

考虑异质性的高速公路客车交通事故频次 影响因素分析

曹振,刘洪瑜*,田准

西安建筑科技大学城市发展与现代交通学院,陕西 西安 710055

摘要:为提高高速公路客车交通安全,降低交通事故率,采集我国华东地区3条高速公路部分路段的客车事故频次、交通量、道路特征、气象条件和节假日占比共2320条样本数据,考虑事故异质性,建立泊松-对数正态(Poisson-lognormal, PLN)模型、随机参数泊松-对数正态(random parameters Poisson-lognormal, RP-PLN)模型和随机效应泊松-对数正态(random effects Poisson-lognormal, RE-PLN)模型,采用贝叶斯方法和马尔科夫链蒙特卡罗仿真估计模型参数的后验分布,由建模结果定量分析各因素对高速公路客车交通事故频次的影响程度。结果表明:RP-PLN模型的拟合效果优于PLN模型和RE-PLN模型;路段长度、交通量和节假日占比与高速公路客车交通事故频次正相关,温度、相对湿度和能见度与高速公路客车交通事故频次负相关。当其他因素不变时,路段长度、交通量分别增加1%时,高速公路客车交通事故频次分别增大0.90%、0.62%,节假日占比增加1个单位时,高速公路客车交通事故频次增大8.30%;温度、相对湿度和能见度分别增加1个单位时,高速公路客车交通事故频次分别减小5.72%、8.82%、6.55%。

关键词:高速公路;客车交通事故;异质性;RP-PLN模型

中图分类号:U491.31

文献标志码:A

文章编号:1672-0032(2025)03-0012-10

引用格式:曹振,刘洪瑜,田准.考虑异质性的高速公路客车交通事故频次影响因素分析[J].山东交通学院学报,2025,33(3):12-21.

CAO Zhen, LIU Hongyu, TIAN Zhun. Analysis of factors influencing highway bus accidents frequency considering heterogeneity[J]. Journal of Shandong Jiaotong University, 2025, 33(3): 12-21.

0 引言

随着我国高速公路网络的持续扩展及机动车保有量的持续增长,高速公路交通安全问题日益凸显。截至2023年底,我国高速公路总里程达18.36万km,载客汽车保有量增至29395.59万辆,占民用汽车保有量的89.3%^[1]。近年来,高速公路客车事故呈上升趋势,此类事故往往造成重大人员伤亡和巨额财产损失,对社会稳定和公共安全构成严重威胁^[2]。研究高速公路客车事故的影响因素及规律,对提高道路交通安全水平具有重要意义。

目前交通安全研究多依托历史交通事故数据构建预测模型,分析各类风险因素与事故频次间的关系。泊松回归模型形式简洁、解释性强,应用广泛,但无法处理过离散性的实际事故数据,导致模型拟合效果和预测精度有限^[3]。为克服泊松回归模型的局限性,相继改进得到泊松-伽玛模型^[4-5]和泊松-对数正态(Poisson-lognormal, PLN)模型^[6-7]。泊松-伽玛模型通过伽玛分布调整泊松分布的均值,简化后验推

收稿日期:2024-06-13

基金项目:陕西省重点研发计划(2020SF-390)

第一作者简介:曹振(1978—),男,西安人,教授,硕士研究生导师,工学博士,主要研究方向为城市公共交通空间综合规划与开发,E-mail:653102531@qq.com。

*通信作者简介:刘洪瑜(1999—),女,山东聊城人,硕士研究生,主要研究方向为交通运输规划与管理、交通安全,E-mail:1984766494@qq.com。

断过程,可更好地捕捉未观测因素引起的频次变化。PLN模型假设误差项服从对数正态分布,能处理异常值和过离散性,提高模型的鲁棒性。未观测因素与已观测因素间的交互作用导致样本间存在安全效应异质性,而上述固定参数模型无法准确表征异质性,因此引入随机效应和随机参数模型。随机效应模型通过嵌入一个随机误差项总体表征样本间异质性,提高拟合度和参数估计的准确性^[8]。随机参数模型允许回归系数在不同样本中发生变化,可更好地反映各样本中存在的未观测异质性。

在交通事故频次影响因素方面,现有研究主要关注气象条件、交通状况、道路特征和路面状况等因素对交通事故频次的影响。曾强等^[9]发现风速及降水量与交通事故频次显著正相关。温惠英等^[10]、孙佳人^[11]依据现行机动车类型划分标准,对广东省开阳高速公路通行车辆实施五级分类,通过建立回归模型,分析不同车型构成比例与交通事故频次的定量关系。郭延永等^[12]构建信号交叉口交通冲突模型,分析直行交通量、右转交通量、渠化岛、右转让行标志和右转半径等对交叉口交通冲突的影响。漆巍巍等^[13]基于空间溢出效应模型得出路段长度、圆曲线长度等道路几何特征等与交通事故频次显著相关。蔡明懋^[14]考虑数据异质性,建立事故频数模型,从隧道特征、道路设计特征、交通状况、路面性能、天气状况等方面定量评价高速公路隧道交通安全。唐峰^[15]以广东省山区高速公路为研究对象,构建包括道路设计特征、交通状况、路面性能和天气条件的常规路段和隧道路段多维事故数据集,评价高速公路安全性。Huo等^[16]构建随机参数模型,发现在服务区及分、合流立交区的交通事故频次更高。侯芹忠^[17]基于随机参数模型分析交通事故频次的影响因素,发现交通量与交通流组成因素与交通事故频次显著相关。Zhao等^[18]研究月度天气(温度和降水等)变化与高速公路交通事故间的关系,结果显示恶劣天气条件会增大事故发生风险,低温条件下交通事故频次也会增大。Hou等^[19]分析交通流特征与高速公路几何设计对交通事故频次的影响。Hou等^[20]采用随机效应负二项和随机参数负二项模型研究货车比例、爬坡车道、雾霾天气等因素对高速公路事故频次的影响。Wen等^[21]分析道路特征与天气条件对高速公路事故的主效应和交互作用,发现风速与纵坡、降水与弯道等特定组合会显著影响事故风险。节假日期间高速公路超饱和运行和不稳定驾驶行为导致交通拥堵,易引发交通事故^[22-24],但已有研究未充分考虑节假日天数对交通事故频次的潜在影响。

本文基于高速公路客车交通事故数据,结合交通量、道路特征、气象数据及节假日天数等数据,分析高速公路客车事故频次影响因素。构建 PLN 模型、随机参数泊松-对数正态(random parameters Poisson-lognormal, RP-PLN)模型、随机效应泊松-对数正态(random effects Poisson-lognormal, RE-PLN)模型分析高速公路客车交通事故频次的影响因素,通过贝叶斯估计方法标定模型参数,定量评估各因素对高速公路客车交通事故频次的影响程度,以期为交通安全政策制定和管理决策提供科学依据。

1 数据采集与变量检验

1.1 数据采集

以我国华东地区3条高速公路部分路段上2021年5月至2022年12月的客车交通事故数据为研究样本。研究路段总长约206.2 km,从相关交通管理部门获取的高速公路路线几何特征数据中,提取高速公路的曲线半径和坡度等基本信息。基于路线几何要素,将3条高速公路划分为116个同质性路段,每个路段的曲率和坡度均相同;研究路段共有16个收费站,15个区段,每个区段的交通构成与交通量相同,通过该地区高速公路联网收费系统,采集各交通区段的月度总交通量及交通组成数据,根据文献[25]的车型分类标准,将车辆划分为1~6类,以对应换算系数1.0、1.5、2.0、3.0、3.5、4.0将其分别换算为标准车,计算月平均日交通量及各类车型占比;通过中国气象数据网获取交通事故发生时的气象数据;通过国务院办公厅发布的节假日安排通知确定月节假日占比。

将统计得到的道路特征、交通状况、气象条件及节假日占比作为客车交通事故频次 PLN、RP-PLN、RE-PLN 模型的影响因素集,共计2 320条样本数据。客车交通事故频次模型变量的描述性统计结果如表1所示。

表 1 客车交通事故频次模型变量的描述性统计结果

变量	变量定义及单位	最小值	最大值	均值	标准差
客车交通事故频次	路段在某月内发生的客车交通事故数,起/月	0	18.00	0.29	0.93
交通量	路段的月平均日标准车流量,10 ³ pcu/d	0.65	30.64	10.52	4.85
1类车占比	路段的月均1类车数量占有所有类型车总数之比, %	18.47	90.13	53.11	13.32
2类车占比	路段的月均2类车数量占有所有类型车总数之比, %	3.12	15.05	9.40	2.38
3类车占比	路段的月均3类车数量占有所有类型车总数之比, %	6.13	50.06	28.43	9.01
4类车占比	路段的月均4类车数量占有所有类型车总数之比, %	0.99	11.76	4.82	1.86
5类车占比	路段的月均5类车数量占有所有类型车总数之比, %	0	3.55	1.13	0.39
6类车占比	路段的月均6类车数量占有所有类型车总数之比, %	0.01	29.05	9.04	3.88
路段长度	路段的长度,km	0.31	4.43	1.43	0.70
路基宽度	路段的路基宽度,m	26.00	33.50	27.05	2.30
车道数	路段车道数	4.00	6.00	4.21	0.61
曲率	路段平面曲线半径的倒数,km ⁻¹	0	0.91	0.20	0.22
坡度	路段纵向坡度, %	0	3.40	0.98	0.68
温度	月平均温度,℃	3.80	31.20	18.58	8.60
相对湿度	月平均相对湿度, %	64.00	87.00	76.84	5.73
能见度	月平均能见度,km	5.40	38.10	16.17	5.73
降水量	月平均日降水量,mm	0	12.18	3.13	2.90
风速	月平均风速,km/h	5.90	18.00	8.63	1.83
异常天气占比	月雨、雪、雾、霾天气数与该月总天数之比, %	0	58.06	27.80	16.61
日照时数	月平均日照时数,h/d	10.08	14.11	12.29	1.34
空气质量优、良率	月空气质量优、良天数与该月总天数之比, %	6.67	100.00	56.16	26.75
节假日占比	月节假日天数与该月总天数之比, %	0	22.58	8.06	8.13

1.2 变量共线性检验

采用软件 SPSS 对模型的解释变量进行相关性检验,结果如图 1 所示。

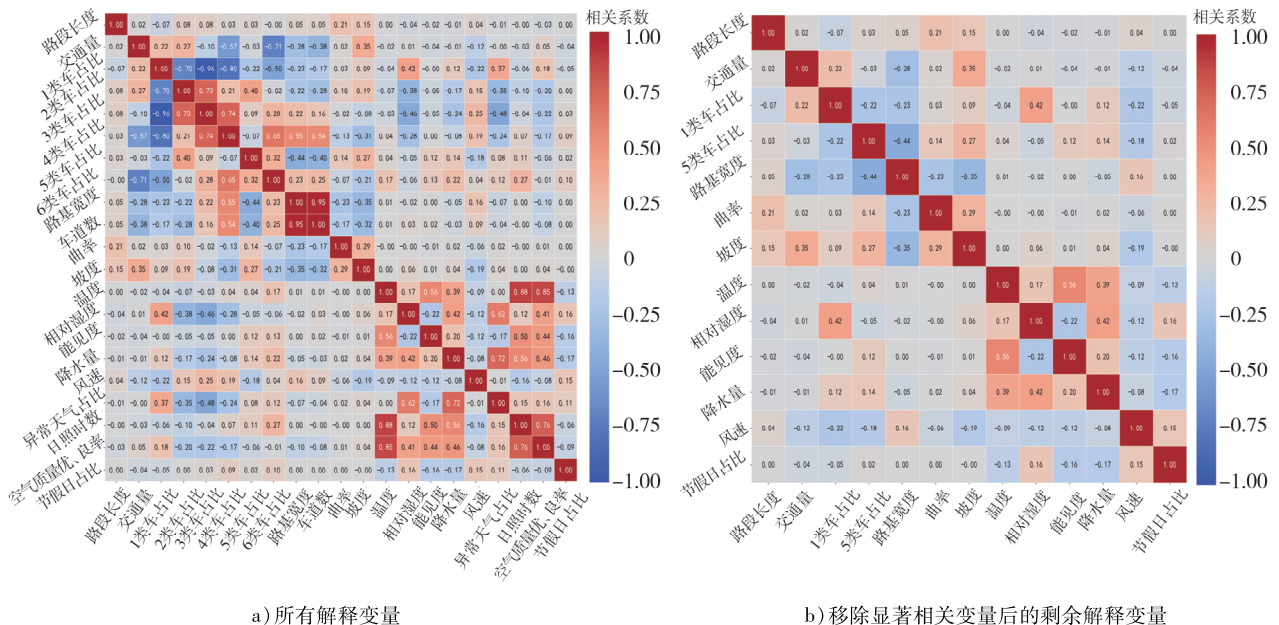


图 1 解释变量的相关性热力图

由图 1a)可知:部分变量间的皮尔逊相关系数绝对值大于 0.6,说明这些变量间显著相关。为避免显著相关变量对参数估计的影响,从解释变量中移除 2、3、4、6 类车占比、车道数、日照时数、空气质量优、良率、异常天气占比等,对剩余解释变量进行多重共线性诊断,相关性热力图如图 1b)所示,未发现显著共线性变量,因此将剩余解释变量用于构建高速公路客车交通事故频次模型。

2 模型建立与评价

2.1 模型建立

2.1.1 PLN 模型

PLN 模型在标准泊松回归的基础上引入服从对数正态分布的随机误差项,提高模型的灵活性和拟合能力。

设 Y_i 为路段 i 在观测期内可能发生的客车交通事故频次,服从泊松分布,其概率分布函数为:

$$P(Y_i = k) = \frac{\lambda_i^k e^{-\lambda_i}}{k!},$$

式中: k 为路段 i 实际记录的客车交通事故频次, λ_i 为路段 i 的期望客车交通事故频次。

设 λ_i 与交通事故影响因素(解释变量)间存在的函数关系为:

$$\ln \lambda_i = \beta_0 + \sum_{m=1}^M \beta_m x_{m,i} + \epsilon_i,$$

式中: β_0 为常数项; β_m 为第 m 个解释变量的系数; $x_{m,i}$ 为第 m 个解释变量在路段 i 上的观测值; M 为解释变量总个数; ϵ_i 为随机误差项, ϵ_i 服从均值为 0、精度为 τ_ϵ 的正态分布,即 $\epsilon_i \sim N(0, 1/\tau_\epsilon)$ 。

针对本文选取的解释变量,PLN 模型的公式为:

$$\ln \lambda_i = \beta_0 + \beta_1 \ln L_i + \beta_2 \ln T_i + \sum_{m=3}^M \beta_m x_{m,i} + \epsilon_i,$$

式中: β_1 为 L_i 的系数, L_i 为路段 i 的路段长度, β_2 为 T_i 的系数, T_i 为路段 i 的交通量。

2.1.2 RP-PLN 模型

为研究不同观测样本中处理交通事故影响因素作用差异的问题,假设 PLN 模型的解释变量系数非固定值^[12],而是服从特定概率分布,进而构建 RP-PLN 模型,其公式为:

$$\ln \lambda_i = \beta_{0,i} + \sum_{m=1}^M \beta_{m,i} x_{m,i} + \epsilon_i,$$

式中 $\beta_{j,i} (j=0,1,\dots,m)$ 为解释变量的系数,服从均值为 β_j 、精度为 τ_{β_j} 的正态分布,即 $\beta_{j,i} \sim N(\beta_j, 1/\tau_{\beta_j})$ 。

2.1.3 RE-PLN 模型

描述交通事故的差异性问题时,也可在 PLN 模型中引入随机效应项,构建 RE-PLN 模型,公式为:

$$\ln \lambda_i = \beta_0 + \sum_{m=1}^M \beta_m x_{m,i} + \mu_i + \epsilon_i,$$

式中: μ_i 为随机效应项, μ_i 服从均值为 0、精度为 τ_μ 的正态分布,即 $\mu_i \sim N(0, 1/\tau_\mu)$ 。

为定量分析解释变量对客车交通事故频次的影响程度,根据模型参数估计结果,计算各解释变量的发生率比值(incidence rate ratio, IRR)^[9]。路段长度的发生率比值 $E_{\text{IRR},L}$ 和交通量的发生率比值 $E_{\text{IRR},T}$ 分别表示其他解释变量保持不变时,路段长度和交通量每增加 1%,期望客车交通事故频次扩大的倍数^[11],公式分别为:

$$E_{\text{IRR},L} = \frac{E(Y_i | X_i, T_i, 1.01L_i)}{E(Y_i | X_i, T_i, L_i)} = 1.01^{\beta_1}, \quad (1)$$

$$E_{\text{IRR},T} = \frac{E(Y_i | X_i, L_i, 1.01T_i)}{E(Y_i | X_i, L_i, T_i)} = 1.01^{\beta_2}, \quad (2)$$

式中 X_i 为由所有解释变量在路段 i 上的观测值的集合。

除交通量与路段长度以外的解释变量 $x_{m,i}$ 的发生率比值 $E_{IRR,m}$ 表示其他解释变量保持不变时, $x_{m,i}$ 每增加 1 个单位,期望客车交通事故频次扩大的倍数,即:

$$E_{IRR,m} = \frac{E(Y_i | \dot{X}_i, T_i, L_i, x_{m,i} + 1)}{E(Y_i | \dot{X}_i, T_i, L_i, x_{m,i})} = e^{\beta_m}, \quad (3)$$

式中 $\dot{X}_i$ 为除 $x_{m,i}$ 外的其他解释变量在路段 i 上的观测值的集合。

2.2 贝叶斯参数估计

贝叶斯方法不依赖模型的闭式似然函数,能处理复杂模型的参数估计问题,已成为交通安全统计分析领域的主要参数估计工具。WinBUGS 是进行贝叶斯推断最常用的软件之一,该软件采用马尔科夫链蒙特卡洛(Markov chain Monte Carlo, MCMC)仿真和梅特罗波利斯-黑斯廷斯(Metropolis-Hastings)算法,根据参数的先验分布和观测数据,推断参数的后验分布^[9]。

采用贝叶斯方法估计 PLN、RP-PLN、RE-PLN 模型中的参数。设定 2 条链进行 20 000 次迭代,将前 4 000 次迭代作为燃烧样本。回归系数的先验分布采用方差较大的正态分布 $N(0, 10^3)$ ^[26]。随机误差项和随机效应项的精度(方差的倒数)被赋予先验伽马分布 $\text{Gamma}(1, 0.1)$ 。根据软件 WinBUGS 中的贝叶斯-格尔曼-鲁宾(Bayesian-Gelman-Rubin, BGR)统计量判断 MCMC 仿真的收敛情况,当 BGR 统计量小于 1.2 时认为 MCMC 算法已收敛^[27];也可通过迭代历史图判断模型收敛情况。模型参数均达到收敛,部分模型参数收敛情况如图 2、3 所示。

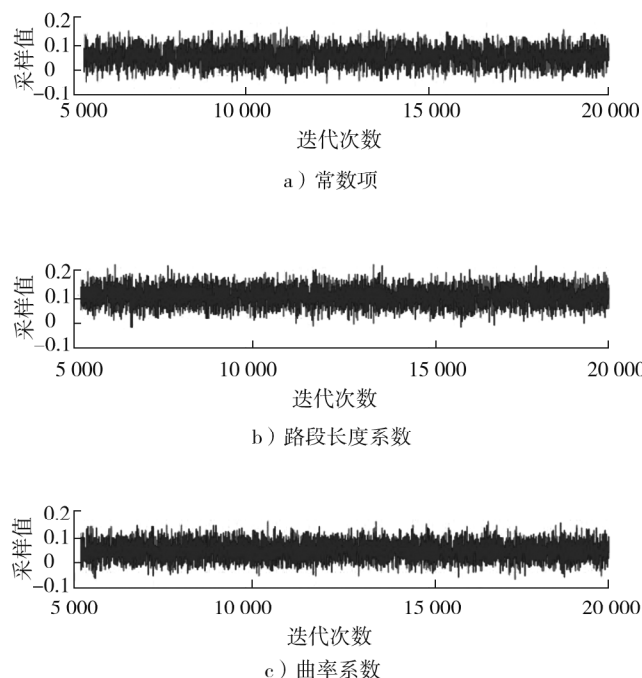


图 2 常数项、路段长度系数及曲率系数的迭代历史图

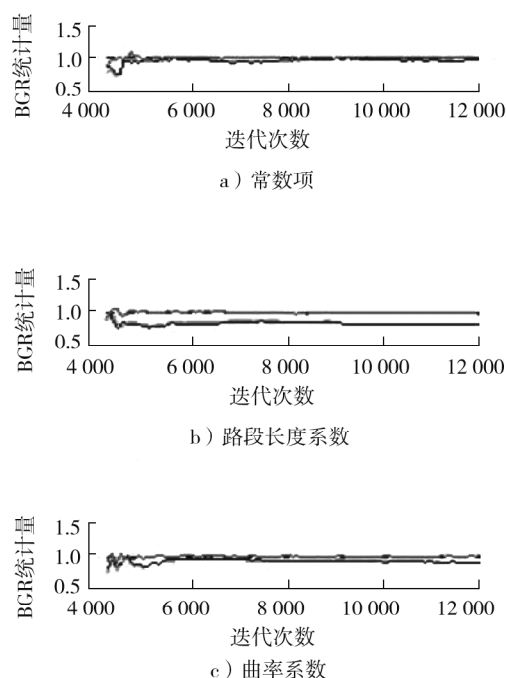


图 3 常数项、路段长度系数及曲率系数的 BGR 历史图

2.3 模型评价指标

采用偏差信息准则(deviance information criterion, DIC) E_{DIC} 、平均绝对偏差(mean absolute deviation, MAD) E_{MAD} 、均方预测误差(mean squared prediction error, MSPE) E_{MSPE} 三个指标比较 PLN 模型、RE-PLN 模型、RP-PLN 模型的性能^[9]。

采用 E_{DIC} 综合评估模型的适应程度和复杂性,公式为:

$$E_{DIC} = \bar{D} + p_D,$$

式中: $\bar{D}$ 为后验平均偏差,用作模型拟合度的度量指标; p_D 为参数的有效数量,用于评估模型的复杂度。

通常 E_{DIC} 越小表示模型表现越佳,若模型间的 E_{DIC} 差异大于 5,则可推断模型间存在显著差异^[8,11]。 E_{MAD} 通过绝对值平处理所有误差且对异常值不敏感, E_{MSPE} 在模型训练中更关注较大误差且对异常值敏感。将 E_{MAD} 和 E_{MSPE} 结合使用能全面评估模型性能,既了解一般误差水平,又识别较大误差,计算公式分别为:

$$E_{MAD} = \frac{1}{N} \sum_{i=1}^N |Y_i - \lambda_i|,$$

$$E_{MSPE} = \frac{1}{N} \sum_{i=1}^N (Y_i - \lambda_i)^2,$$

式中 N 为路段总数。

3 模型结果分析

3.1 模型对比分析

PLN 模型、RE-PLN 模型、RP-PLN 模型的贝叶斯参数估计结果如表 2~4 所示,评价结果如表 5 所示。

表 2 PLN 模型的贝叶斯参数估计结果

参数	均值	90% I_{BC}	95% I_{BC}
常数项	-0.004 7	(-0.056 7, 0.046 4)	(-0.065 5, 0.058 0)
路段长度系数	0.059 6	(0.008 6, 0.110 9)	(-0.001 6, 0.120 4)
交通量系数	0.054 3	(0.005 5, 0.102 8)	(-0.004 8, 0.112 3)
1 类车占比系数	0.015 6	(0.008 5, 0.023 7)	(0.007 5, 0.025 3)
5 类车占比系数	0.003 9	(-0.046 4, 0.053 0)	(-0.056 0, 0.061 1)
路基宽度系数	-0.083 3	(-0.113 7, -0.051 8)	(-0.119 0, -0.047 7)
曲率系数	-0.007 8	(-0.060 1, 0.041 8)	(-0.068 4, 0.050 9)
坡度系数	0.034 8	(-0.014 8, 0.084 5)	(-0.023 8, 0.093 2)
温度系数	-0.023 8	(-0.036 7, -0.010 6)	(-0.039 2, -0.007 9)
相对湿度系数	-0.004 4	(-0.018 5, 0.009 6)	(-0.020 3, 0.011 6)
能见度系数	-0.018 6	(-0.037 3, -0.000 0)	(-0.041 3, 0.002 7)
降水量系数	0.010 5	(-0.020 7, 0.042 2)	(-0.027 9, 0.048 0)
风速系数	-0.036 6	(-0.071 5, -0.001 6)	(-0.078 5, 0.005 4)
节假日占比系数	0.052 7	(0.041 7, 0.063 8)	(0.039 6, 0.065 9)
τ_e	0.701 8	(0.575 3, 0.846 8)	(0.556 5, 0.877 8)

注: I_{BC} 为贝叶斯置信区间(Bayesian confidence interval);黑体为在置信区间内显著。

表 3 RE-PLN 模型的贝叶斯参数估计结果

参数	均值	90% I_{BC}	95% I_{BC}
常数项	-0.004 5	(-0.057 2, 0.047 0)	(-0.067 3, 0.058 0)
路段长度系数	0.060 0	(0.009 5, 0.110 0)	(0.001 0, 0.119 9)
交通量系数	0.053 2	(0.006 6, 0.101 3)	(-0.002 6, 0.110 5)
1 类车占比系数	0.015 7	(0.007 7, 0.023 9)	(0.006 1, 0.025 7)
5 类车占比系数	0.003 8	(-0.047 7, 0.054 5)	(-0.058 0, 0.063 5)
路基宽度系数	-0.082 4	(-0.113 1, -0.054 4)	(-0.118 5, -0.051 0)
曲率系数	-0.007 6	(-0.058 4, 0.043 6)	(-0.068 7, 0.052 6)

表 3(续)

变量	均值	90% I_{BC}	95% I_{BC}
坡度系数	0.034 8	(-0.012 7, 0.083 8)	(-0.023 3, 0.094 7)
温度系数	-0.024 3	(-0.038 0, -0.010 3)	(-0.040 5, -0.007 8)
相对湿度系数	-0.004 9	(-0.018 4, 0.008 6)	(-0.020 2, 0.009 8)
能见度系数	-0.018 0	(-0.037 9, 0.000 4)	(-0.041 1, 0.003 6)
降水量系数	0.011 5	(-0.019 0, 0.042 8)	(-0.024 9, 0.048 2)
风速系数	-0.035 9	(-0.071 6, 0.000 2)	(-0.078 1, 0.008 0)
节假日占比系数	0.052 9	(0.041 9, 0.064 2)	(0.039 7, 0.066 4)
τ_{ε}	0.888 9	(0.641 7, 1.336 0)	(0.617 3, 1.474 0)
τ_{μ}	6.687 0	(1.648 0, 19.930 0)	(1.484 0, 25.910 0)

表 4 RP-PLN 模型的贝叶斯参数估计

参数	均值	90% I_{BC}	95% I_{BC}
常数项	-0.004 6	(-0.058 6, 0.047 0)	(-0.004 6, 0.057 2)
路段长度系数	0.895 9	(0.470 8, 1.310 0)	(0.386 6, 1.381 0)
路段长度系数分布标准差	0.301 7	(0.188 7, 0.481 6)	(0.181 4, 0.539 2)
交通量系数	0.617 6	(0.162 3, 1.062 0)	(-0.077 8, 1.128 0)
交通量系数分布标准差	0.301 7	(0.188 7, 0.481 6)	(0.181 4, 0.539 2)
1 类车占比系数	0.014 6	(-0.038 9, 0.042 9)	(-0.064 8, 0.044 8)
5 类车占比系数	0.000 2	(-0.048 8, 0.048 5)	(-0.059 3, 0.057 0)
路基宽度系数	-0.031 8	(-0.081 2, 0.030 0)	(-0.086 3, 0.043 7)
曲率系数	-0.000 2	(-0.052 0, 0.050 5)	(-0.052 0, 0.050 5)
坡度系数	0.004 0	(-0.045 8, 0.054 4)	(-0.054 9, 0.062 7)
温度系数	-0.058 9	(-0.088 8, -0.028 8)	(-0.094 3, -0.025 1)
温度系数分布标准差	0.087 4	(0.066 5, 0.136 7)	(0.064 5, 0.178 6)
相对湿度系数	-0.092 3	(-0.125 7, -0.062 7)	(-0.158 5, -0.059 6)
相对湿度系数分布标准差	0.043 3	(0.036 2, 0.071 3)	(0.035 6, 0.097 1)
能见度系数	-0.067 8	(-0.102 4, -0.023 8)	(-0.107 2, -0.017 0)
能见度系数分布标准差	0.088 7	(0.070 5, 0.131 5)	(0.068 4, 0.142 4)
降水量系数	-0.000 9	(-0.052 2, 0.039 1)	(-0.062 4, 0.044 5)
风速系数	-0.022 6	(-0.066 7, 0.029 3)	(-0.074 1, 0.037 0)
节假日占比系数	0.079 8	(0.044 1, 0.110 1)	(0.036 8, 0.115 8)
节假日占比系数分布标准差	0.104 3	(0.080 7, 0.160 5)	(0.078 4, 0.190 7)
τ_{ε}	13.880 0	(4.311 0, 28.070 0)	(3.439 0, 30.400 0)

由表 3、4 可知:RE-PLN 模型的 τ_{μ} 及 RP-PLN 模型的显著变量的参数均服从正态分布,表明 RE-PLN 模型、RP-PLN 模型均能有效处理事故的异质性。由表 5 可知:RP-PLN 模型的 p_D 比 RE-PLN 模型、PLN 模型稍大,但前者的 $\bar{D}$ 、 E_{DIC} 、 E_{MAD} 、 E_{MSPE} 均明显小于 RE-PLN 模

表 5 PLN、RP-PLN、RE-PLN 模型的评价结果

模型	$\bar{D}$	p_D	E_{DIC}	E_{MAD}	E_{MSPE}
PLN	2 228.790	422.614	2 651.400	0.262	0.151
RP-PLN	1 675.750	516.325	2 192.080	0.194	0.093
RE-PLN	2 224.430	424.166	2 648.600	0.262	0.150

型和 PLN 模型,说明 RP-PLN 模型拟合程度较高,模型较优。

3.2 影响因素分析

由于 RP-PLN 模型的拟合度更佳,根据 RP-PLN 模型的系数估计结果,由式(1)~(3)计算解释变量的发生率比值 E_{IRR} ,计算结果如表 6 所示。根据表 4、6 分析各解释变量对高速公路客车交通事故频次的影响程度。

路段长度和交通量的 90% 贝叶斯置信区间未包含 0 且后验均值均为正,二者与客车交通事故频次显著正相关,与以往研究结论基本一致^[11,13]。路段长度的发生率比值为 1.009 0,路段长度每增加 1%,客车交通事故频次增大 0.90%;交通量的发生率比值为 1.006 2,交通量每增加 1%,客车交通事故频次增大 0.62%。模型标定结果显示路段长度系数服从均值为 0.895 9、标准差为

表 6 解释变量的发生率比值

变量	E_{IRR}	变量	E_{IRR}
路段长度	1.009 0	温度	0.942 8
交通量	1.006 2	相对湿度	0.911 8
1 类车占比	1.014 7	能见度	0.934 5
5 类车占比	1.000 2	降水量	0.999 1
路基宽度	0.968 7	风速	0.977 7
曲率	0.999 8	节假日占比	1.083 0
坡度	1.004 1		

0.301 7 的正态分布,交通量系数服从均值为 0.617 6、标准差为 0.301 7 的正态分布,二者在不同路段上对客车交通事故频次的影响具有异质性。路段长度和交通量系数的概率密度分布如图 4a)、b) 所示。由图 4a)、b) 可知:99.85% 的路段中客车交通事故频次随路段长度的增大而增大,0.15% 的路段中客车交通事故频次随路段长度增大而减小;97.97% 的路段中客车交通事故频次随交通量增大而增大,2.03% 的路段中客车交通事故频次随交通量的增大而减小。

温度、相对湿度、能见度的贝叶斯置信区间未包含 0 且后验均值均为负,三者与客车交通事故频次均显著负相关。温度的发生率比值为 0.942 8,温度每增加 1℃,客车交通事故频次减小 5.72%;温度系数服从均值为 -0.058 9、标准差为 0.087 4 的正态分布,温度系数的概率密度分布如图 4c) 所示。由图 4c) 可知:在 74.98% 的路段中客车交通事故频次随温度增加而减小,可能因为温度升高改善了路面状况,提高了车辆的稳定性和司机的可见性;在 25.02% 的路段中交通事故频次随温度增加而增大,可能因为高温易导致驾驶员疲劳、车辆胎压增大易导致车辆爆胎,进而增大交通事故的风险。

相对湿度的发生率比值为 0.911 8,相对湿度每增加 1%,客车交通事故频次减小 8.82%,相对湿度系数服从均值为 -0.092 3、标准差为 0.043 3 的正态分布,概率密度分布如图 4d) 所示。由图 4d) 可知:在 98.35% 的路段中客车交通事故频次随相对湿度增加而减小,可能是大部分驾驶员由于道路湿滑而减缓速度,小心驾驶,交通事故发生风险下降;在 1.65% 的路段中客车交通事故频次随相对湿度增加而增大,因为高湿度会降低路面的摩擦系数,增大车辆打滑的风险,并会影响驾驶员的能见度,导致交通事故发生率增大。

能见度的发生率比值为 0.934 5,能见度每增加 1 km,客车交通事故频次降低 6.55%。能见度系数服从均值为 -0.067 8、标准差为 0.088 7 的正态分布,概率密度分布如图 4e) 所示。由图 4e) 可知:在 77.77% 的路段中交通事故频次随能见度增加而减小,能见度高时,驾驶员能清晰地看到前方的道路和其他车辆,驾驶员的反应速度较快,减少碰撞和失控的可能性;在 22.23% 的路段中客车交通事故频次随能见度增加而增大,可能是因为良好的能见度导致驾驶员放松警惕,影响其决策和驾驶行为,增大了交通事故风险。

节假日占比的贝叶斯置信区间未包含 0 且后验均值均为正,与客车交通事故频次显著正相关。节假日占比的发生率比值为 1.083 0,节假日占比每增加 1%,客车交通事故频次增大 8.30%。节假日占比系数服从均值为 0.079 8、标准差为 0.010 4 的正态分布,概率密度分布如图 4f) 所示。由图 4f) 可知:77.79% 的路段中客车交通事故频次随节假日占比的增大而增大,节假日期间高速公路交通量普遍增大,交通拥堵和驾驶压力增大,交通事故风险增大;22.21% 的路段中客车交通事故频次随节假日占比的增大而减小,可能是节假日期间执法和安全宣传教育力度增大,道路维护和交通管制措施加强,客车交通事故率明显减少。

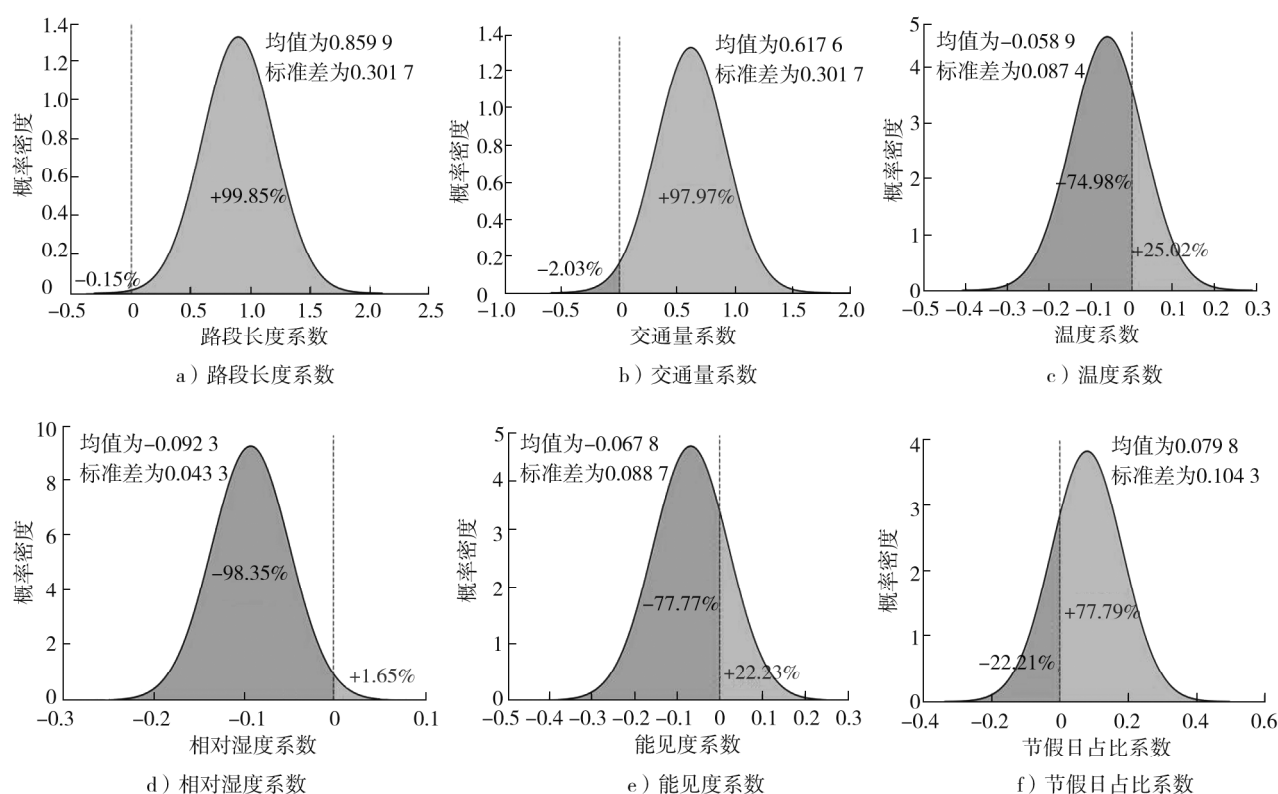


图4 RP-PLN 模型参数概率密度分布

为降低高速公路客车交通事故频次,可结合智能技术、工程改良、管理优化、公众教育等措施构建多维度防控体系:采用智能交通系统监控交通,结合大数据分析车流量,动态调整限速,联合导航软件及时推送前方事故或施工信息,帮助驾驶员规避高风险路段;在高温地区使用耐高温路面材料,加强车辆维护,减少爆胎风险;在弯道、坡道等事故高发段铺设排水沥青或陶瓷颗粒防滑层,降低湿滑事故率;增设太阳能主动发光轮廓标和反光道钉,强化夜间及低能见度条件下的车道引导;春节、国庆等节假日期间加强交通执法,在服务区增设临时检查点,严查超员等违规行驶;通过交通广播、社交媒体等加强安全行车宣传,提高公众主动安全意识。

4 结论

为分析高速公路客车交通事故频次的影响因素,以交通量、路段长度、温度、相对湿度、能见度、节假日占比等影响因素为解释变量,以高速公路客车事故频次为因变量,构建 PLN 模型、RP-PLN 模型及 RE-PLN 模型,RP-PLN 模型的偏差信息准则、平均绝对偏差和均方预测误差均小于 PLN 模型和 RE-PLN 模型,表明 RP-PLN 模型拟合效果更优,且能有效处理高速客车事故的异质性。由 RP-PLN 模型的贝叶斯参数估计结果可知:路段长度、交通量和节假日占比与高速公路客车交通事故频次正相关,当其他因素不变,路段长度、交通量每增加 1%时,高速公路客车交通事故频次分别增大 0.90%、0.62%,节假日占比每增加 1 个单位时,高速公路客车交通事故频次增大 8.30%;温度、相对湿度和能见度与高速公路客车交通事故频次负相关,温度、相对湿度和能见度每增加 1 个单位时,高速公路客车交通事故频次分别减少 5.72%、8.82%和 6.55%。建议结合智能交通监控技术、道路工程改良、交通管理优化、公众教育等多方面构建综合防控体系,降低高速公路交通事故频次,提高高速公路客车行车安全。

本文仅分析不同观测样本间的异质性,未考虑时空效应。未来研究可引入时空效应模型,分析交通事故在不同地理位置和时间上的分布规律。

参考文献:

- [1] 国家统计局. 中国统计年鉴 2024[M]. 北京: 中国统计出版社, 2024.
- [2] 中国法院网. 最高法发布交通肇事罪司法大数据专题报告[EB/OL]. (2020-08-06)[2024-05-08]. <https://www.chinacourt.org/article/detail/2020/08/id/5394281.shtml>.
- [3] KAPLAN S E, WILLIAMS D D. Do going concern audit reports protect auditors from litigation? A simultaneous equations approach[J]. *The Accounting Review*, 2013, 88(1): 199-232.
- [4] FARID A, ABDEL-ATY M, LEE J. Transferring and calibrating safety performance functions among multiple States[J]. *Accident Analysis & Prevention*, 2018, 117: 276-287.
- [5] 王雪松, 宋洋, 黄合来, 等. 基于分层负二项模型的城郊公路安全影响因素研究[J]. *中国公路学报*, 2014, 27(1): 100-106.
- [6] WANG K, IVAN J N, RAVISHANKER N, et al. Multivariate Poisson lognormal modeling of crashes by type and severity on rural two lane highways[J]. *Accident Analysis & Prevention*, 2017, 99: 6-19.
- [7] LORD D, MIRANDA-MORENO L F. Effects of low sample mean values and small sample size on the estimation of the fixed dispersion parameter of Poisson-gamma models for modeling motor vehicle crashes: a Bayesian perspective[J]. *Safety Science*, 2008, 46(5): 751-770.
- [8] HOU Q Z, MENG X H, HUO X Y, et al. Effects of freeway climbing lane on crash frequency: application of propensity scores and potential outcomes[J]. *Physica A: Statistical Mechanics and Its Applications*, 2019, 517: 246-256.
- [9] 曾强, 王雪松, 张璇, 等. 基于时空交互模型的高速公路季节事故频次影响因素分析[J]. *中国公路学报*, 2020, 33(11): 255-263.
- [10] 温惠英, 孙佳人, 曾强, 等. 高速公路交通构成的安全效应计量评价[J]. *华南理工大学学报(自然科学版)*, 2018, 46(6): 1-7.
- [11] 孙佳人. 基于贝叶斯时空建模的高速公路交通安全计量评价研究[D]. 广州: 华南理工大学, 2020.
- [12] 郭延永, 刘攀, 吴瑶, 等. 考虑异质性的贝叶斯交通冲突模型[J]. *中国公路学报*, 2018, 31(4): 296-303.
- [13] 漆巍巍, 覃铄磊, 朱淑芳, 等. 考虑空间溢出效应的高速公路交通事故随机参数频次模型[J]. *交通运输系统工程与信息*, 2024, 24(2): 263-271.
- [14] 蔡明懋. 考虑数据零膨胀特征与异质性的高速公路隧道路段交通安全计量研究[D]. 广州: 华南理工大学, 2022.
- [15] 唐峰. 基于事故频数建模的高速公路交通安全计量评价研究[D]. 广州: 华南理工大学, 2022.
- [16] HUO X Y, LENG J Q, HOU Q Z, et al. Assessing the explanatory and predictive performance of a random parameters count model with heterogeneity in means and variances[J]. *Accident Analysis & Prevention*, 2020, 147: 105759.
- [17] 侯芹忠. 考虑异质性与内生性的高速公路交通事故随机参数模型[D]. 哈尔滨: 哈尔滨工业大学, 2019.
- [18] ZHAO S S, WANG K, LIU C H, et al. Investigating the effects of monthly weather variations on Connecticut freeway crashes from 2011 to 2015[J]. *Journal of Safety Research*, 2019, 71: 153-162.
- [19] HOU Q Z, MENG X H, LENG J Q, et al. Application of a random effects negative binomial model to examine crash frequency for freeways in China[J]. *Physica A: Statistical Mechanics and its Applications*, 2018, 509: 937-944.
- [20] HOU Q Z, TARKO A P, MENG X H. Investigating factors of crash frequency with random effects and random parameters models: New insights from Chinese freeway study[J]. *Accident Analysis & Prevention*, 2018, 120: 1-12.
- [21] WEN H Y, ZHANG X, ZENG Q, et al. Bayesian spatial-temporal model for the main and interaction effects of roadway and weather characteristics on freeway crash incidence[J]. *Accident Analysis & Prevention*. 2019, 132: 105249.
- [22] 汤海学. 江苏省节假日高速公路交通拥堵成因及大流量分级管控策略探究[J]. *中国公路*, 2022(16): 104-107.
- [23] 王思彤, 贵莉莉, 李建东. 法定节假日高速公路拥堵成因及对策研究: 以江苏为例[J]. *统计科学与实践*, 2023(3): 19-22.
- [24] 周胜. 关于江苏省高速公路节假日交通大流量应对处置工作的调查研究[J]. *道路交通管理*, 2024(1): 52-55.
- [25] 全国交通工程设施(公路)标准化技术委员会. 收费公路车辆通行费车型分类: JT/T 489—2019[S]. 北京: 人民交通出版社, 2019.
- [26] 胡郁葱, 韦湖, 曾强. 基于空间广义有序 Probit 模型的高速公路事故严重程度分析[J]. *华南理工大学学报(自然科学版)*, 2023, 51(1): 114-122.

(下转第 42 页)

patterns at critical positions of the supports. The results show that the deformation data measured by the gyroscope closely match the simulated data, and the angular change threshold for construction process monitoring is feasible. Insufficient foundation bearing capacity leads to three different settlement conditions; first, when the single-point foundation bearing capacity is insufficient, forced displacements are applied to the four corner support rods at the bottom plane of the middle part of the support structure, resulting in significant angular changes at four measurement points, but none exceeds the warning threshold; settlement occurred at the position where the largest reaction force from the support is applied at the bottom of the support; second, when the foundation bearing capacity is insufficient in certain areas, forced displacements are applied to the middle, left, and right parts of the support structure, significantly affecting the key monitoring sections, with noticeable changes in local structural displacements; the angular changes remain within the permissible thresholds, and the reduction rates of displacement and angle during the pouring condition show minimal variation; third, when the foundation bearing capacity at the bottom of the members where each measurement point is located is insufficient, simultaneous node displacements are applied to the bottom nodes of the monitoring sections, with major settlement occurring in the regions around the bottom of the measurement points; the angular threshold increases with the increase in load, validating the effectiveness of the settlement identification and early warning method based on angular thresholds.

Keywords: foundation bearing capacity; support; deformation; finite element modeling analysis; angle threshold

(责任编辑:王惠)

(上接第 21 页)

[27] 张璇, 温惠英. 基于贝叶斯空间有序 logit 模型的高速公路货车事故严重程度分析[J]. 甘肃科学学报, 2022, 34(6): 78-84.

Analysis of factors influencing highway bus accidents frequency considering heterogeneity

CAO Zhen, LIU Hongyu^{*}, TIAN Zhun

Collage of Urban Development and Modern Transportation, Xi'an University of Architecture and Technology, Xi'an 710055, China

Abstract: To enhance highway bus traffic safety and reduce accident rates, data from 2,320 samples are collected, including bus accident frequency, traffic volume, road characteristics, weather conditions, and holiday proportions, from three highways in eastern China. Accounting for accident heterogeneity, Poisson-lognormal (PLN), random parameters Poisson-lognormal (RP-PLN), and random effects Poisson-lognormal (RE-PLN) models are established. The Bayesian method and Markov chain Monte Carlo simulation estimate the posterior distribution of model parameters. The modeling results quantitatively assess the impact of various factors on highway bus accident frequency. Findings indicate that the RP-PLN model exhibits superior goodness-of-fit compared to the PLN and RE-PLN models. Road segment length, traffic volume, and holiday proportion show positive correlations with accident frequency, while temperature, relative humidity, and visibility demonstrate negative correlations. Holding other factors constant, a 1% increase in road segment length and traffic volume increases accident frequency by 0.90% and 0.62%, respectively. A one-unit increase in holiday proportion increases accident frequency by 8.30%. Conversely, a one-unit increase in temperature, relative humidity, and visibility reduces accident frequency by 5.72%, 8.82%, and 6.55%, respectively.

Keywords: highway; bus accident; heterogeneity; RP-PLN model

(责任编辑:赵玉真)