

高速公路路侧事故起数预测模型

马壮林¹, 张宏璐², 张祎祎¹, 王 晋³

(1. 长安大学 汽车学院, 陕西 西安 700064; 2. 德邦物流有限公司, 上海 201702;
3. 云南省交通科学研究院, 云南 昆明 650011)

摘 要:为识别高速公路路侧事故的主要诱导因素,分析路侧事故起数与道路线形和交通条件之间的关系,以广珠(广州—珠海)东线高速公路 3 年中发生的 178 起路侧事故为基础,分别采用定长法和不定长法划分路段单元,从道路线形和交通条件中选取 11 个自变量,采用零堆积负二项回归模型建立路侧事故起数预测模型。选择 Vuong 检验统计量、对数似然值和 AIC 信息准则 3 个指标进行模型的拟合优度检验,选择相对误差和累积残差 2 个指标进行模型的拟合准确性检验;通过对比分析负二项回归模型和零堆积负二项回归模型的拟合优度和拟合准确性检验结果判断其优劣性,并采用弹性分析确定较优模型中显著自变量对因变量的影响程度。研究结果表明:无论采用定长法划分还是不定长法划分路段单元,零堆积负二项回归模型构建的路侧事故起数预测模型明显优于负二项回归模型;采用零堆积负二项回归模型构建的路侧事故起数预测模型,其定长法划分的路段单元模型的拟合准确性优于不定长法;对于定长法划分的路段单元,基于零堆积负二项回归模型的路侧事故起数预测模型有 5 个自变量对路侧事故起数均有显著影响,影响程度大小依次为车道数、曲率变化率、曲线比例、曲度和平均纵坡坡度。

关键词:交通工程;路侧事故起数预测;负二项回归模型;零堆积负二项回归模型;弹性分析
中图分类号:U491.3 **文献标志码:**A

Roadside accident frequency prediction model on expressway

MA Zhuang-lin¹, ZHANG Hong-lu², ZHANG Yi-yi¹, WANG Jin³

(1. School of Automobile, Chang'an University, Xi'an 710064, Shaanxi, China; 2. Deppon Logistics Co. Ltd., Shanghai 201702, China; 3. Yunnan Science Research Institute of Communication & Transportation, Kunming 650011, Yunnan, China)

Abstract: To identify the main induction factors of roadside accident and analyze the relationship between roadside accident frequency and main influence factors such as road geometry, traffic conditions of expressway, based on 178 cases of roadside accidents occurred in Eastern Guangzhou to Zhuhai Expressway in three years, this paper used the fixed-length method and homogeneous longitudinal grade method to divide the study section, selected eleven independent variables from the aspects of road geometry and traffic conditions, and adopted zero-inflated negative binomial regression model to establish roadside accident frequency prediction model. Three indicators including Vuong test statistics, log likelihood value, and Akaike's information

收稿日期:2016-12-03

基金项目:国家自然科学基金项目(51208052);高等学校青年骨干教师出国研修项目(201406565054);

中央高校基本科研业务费专项资金项目(310822151024, 310822161007)

作者简介:马壮林(1980-),男,辽宁鞍山人,副教授,工学博士,E-mail:zhuanglinma@chd.edu.cn.

criterion were used to test the goodness-of-fit of the model. Two indicators including relative error and cumulative residual were used to test the prediction accuracy of the model. To further judge the better model, negative binomial regression model and zero-inflated negative binomial regression model were compared by the goodness-of-fit and prediction accuracy. Elastic analysis was used to determine the influence degree of the independent variables on the dependent variable in the better model. The results show that the roadside accident frequency prediction model based on zero-inflated negative binomial regression model is better than that of negative binomial regression model both for fixed-length segmentation method and homogeneous longitudinal grade segmentation methods. For roadside accident frequency prediction model based on zero-inflated negative binomial regression model, the prediction accuracy of fixed-length segmentation method is better than that of homogeneous longitudinal grade segmentation method. For fixed-length segmentation method, five independent variables in roadside accident frequency prediction model based on zero-inflated negative binomial regression model have significant impact on roadside accident frequency, and the influence degree in descending order is number of lane, curvature change rate, curve ratio, curve degree, and average longitudinal gradient grade. 8 tabs, 2 figs, 22 refs.

Key words: traffic engineering; road side accident frequency prediction; negative binomial regression model; zero-inflated negative binomial regression model; elastic analysis

0 引言

路侧事故指车辆偏离行车道,驶入路肩带或路肩带以外的区域,并与护栏、车辆、行人、交通标志或其他坚硬危险物发生碰撞的事故,也包括翻车、坠车等事故。据统计,2011 年中国高速公路发生 9 583 起交通事故,其中路侧事故 3 657 起,占高速公路事故总起数的 38.16%;此外,路侧事故也造成严重的人员伤亡,直接导致 2 234 人死亡和 4 313 人受伤,分别占高速公路交通事故死亡和受伤人数的 34.65%和 33.16%^[1]。因此,以路侧事故历史统计数据为基础,构建高速公路路侧事故起数预测模型,可以发掘路侧事故与各种潜在风险因素之间的关系,有助于减少路侧事故的发生概率,从而提高中国高速公路的整体安全水平。

近年来,国内外学者对路侧事故进行了研究,并取得了一定的研究成果。Zegger 等分析了边坡、路侧净空、路侧障碍等对路侧事故的影响,发现当边坡比从 3:1 变为 7:1 时,路侧事故率显著下降,而当边坡比从 2:1 变为 3:1 时,路侧事故率仅下降了 2%^[2];Dumbaugh 分析了 3 种路侧设计方法对路侧事故的影响,发现增加路肩带宽度会导致路侧事故起数的增加^[3];李长城等采用统计分析方法,研究了低等级双车道公路路侧事故的发生规律^[4];张铁军等基于道路线形与交通状态等因素,建立了双车道路侧事故预测模型,发现交通量、货车比例、平曲线

曲度、所在地域等因素对路侧事故的发生有较大影响^[5];张铁军等还建立了山区双车道公路路侧危险度事故修正系数函数,发现路侧危险度对路侧事故的影响为抛物线形状,并呈持续下降趋势^[6];游克思等提出了用于评价路侧安全的路侧危险指数,并采用 Kappa 统计量对评价者主观打分的一致性进行分析^[7];刘彦等采用路侧事故风险指标作为中间变量,提出评价山区高速公路路侧事故风险等级的定量方法,该方法可分析局部路段和全路段的路侧安全隐患^[8]。

与事故总起数相比,路侧事故仅占其 1/3 左右,因此在一定的统计时间间隔内,一些路段上没有路侧事故发生。虽然负二项(negative binomial, NB)回归模型解决了样本数据过离散的问题,但其不能解释样本数据中零值较多的现象,因此一些学者采用零堆积回归模型分析具有较多零值的样本数据^[9]。零堆积回归模型最早于 20 世纪 60 年代提出,包括零堆积泊松(zero-inflated Poisson, ZIP)模型和零堆积负二项(zero-inflated negative binomial, ZINB)回归模型 2 种。近年来,该类模型已经广泛应用于交通事故分析与预测领域。

Shankar 等分别采用泊松、NB、ZIP、ZINB 模型建立交通事故起数预测模型,并对比分析这 4 种模型的拟合优度检验指标,发现 ZIP、ZINB 模型可以解决过多零值问题,预测效果比泊松模型和 NB 模型好^[10];Qin 等采用 ZIP 模型建立双车道公路不同

类型交通事故起数预测模型,发现交通事故起数与年平均交通量呈非线性关系^[11];Kumara 等采用 ZINB 模型建立交叉口交通事故起数预测模型,发现是否设置非控制左转车道、曲率和视距对交通事故起数均有显著影响^[12];Daniel 等采用 ZIP 和 ZINB 模型建立城市主干道货车事故起数预测模型,发现信号设置和路段长度对货车事故起数有显著影响^[13];Easa 等分别采用泊松、NB、ZIP、ZINB 模型建立华盛顿州道路弯道交通事故起数预测模型,发现曲度、车道宽度、出入口密度、路段长度、坡长和日平均交通量对弯道交通事故起数有显著影响^[14];高海龙等分别采用泊松、NB、ZIP、ZINB 模型建立双车道公路路侧事故起数预测模型,发现 ZIP 和 ZINB 模型更适合建立路侧事故起数、死亡人数和受伤人数预测模型^[15];唐铮铮等分别采用泊松、NB、ZIP、ZINB 模型建立双车道公路交通事故起数预测模型,发现交通量、货车比例、曲度和出入口密度对事故起数有显著影响^[16];Ayati 等分别采用泊松、NB、ZIP、ZINB 模型建立城市道路交通事故起数预测模型,发现交通量、车速、车道数、平曲线数量和出入口数量对交通事故起数有显著影响,且 ZINB 模型的预测效果最好^[17]。由此可见,ZINB 模型可以有效解释交通事故起数较少的情况,并得到了研究人员的一致认同。

但目前尚没有研究对交通事故起数较少进行量化分析。因此,为了明确 NB、ZINB 模型的适用范围,本文针对广珠(广州—珠海)东线高速公路 3 年中发生的 178 起路侧事故,分别采用定长法和不定长法将研究路段划分为若干个路段单元,从道路线形和交通条件 2 个方面选取 11 个自变量,建立基于 ZINB 模型的路侧交通事故起数预测模型,从模型的拟合优度和拟合准确性 2 个方面对比分析 NB 模型与 ZINB 模型的优劣,并采用弹性分析法量化不同自变量对路侧事故起数的影响程度。该研究成果有助于完善交通事故预测的相关理论,可为高速公路管理部门实施有效的预防措施提供科学依据。

1 基础数据收集及分析

1.1 数据收集

广珠东线高速公路起点为广州市南沙区亭角,终点为珠海市高新区金鼎,全程 62.4 km,设有 8 个收费站。其中坦尾至南朗为双向 6 车道,设计时速 120 km/h;南朗至金鼎为双向 4 车道,设计时速 100 km/h。考虑采集数据的完整性,以该高速

公路(K22+690~K72+690)3 年中发生的 178 起路侧事故为基础,采集的数据主要包括交通事故、道路线形和交通流量 3 个部分:

(1)交通事故数据来源于公安交警事故统计系统,包括每起路侧事故发生的时间、地点、事故原因和事故类型等;

(2)道路线形数据来源于高速公路施工图设计文件,包括道路的平、纵、横线形指标以及相应的隧道、桥梁等特殊路段;

(3)交通流量数据来源于高速公路管理部门安装的流量检测系统,包括不同路段单元的日交通量。

1.2 路段单元划分

路段单元划分是进行交通事故起数预测研究的前提条件,对预测模型的拟合优度和预测准确性有直接影响。目前,路段单元划分方法主要有定长法和不定长法 2 种,2 种方法各有优缺点^[18-22]。定长法是以固定的路段长度为划分单元,路段单元通常取 0.5 或 1 km。不定长法是以道路属性的变化为划分依据,每个路段单元的长度各不相同,但其最小长度一般不小于 400 m。为了比较这 2 种路段单元划分方法对路侧事故起数预测模型的影响,本文分别采用定长法和不定长法对研究路段进行划分,并构建路侧事故起数预测模型,对比分析不同路段划分情况下路侧事故起数预测模型的优劣。考虑到上、下行方向道路平、纵线形的差异,将上、下行方向分别进行研究。考虑到高速公路的线形指标较好,以 0.5 km 为固定长度统计交通事故起数可能导致事故起数分布趋于平均,无法突出危险路段。因此,定长法以 1 km 为固定长度,将研究路段划分为 100 个长度相同的路段单元,其中最小、最大纵坡坡度分别为-2.315%和+2.315%;不定长法则以相同坡度的路段为路段单元,将研究路段划分为 154 个坡度相同的路段单元,其中最小、最大路段长度分别为 400、1 430 m,平均路段长度为 649 m。

2 路侧事故起数预测模型

2.1 模型构建

NB 模型可以解决数据过度离散的问题,因此被广泛应用于交通事故起数预测研究。如果样本数据中有很多的零值时,采用 NB 模型分析会不可避免地导致参数估计偏差;相反,采用 ZINB 模型分析可以有效避免这个问题。ZINB 模型是一种双状态模型,第 1 个状态是事故起数为 0 的状态,即在一定时间间隔内第 i 个路段没有交通事故发生,其发生

概率为 P_i , 称为结构零值; 第 2 个状态是事故起数大于 0 的状态, 即在一定时间间隔内第 i 个路段有交通事故发生, 其发生概率为 $1 - P_i$, 称为抽样零值。则 ZINB 模型的概率分布 $P(Y_i)$ 为

$$P(Y_i = y_i) = \begin{cases} P_i + (1 - P_i) \frac{1}{(1 + K\lambda_i)^{1/K}} & y_i = 0 \\ (1 - P_i) \frac{\Gamma(y_i + (1/K))}{\Gamma(y_i + 1)\Gamma(1/K)} \cdot \frac{(K\lambda_i)^{y_i}}{(1 + K\lambda_i)^{y_i + (1/K)}} & y_i > 0 \end{cases} \quad (1)$$

式中: Y_i 为随机变量; y_i 为第 i 个路段路侧事故起数的观测值; λ_i 为第 i 个路段路侧事故起数的期望值; K 为过离散参数; $\Gamma(\cdot)$ 为 ZINB 模型的伽马函数。

ZINB 模型的均值 $E(Y_i)$ 和方差 $S(Y_i)$ 分别为

$$E(Y_i) = (1 - P_i)\lambda_i \quad (2)$$

$$S(Y_i) = \lambda_i(1 - P_i)(1 + P_i\lambda_i + \lambda_i/K) \quad (3)$$

当 $P_i = 0$ 时, 随机变量 Y_i 服从均值为 λ_i 、方差为 $\lambda_i(1 + \lambda_i/K)$ 的负二项分布, 即 ZINB 模型转变为 NB 模型。本文采用二元 Logit 模型确定 ZINB 模型中事故发生的概率, 则 ZINB 模型的形式为

$$\lambda_i = e^{\mathbf{x}_i^T \boldsymbol{\beta} + \sigma} \quad (4)$$

$$\lg(P_i/1 - P_i) = \mathbf{x}_i^T \boldsymbol{\gamma} + \epsilon \quad (5)$$

式中: \mathbf{x}_i^T 为影响 P_i 的协方差向量的转秩; $\boldsymbol{\beta}$ 、 $\boldsymbol{\gamma}$ 均为模型回归系数的向量; σ 、 ϵ 均为模型的误差项。

2.2 模型检验

2.2.1 拟合优度检验

选择 Vuong 检验统计量、对数似然值和 AIC 准则 3 个指标进行模型的拟合优度检验, 各检验指标具体含义如下所示。

(1) Vuong 检验统计量 V

Vuong 检验统计量由 Vuong 于 1989 年提出, 用于非嵌套模型, 它可以比较 ZINB 模型和 NB 模型对样本数据的拟合效果。V 值可以通过式(6)计算得到

$$V = \frac{\sqrt{n}(1/n \sum_{i=1}^n m_i)}{\sqrt{1/n \sum_{i=1}^n (m_i - \bar{m})^2}} = \frac{\sqrt{nm}}{S_{m_i}} \quad (6)$$

$$m_i = \lg(f_1(y_i)/f_2(y_i)) \quad (7)$$

式中: n 为观测样本数; m_i 为 ZINB 模型与 NB 模型的概率密度函数的商的对数值; \bar{m} 为 m_i 的均值; S_{m_i} 为 m_i 的方差; $f_1(y_i)$ 、 $f_2(y_i)$ 分别为 ZINB 模型、NB 模型的概率密度函数。

从统计学的角度来看, V 值近似服从标准正态

分布, 当 $V \geq 1.96$ 时, 选择 ZINB 模型; 当 $V \leq -1.96$ 时, 选择 NB 模型; 当 $-1.96 < V < 1.96$ 时, 暂时无法判断哪种模型更优。

(2) 对数似然值

对数似然值又称拟合模型的对数似然值, 它可以通过最大似然法求出, 其取值范围为 $(-\infty, 0)$ 。采用最大似然法进行求解, 需要构建将观测数据的概率表述为未知模型参数的函数, 该函数称为似然函数。似然函数最大化的过程非常困难, 一般通过使似然函数的自然对数变化式最大的方法进行求解。因此, 对数似然值的绝对值越小, 表示模型的拟合效果越好。

(3) AIC 准则

AIC(Akaike's information criterion)准则由日本统计学家 Akaike 提出, 它建立在熵的概念基础上, 用来权衡估计模型的复杂程度与该模型拟合数据的优良性, 其计算公式为

$$A = -2\lg(L) + 2k \quad (8)$$

式中: A 为 AIC 准则值; L 为拟合模型的最大似然估计值; k 为拟合模型的参数数量。

AIC 准则鼓励数据拟合的优良性, 但是要尽量避免出现过度拟合的情况, 因此优先考虑 A 值最小的模型。A 值越小, 表示模型的拟合效果越好。

2.2.2 准确性检验

选择相对误差和累积残差 2 个指标进行模型的准确性检验, 各检验指标具体含义如下所示。

(1) 相对误差

相对误差越小, 则拟合模型的拟合准确性越好。相对误差 R_e 值的计算公式为

$$R_e = \left| \left(\sum_{i=1}^u y_i - \sum_{i=1}^u y'_i \right) / \sum_{i=1}^u y_i \right| \times 100\% \quad (9)$$

式中: y'_i 为第 i 个路段拟合模型路侧事故起数的预测值; u 为划分的路段单元总数。

(2) 累积残差

累积残差值越小, 表明拟合模型的拟合准确性越好。累积残差计算公式为

$$R_i = (y_i - y'_i) / (y'_i + Ky_i^2)^{-\frac{1}{2}} \quad (10)$$

$$C = \sum_{i=1}^u R_i \quad (11)$$

式中: R_i 为第 i 个路段的标准残差; C 为累积残差值。

2.3 弹性分析

采用弹性分析法分析、判断自变量的增加对因

变量产生的效用是递增或递减的趋势,以及这种趋势变化的幅度。则第 i 路段第 j 个自变量 x_{ij} 的弹性系数 $E_{x_{ij}}^{\lambda_i}$ 为

$$E_{x_{ij}}^{\lambda_i} = \partial \lambda_i / \partial x_{ij} (x_{ij} / \lambda_i) \tag{12}$$

在本文选择的自变量中,既有连续型自变量,还有离散型自变量。而连续型自变量和离散型自变量的求导方法不同,为了简化计算,第 j 个连续型自变量、离散型自变量的弹性系数 E_{c_j} 、 E_{d_j} 分别为

$$E_{c_j} = \beta_j \bar{x}_j \tag{13}$$

$$E_{d_j} = (e^{\beta_j} - 1) / e^{\beta_j} \tag{14}$$

式中: β_j 为第 j 个自变量的回归系数; \bar{x}_j 为第 j 个自变量的平均值。

3 实证性分析

3.1 基于定长法的事事故起数预测模型

3.1.1 自变量的选取

选择时间间隔为 1 年,分别统计每个路段的道路几何线形、每年的年交通事故起数和年平均日交通量,从而得到 300 个不同路段、不同时间间隔的样本数据。选择 11 个候选自变量,涵盖了平面线形(曲度、曲线比例、曲率变化率)、纵断面线形(平均纵坡坡度、相邻路段坡度差)、横断面线形(超高、车道数、左侧路肩宽度)和交通环境(弯坡组合、特殊路段、年平均日交通量),这些自变量的定义和统计描述分别见表 1、表 2^[22]。

表 1 自变量的定义

Tab. 1 Definition of independent variables

自变量	定义
曲度 D/m^{-1}	每 100 英尺曲线的弧度,当曲线半径 R 的单位为“m”时, $D=1\ 748/R$ 。
曲线比例/%	路段平面曲线长度所占比例
曲率变化率/ [$(^\circ) \cdot \text{km}^{-1}$]	平面线形中角度改变的绝对值与路线长度的比值
平均纵坡坡度/%	路段不同坡度的加权平均值
相邻路段坡度差/%	路段坡度与上游相邻路段坡度的差值
超高	如果路段内存在超高,用 1 表示;反之,用 0 表示。
车道数	路段包含的行车道数量
左侧路肩宽度/m	左侧硬路肩的宽度
弯坡组合/ $(\% \cdot \text{km}^{-1})$	路段纵坡坡度与平曲线半径之比的绝对值
特殊路段	如果路段中包含桥梁、隧道、服务区等特殊构造物,用 1 表示;反之,用 0 表示。
年平均日交通量/ ($10^4 \text{ veh} \cdot \text{d}^{-1}$)	1 年的交通量总和除以该年的总天数

表 2 基于定长法的自变量统计描述

Tab. 2 Statistical description of independent variables based on fixed-length segment method

自变量	最大值	最小值	平均值	标准差
曲度/ m^{-1}	1.75	0	0.49	0.38
曲线比例/%	1.00	0	0.75	0.35
曲率变化率/[$(^\circ) \cdot \text{km}^{-1}$]	49.15	0	14.94	9.94
平均纵坡坡度/%	2.32	-2.32	1.17	0.72
相邻路段坡度差/%	6.69	-5.70	-0.02	2.34
超高	1.00	0	0.38	0.49
车道数	3.00	2.00	2.74	0.44
左侧路肩宽度/m	0.75	0.75	0.75	0
弯坡组合/[$\% \cdot \text{km}^{-1}$]	2.00	0	0.40	0.39
特殊路段	1.00	0	0.70	0.46
年平均日交通量/($10^4 \text{ veh} \cdot \text{d}^{-1}$)	0.63	0.47	0.53	0.04

3.1.2 模型构建

采用极大似然估计法求解模型参数,取显著性水平 0.05,采用混合逐步选择法筛选显著自变量,借助统计分析软件 Stata12.0 完成求解。当自变量的 p 值小于 0.05 时,保留该自变量;反之,删除该自变量。NB 模型和 ZINB 模型回归系数的估计结果见表 3。从表 3 可以看出:①无论是 NB 模型,还是 ZINB 模型,曲度、曲线比例和车道数对路侧事故起数均有显著正影响,曲率变化率和平均纵坡坡度对路侧事故起数有显著负影响;②ZINB 模型的 Vuong 检验统计量为 3.07,大于临界值 1.96,NB 模型的对数似然值的绝对值和 AIC 准则值均大于 ZINB 模型,表明 NB 模型的拟合效果不如 ZINB 模型;③NB 模型的相对误差为 2.28%,大于 ZINB 模型的 1.68%,NB 模型的累积残差为 163.30,大于 ZINB 模型的 148.27,表明 ZINB 模型的预测准确性优于 NB 模型。

表 3 基于定长法的 NB 模型和 ZINB 模型回归系数的估计结果

Tab. 3 Estimation results of NB model and ZINB model based on fixed-length segment method

变量	NB 模型		ZINB 模型	
	回归系数	p 值	回归系数	p 值
截距	-8.386	0	-4.565	0.001
曲度	3.383	0.007	2.830	0.024
曲线比例	6.400	0	2.561	0.029
曲率变化率	-0.164	0	-0.136	0.003
平均纵坡坡度	-0.681	0	-0.601	0
车道数	1.070	0	1.029	0
Vuong 检验统计量			3.07	
对数似然值	-249.431		-242.308	
AIC 准则值	512.862		504.618	
相对误差/%	2.28		1.68	
累积残差	163.30		148.27	

基于定长法的路侧事故起数年均实际值、NB 模型和 ZINB 模型的年预测值在不同路段上的分布状况见图 1。

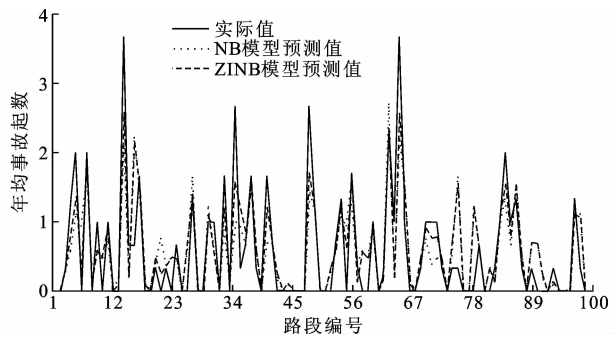


图 1 基于定长法的路侧事故起数实际值与预测值在不同路段的分布

Fig. 1 Actual values and prediction values of roadside accidents distributed on different segments based on fixed-length segment method

从图 1 可以看出:①在 100 个路段单元中,3 年中实际平均路侧事故起数最大值、最小值和平均值分别为 3.67、0 和 0.59;②NB 模型预测最大值、最小值和平均值分别为 2.58、0 和 0.63;③ZINB 模型预测最大值、最小值和平均值分别为 2.7、0 和 0.6;④ZINB 模型比 NB 模型的预测值更接近路侧事故起数的实际值,表明 ZINB 模型比 NB 模型能更好地反映路侧事故在不同路段上的分布特征。

3.1.3 弹性分析

根据式(13)、式(14)计算得到 5 个显著自变量对路侧事故起数的相对影响程度,见表 4。

表 4 基于定长法的弹性分析结果

Tab. 4 Elasticity analysis results based on fixed-length segment method

自变量	均值	系数	弹性系数
曲度	0.493	2.830	1.395
曲线比例	0.750	2.561	1.921
曲率变化率	14.938	-0.136	-2.031
平均纵坡坡度	1.171	-0.601	-0.704
车道数	2.740	1.029	2.819

从表 4 可以看出:①5 个显著自变量对路侧事故起数影响程度大小依次为车道数、曲率变化率、曲线比例、曲度和平均纵坡坡度;②在其他条件不变的情况下,当曲度、曲线比例、车道数分别增加 1%时,路侧事故起数分别增加 1.40%、1.92%、2.82%,当曲率变化率、平均纵坡坡度分别增加 1%时,路侧事故起数分别减少 2.03%、0.70%。

3.2 基于不定长法的事事故起数预测模型

3.2.1 自变量的选取

选择时间间隔为 1 年,统计每个路段的道路几何线形、每年的年交通事故起数和年平均日交通量,从而得到 462 个不同路段、不同时间间隔的样本数据。自变量的选取与第 3.1.1 节相同,见表 5。

表 5 基于不定长法的自变量统计描述

Tab. 5 Statistical description of independent variables based on homogeneous longitudinal grade segment method

自变量	最大值	最小值	平均值	标准差
曲度/ m^{-1}	1.75	0	0.45	0.36
曲线比例/%	1.00	0	0.76	0.36
曲率变化率/ $[(^{\circ}) \cdot km^{-1}]$	49.15	0	13.54	10.10
纵坡坡度/%	3.99	-3.80	1.17	1.58
相邻路段坡度差/%	6.69	-5.70	-0.14	2.85
超高	1.00	0	0.35	0.48
车道数	3.00	2.00	2.77	0.42
左侧路肩宽度/m	0.75	0.75	0.75	0
弯坡组合/ $(\% \cdot km^{-1})$	2.00	0	0.29	0.35
特殊路段	1.00	0	0.58	0.49
年平均日交通量/ $(10^4 veh \cdot d^{-1})$	0.63	0.47	0.53	0.05

3.2.2 模型构建

采用极大似然估计法求解模型参数,NB 模型和 ZINB 模型回归系数的估计结果见表 6。

表 6 基于不定长法的 NB 模型和 ZINB 模型回归系数的估计结果

Tab. 6 Estimation results of NB model and ZINB model based on homogeneous longitudinal grade segment method

变量	NB 模型		ZINB 模型	
	回归系数	p 值	回归系数	p 值
截距	-6.961	0	-6.629	0
曲线比例	4.059	0	3.996	0
纵坡坡度	-0.549	0	-0.523	0
车道数	0.879	0.002	0.936	0
弯坡组合	1.048	0.004	1.064	0.002
Vuong 检验统计量			4.530	
对数似然值	-343.128		-341.337	
AIC 准则值	698.256		696.724	
相对误差/%	3.370		2.240	
累积残差	240.130		230.040	

由表 6 可以看出:①无论是 NB 模型还是 ZINB 模型,曲线比例、弯坡组合和车道数对路侧事故起数均有显著正影响,纵坡坡度对路侧事故起数有显著负影响;② ZINB 模型的 Vuong 检验统计量值为 4.53,大于临界值 1.96,NB 模型的对数似然值的绝对值和 AIC 准则值均大于 ZINB 模型,表明 NB 模型的拟合效果不如 ZINB 模型;③NB 模型的相对误差为 3.37%,大于 ZINB 模型的 2.24%,NB 模型的

累积残差为 240.13,大于 ZINB 模型的 230.04,表明 ZINB 模型的预测准确性优于 NB 模型。

基于不定长法的路侧事故起数年实际值、NB 模型和 ZINB 模型年预测值在不同路段上的分布状况见图 2。

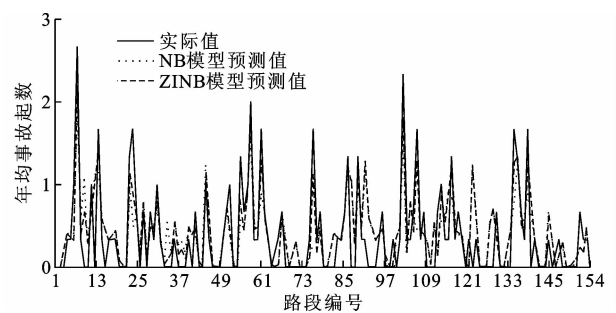


图 2 基于不定长法的路侧事故起数实际值与预测值的路段分布
Fig.2 Actual values and prediction values of roadside accidents distributed on different segments based on homogeneous longitudinal grade segment method

从图 2 可以看出:①在 154 个路段单元中,3 年中实际平均路侧事故起数最大值、最小值和平均值分别为 2.67、0 和 0.39;②NB 模型预测最大值、最小值和平均值分别为 1.89、0 和 0.44;③ZINB 模型预测最大值、最小值和平均值分别为 2.22、0 和 0.41;④ZINB 模型的预测值比 NB 模型更接近路侧事故起数的实际值,表明 ZINB 模型比 NB 模型能更好地反映路侧事故在不同路段上的分布特征。

3.2.3 弹性分析

根据式(11)、式(12)计算得到 4 个显著自变量对路侧事故起数的相对影响程度,见表 7。

表 7 基于不定长法的弹性分析结果

Tab.7 Elasticity analysis results based on homogeneous longitudinal grade segment method			
自变量	均值	系数	弹性系数
曲线比例	0.762	3.996	3.045
纵坡坡度	1.170	-0.523	-0.612
车道数	2.766	0.936	2.589
弯坡组合	0.288	1.064	0.306

从表 7 可以看出:①4 个显著自变量对路侧事故起数影响程度大小依次为曲线比例、车道数、纵坡坡度和弯坡组合;②在其他条件不变的情况下,当曲线比例、车道数、弯坡组合分别增加 1%时,路侧事故起数分别增加 3.05%、2.59%、0.30%,当纵坡坡度增加 1%时,路侧事故起数减少 0.61%。

3.3 对比分析

2 种路段单元划分方法下路侧事故起数的 ZINB 模型拟合效果见表 8。

表 8 两种方法下路侧事故起数的 ZINB 模型拟合效果

Tab.8 Fitting results of ZINB model with two segment methods

路段划分方法	相对误差/%	累积残差
定长法	1.68	148.27
不定长法	2.24	230.04

从表 8 可以看出:①定长法的相对误差为 1.68%,而不定长法的相对误差为 2.24%;②定长法的累积残差为 148.27,而不定长法的累积残差为 230.04;③基于定长法的路侧事故起数 ZINB 模型的拟合准确性优于不定长法。

4 结 语

(1)无论采用定长法还是不定长法划分路段单元,从模型的拟合优度检验与拟合准确性来看,ZINB 模型均优于 NB 模型;从弹性分析结果来看,曲线比例和车道数都对路侧事故起数有显著正影响。

(2)基于定长法的路侧事故起数 ZINB 模型的拟合准确性优于不定长法。

(3)由于数据采集的局限性,没有收集到气象条件、行车速度、照明条件等统计数据,因此本文没有考虑这些变量对路侧事故起数的影响,这将是下一步研究的重点。

参考文献:

References:

[1] 公安部交通管理局. 中华人民共和国道路交通事故统计年报(2011 年度)[R]. 无锡:公安部交通管理科学研究所,2012.
Traffic Management Bureau of the Ministry of Public Security. Road traffic accident statistic annual report of the People's Republic of China (2011)[R]. Wuxi: Traffic Management Research Institute of the Ministry of Public Security,2012.

[2] ZEGGER C V,REINFURT D W. Accident effects of side slope and other roadside feature on two-lane roads[J]. Transportation Research Record, 1988 (1195):33-47.

[3] DUMBAUGH E. Design of safe urban roadside: An empirical analysis[J]. Transportation Research Record,2006(1961):74-82.

[4] 李长城,阚伟生. 我国双车道等级公路路侧事故规律统计分析[J]. 道路交通与安全,2006,6(10):15-18.
LI Chang-cheng,KAN Wei-sheng. Analysis on roadside accidents of two-lane highway[J]. Road Traffic and Safety,2006,6(10):15-18.

- [5] 张铁军,唐铮铮.我国双车道公路普通路段路侧事故预测模型研究[J].公路交通科技,2008,25(4):147-150,158.
ZHANG Tie-jun, TANG Cheng-cheng. Research of roadside accident prediction models for common segment of two-lane roads in China[J]. Journal of highway and transportation research and development, 2008,25(4):147-150,158.
- [6] 张铁军,唐铮铮.山区双车道公路事故预测模型路侧危险度事故修正系数研究[J].公路,2007(9):116-122.
ZHANG Tie-jun, TANG Cheng-cheng. A study on accident modification factor about roadside hazard rating in accident prediction models for two-lane highway in mountainous area[J]. Highway, 2007(9):116-122.
- [7] 游克思,孙璐,顾文钧.山区公路路侧安全量化评价[J].交通运输工程与信息学报,2010,8(3):49-55.
YOU Ke-si, SUN Lu, GU Wen-jun. Quantitative assessment of roadside safety on mountain highway[J]. Journal of Transportation Engineering and Information, 2010,8(3):49-55.
- [8] 刘彦,陈健蕾,刘志刚,等.山区高速公路路侧风险等级评价方法研究[J].重庆交通大学学报:自然科学版,2015,34(1):87-90.
LIU Yan, CHEN Jian-lei, LIU Zhi-gang, et al. Assessment method for roadside risk grade of freeway in mountainous areas[J]. Journal of Chongqing Jiaotong University: Natural Science, 2015,34(1):87-90.
- [9] LORD D, MANNERING F. The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives[J]. Transportation Research Part A, 2010,44(5):291-305.
- [10] SHANKAR V, MILTON J, MANNERING F L. Modeling accident frequency as zero-altered probability processes: An empirical inquiry[J]. Accident Analysis and Prevention, 1997,29(6):829-837.
- [11] QIN X, IVAN J N, RAVISHANKAR N. Selecting exposure measures in crash rate prediction for two-lane highway segments[J]. Accident Analysis and Prevention, 2004,36(2):183-191.
- [12] KUMARA S P, CHIN H C. Modeling accident occurrence at signalized tee intersections with special emphasis on excess zeros[J]. Traffic Injury Prevention, 2003,4(1):53-57.
- [13] DANIEL J, CHIEN S I J. Truck safety factors on urban arterials[J]. Journal of Transportation Engineering, 2004,130(6):742-752.
- [14] EASA S M, YOU Q C. Collision prediction models for three-dimensional two-lane highways: Horizontal curves[J]. Transportation Research Record, 2009(2092):48-56.
- [15] 高海龙,阚伟生,李长城,等.双车道等级公路路侧事故预测模型研究[J].公路交通科技,2007,24(7):127-131.
GAO Hai-long, KAN Wei-sheng, LI Chang-cheng, et al. Roadside accident prediction model for two-lane highway[J]. Journal of Highway and Transportation Research and Development, 2007,24(7):127-131.
- [16] 唐铮铮,张铁军,何勇.双车道公路交通安全性预测基础模型[J].同济大学学报:自然科学版,2008,36(9):1199-1204.
TANG Cheng-cheng, ZHANG Tie-jun, HE Yong. Basic model of safety performance prediction for two-lane highways[J]. Journal of Tongji University: Natural Science, 2008,36(9):1199-1204.
- [17] AYATI E, ABBASI E. Modeling accidents on Mashhad urban highways[J]. Open Journal of Safety Science and Technology, 2014,4(1):22-35.
- [18] 罗石贵,周伟.道路交通事故多发点鉴定方法探讨[J].西安公路交通大学学报,1999,19(1):30-33.
LUO Shi-gui, ZHOU Wei. Research on the identification of road accident blackspots[J]. Journal of Xi'an Highway University, 1999,19(1):30-33.
- [19] HAUER E, KONOV J, ALLERY B, et al. Screening the road network for sites with promise[J]. Transportation Research Record, 2002(1784):27-32.
- [20] ZHANG C, IVAN J N. Effects of geometric characteristics on head-on crash incidence on two-lane roads in Connecticut[J]. Transportation Research Record, 2005(1908):159-164.
- [21] PARDILO J M, LIAMAS R. Relevant variables for crash rate prediction in Spain's two lane rural roads[R]. Washington DC: National Research Council, 2003.
- [22] 马壮林,邵春福,胡大伟,等.高速公路交通事故起数时空分析模型[J].交通运输工程学报,2012,12(2):93-99.
MA Zhuang-lin, SHAO Chun-fu, HU Da-wei, et al. Temporal-spatial analysis model of traffic accident frequency on expressway[J]. Journal of Traffic and Transportation Engineering, 2012,12(2):93-99.