年得机动车保有量（单位为万辆）。

$\hat{\beta}=(90.9224, 0.0920)'$

$s^2=1\ 080\ 000$

$$(X'X)^{-1}=\begin{pmatrix} 0.0778 & -5.7761E-05 \\ -5.7761E-05 & 9.2276E-08 \end{pmatrix}$$

现有 1993 年机动车保有量为 2 331.64 万辆，那么 1994 年交通事故预测值应为 $Y_f \sim N(118.6078, 1\ 140\ 000)$，这里给出 C 值以后就可以得到 $Y_f < C$ 的概率。

6　结语

该文引入了贝叶斯矩法(BMOM)用于预警交通事故，该方法的主要优点在于有明确的分布函数，可以知道交通事故发生小于某一个值 C 的概率，因而用于交通预测和控制更加方便。当然交通事故的减少和预防，并非单依靠某一个方法就可以达到的。

参考文献

[1]Arnold Zellner.Justin Tobias Ferther Results On Baysian Method Of Moments Analysis Of The Multiple Regression Model [J].International Economic Review,2001,42(1).

[2]宁自军.贝叶斯矩法在粮食安全预警中的应用[J].2003,18(2).

[3]徐吉谦.交通工程总论[M].人民交通出版社.

[4]公安部交通管理局.全国交通事故资料汇编[M].群众出版社,1993.

(上接 190 页) 酯聚脲作为高功能的工程涂料已经在欧美工程实例中得到了证明，也确实为许多工程难题提供了解决方案。我们将会看到，随着这些技术向中国乃至亚太地区的引进和推广，其应用范围将会得到进一步的拓宽，其工程应用的价值也将会得到进一步彰显和加强。

胶州湾隧道年内开建

总投资 31.8 亿元的胶州湾海底隧道项目合作协议日前正式签约，并将于年底前正式开工建设。

海底隧道项目将由青岛国信实业有限公司与上海上实(集团)有限公司共同出资建设。该项目南接薛家岛，北连团岛，下穿胶州湾湾口海域，工程全长 6 170 m。隧道采用双向双洞六车道，设计时速为 80 km，路线等级为城市快速路，预计建设期限为 3 到 4 年半。

经过前期认真详实的地质勘察，有关部门和专家已经确定：胶州湾海域地址情况基本搞清，建设海底隧道的地质条件比较理想，下一步在经过专家进一步论证、设计招标后，将在年底前正式开工建设海底隧道。