

تحلیل و پیش‌بینی فراوانی تصادفات ساعات اوج در تقاطع‌های شهری با استفاده از مدل دوجمله‌ای منفی با اثرات تصادفی (RENB)

مقاله علمی-پژوهشی

*شاهین شعبانی (نویسنده مسئول)، گروه مهندسی عمران، دانشگاه پیام نور، تهران، ایران

امیربهادر عظیمی، دانش پژوه دکتری، گروه مهندسی عمران، دانشگاه پیام نور، تهران، ایران

*پست الکترونیکی نویسنده مسئول: shabani@pun.ac.ir

دریافت: ۱۴۰۵/۰۲/۲۱ - پذیرش: ۱۴۰۵/۰۵/۱۰

صفحه ۶۶-۵۵

چکیده

مدلسازی فراوانی تصادفات در تقاطع‌های شهری به دلیل ماهیت گسسته، تصادفی و بیش‌پراکنده داده‌های تصادف، همواره یکی از چالش‌های اصلی مطالعات ایمنی راه بوده است. پژوهش حاضر با هدف تحلیل و پیش‌بینی فراوانی تصادفات ساعات اوج روزانه در تقاطع‌های شهری، از مدل دوجمله‌ای منفی با اثرات تصادفی (RENB) استفاده نموده است. داده‌های مورد استفاده شامل اطلاعات ۳۴ تقاطع شهری در شهر بروجرد طی دوره پنج ساله ۱۴۰۰ تا ۱۴۰۴ بوده که پس از تجمیع ساعتی، ۲۰۴ مشاهده و مجموعاً ۶۹۸ فقره تصادف را شامل می‌شود. متغیرهای توضیحی شامل حجم ترافیک ساعتی راه اصلی (ADTmaj)، حجم ترافیک راه فرعی (ADTmin)، تعداد شاخه‌های تقاطع، تعداد چراغ راهنما و متغیرهای مجازی ساعات اوج مختلف بودند. نتایج تحلیل آماری وجود بیش‌پراکنندگی معنادار در داده‌ها را تأیید کرد و به دلیل پایین بودن سهم مشاهدات صفر، استفاده از مدل دوجمله‌ای منفی نسبت به مدل‌های پرفرمر مناسب‌تر تشخیص داده شد. همچنین مقایسه مدل دوجمله‌ای منفی کلاسیک و مدل RENB بر اساس شاخص‌های BIC و AIC و آزمون نسبت درست‌نمایی نشان داد که لحاظ کردن اثرات تصادفی موجب بهبود معنادار عملکرد مدل شده است. نتایج نشان داد که حجم ترافیک راه فرعی تأثیر بیشتری نسبت به راه اصلی بر افزایش تصادفات دارد؛ به گونه‌ای که افزایش ۱۰۰۰ خودرو در ساعت در راه فرعی، نرخ مورد انتظار تصادفات را حدود ۱۴۶ درصد افزایش می‌دهد. همچنین افزایش تعداد شاخه‌های تقاطع از سه به چهار شاخه، موجب افزایش ۷۶ درصدی فراوانی تصادفات می‌شود. در میان متغیرهای زمانی، تنها ساعت ۷-۶ صبح کاهش معنادار تصادفات نسبت به ساعت مرجع ۲۰-۱۹ را نشان داد. مقدار R^2 شرطی برابر ۰/۷۱ و همبستگی پیرسون ۰/۶۷ بیانگر برازش مناسب مدل پیشنهادی است. نتایج این پژوهش می‌تواند در اولویت‌بندی اقدامات ایمنی، بهینه‌سازی زمان‌بندی چراغ‌ها و مدیریت تقاطع‌های شهری مورد استفاده قرار گیرد.

واژه‌های کلیدی: مدلسازی تصادفات، ساعات اوج، مدل دوجمله‌ای منفی با اثرات تصادفی، تقاطع‌های شهری

۱- مقدمه

Abdel- (1990; Miaou, 1994) و مدل‌های منفی دوجمله‌ای (Aty & Radwan, 2000; Champahom et al., 2022; Xu & Huang, 2015; Young & Park, 2013) برای تحلیل تصادفات استفاده کرده‌اند و نتایج، در نظر گرفتن عوامل

محققان از ابزارهای آماری مختلفی از جمله رگرسیون خطی چندگانه (Joshua & Garber, 1990; Jovanis & Chang, 1986)، مدل‌های پواسون (Joshua & Garber, 1990)

تصادفی برای در نظر گرفتن ناهمگونی ناشی از عوامل مؤثر استفاده شده‌اند و نتایج قابل اعتمادتری از مدل‌های سنتی نشان داده‌اند (Almutairi, 2024; Azmeri Khan., 2023; Moomen et al., 2020; Gong et al., 2020; Anastasopoulos & Mannering, 2009; Chen and Tarko, 2014).

باتوجه به پیشینه بررسی شده در خصوص مدل‌سازی تصادفات، مشخص می‌شود که با وجود مزایا و پرکاربرد بودن مدل‌های دوجمله‌ای منفی سنتی در مدل‌سازی تصادفات تقاطع‌ها (Tavakoli Kashani et al., 2022)، این مدل‌ها قادر به لحاظ کردن ناهمگونی مشاهده‌نشده بین تقاطع‌ها نیستند و اغلب اثرات خوشه‌ای و مکانی-زمانی را نادیده می‌گیرند. از سوی دیگر، اکثر مطالعات پیشین بر داده‌های تصادفات سالیانه متمرکز بوده‌اند و تحلیل ساعتی تصادفات (که برای مدیریت ساعات اوج و بهینه‌سازی زمان‌بندی چراغ راهنما حیاتی است) کمتر مورد توجه قرار گرفته است که آن هم مربوط به تقاطع‌ها نیست. برای مثال یوان و همکاران (Yuan et al., 2021) مدل دوجمله‌ای منفی را برای بزرگراه‌ها با استفاده از داده‌های آشکارساز ترافیک میکروسکوپی چند-حالتی توسعه داد و آن‌ها را در حجم متوسط ساعتی، حجم روزهای هفته و آخر هفته، حجم ساعات اوج و غیراوج و AADT تحلیل کرد. رویکرد مشابهی نیز توسط فو و همکاران (Fu et al., 2022) برای تحلیل تصادفات در کوتاه‌مدت به منظور به‌دست‌آوردن نرخ‌های ساعتی تصادف مورد بحث قرار گرفت. آن‌ها داده‌های تصادف بزرگراهی در یک ایالت آمریکای شمالی را تحلیل و داده‌های کوتاه‌مدت را به صورت ساعتی تجمیع کردند تا فراوانی تصادفات را پیش‌بینی کنند.

باتوجه به مطالب فوق، نوآوری اصلی پژوهش حاضر در توسعه مدل دوجمله‌ای منفی با اثرات تصادفی^۲ (RENB) برای تحلیل تصادفات ساعتی تقاطع‌های شهری با استفاده از داده‌های ۳۴ تقاطع شهر بروجرد، برای پاسخ به دو سوال اصلی است: ۱) کدام یک از متغیرهای حجم ترافیک راه اصلی و فرعی تاثیر بیشتری بر تصادفات ساعات اوج دارد؟ و ۲) آیا ساعات مختلف اوج روزانه تفاوت معناداری از نظر خطر تصادف با یکدیگر دارند؟.

۲- داده‌ها و انتخاب متغیرها

این پژوهش از دو نوع داده استفاده می‌کند.

-داده‌های سابقه تصادف تقاطع‌های شهر بروجرد از سال ۱۴۰۰ تا ۱۴۰۴ (۵ سال).

مهندسی، محیطی و انسانی به عنوان سه دسته کلی را رضایت‌بخش یافته‌اند. عوامل مهندسی معمولاً شامل ویژگی‌های هندسی راه (تعداد خطوط، عرض خط، عرض میانه و شانه)، حجم ترافیک (روزانه و سالانه)، شرایط ترافیک، زمان‌بندی چراغ‌های راهنمایی و شرایط سطح راه؛ عوامل محیطی شامل شرایط جوی، میزان بارش، دید و غیره؛ و عوامل انسانی شامل جنسیت، سن، خستگی و غیره هستند (Retallack & Ostendorf, 2020).

در حالی که عوامل توضیحی بر نتایج رگرسیون تأثیر می‌گذارند، مدل‌های پیش‌بینی با ویژگی‌های مختلف ممکن است در تحلیل تصادفات که رویدادهای تصادفی با داده‌های پراکنده هستند، متفاوت عمل کنند. به عنوان مثال، مدل رگرسیون خطی چندگانه می‌تواند به‌طور غیرعادی تعداد تصادفات را کمتر از صفر پیش‌بینی کند و چنین محدودیت‌هایی در مدیریت تصادفی بودن، گسسته بودن و غیرمنفی بودن توسط چندین مطالعه برجسته شده است (Alqatawna et al., 2021; Jovanis & Chang, 1986; Zegeer et al., 1991). علاوه بر این، مدل‌های پواسون اگرچه به‌طور گسترده استفاده می‌شوند، اما نمی‌توانند عوامل بیش‌پراکندگی را مدیریت کنند، جایی که واریانس بزرگتر از میانگین است (Abdel-Aty & Radwan, 2000; Shankar et al., 1995). با این حال چندین مطالعه از مدل دوجمله‌ای منفی استفاده کرده‌اند که نشان می‌دهد این مدل می‌تواند بیش‌پراکندگی، غیرمنفی بودن و تصادفی بودن موجود در مجموعه داده‌های تصادف را در نظر بگیرد (Abdel-Aty & Radwan, 2000; Champahom et al., 2022; Chin & Quddus, 2003; Islam et al., 2022; Mathew et al., 2022; Xu & Huang, 2015; Yuan et al., 2021).

پیش از سال ۲۰۰۰، مدل‌های پواسون و رگرسیون دوجمله‌ای منفی^۱ (NB) عمدتاً برای برقراری روابط آماری بین شمارش تصادفات خودرویی و عوامل مؤثر مورد استفاده قرار می‌گرفتند (Gong et al., 2020). با پیشرفت مستمر در ایمنی بزرگراه‌ها، مدل‌های آماری پیچیده‌تری معرفی شدند که هدف آن‌ها دستیابی به نتایج بهتر بود.

این مدل‌های جدید، محدودیت‌های ذاتی چارچوب‌های مدل قبلی را تسهیل کردند و عموماً برازش بهتری نشان دادند. چین و همکاران اثرات تصادفی را به مدل‌های NB برای در نظر گرفتن اثرات زمانی و مکانی پنهان بین دوره‌های زمانی و مکان‌های فضایی مختلف در مطالعه فراوانی تصادفات خود معرفی کردند (Chin & Quddus, 2003). مدل‌های شمارشی با پارامترهای

دوجمله‌ای منفی را به جای مدل پواسون ساده توجیه مینماید. درصد صفرها در numcr حدود ۴/۵ درصد است که کمتر از آستانه ۲۵ درصد (آستانه رایج در ادبیات ایمنی برای استفاده از مدل‌های پرفر) بوده، بنابراین نیازی به مدل پرفر نیست.

مطالعات پیشین (Champahom et al., 2022; Chin & Quddus, 2003; Gu et al., 2020) بر اهمیت تعیین نوع رابطه (ثابت یا تصادفی) بین متغیرهای توضیحی قبل از کالبره کردن مدل تأکید دارند. با توجه به اینکه حجم ترافیک در تقاطع‌های مختلف ممکن است اثرات خوش‌های تصادفی داشته باشد، از رویکرد اثرات تصادفی برای ضرایب $X1$ و $X2$ و همچنین ثابت مدل استفاده شده است.

پیش از برازش مدل، همبستگی بین متغیرهای توضیحی با استفاده از ضریب همبستگی پیرسون بررسی شد. مقدار این ضریب بین ۱- تا ۱ قرار دارد و آستانه ۰/۷ برای تشخیص همخطی به کار رفت (Basso et al., 2018; Q. Cai et al., 2020; Guo et al., 2021).

نتایج نشان داد که تمام همبستگی‌ها کمتر از ۰/۷ هستند، بنابراین هیچ تغییری حذف نشد. بر اساس مقادیر جدول (۲)، بیشتر تقاطع‌ها دارای حجم ترافیک راه اصلی (ADTmaj) بین ۷۲۰ تا ۱۷۰۰ خودرو در ساعت و حجم ترافیک راه فرعی (ADTmin) بین ۲۰۰ تا ۶۱۰ خودرو در ساعت هستند. همچنین اکثر تقاطع‌ها (بیش از ۹۰ درصد) دارای ۴ شاخه هستند و تعداد چراغ راهنما در آنها بین ۳ تا ۵ متغیر است.

جدول ۱. پارامترها، نمادها، متغیرها و توصیف آنها

نوع پارامتر	شرح	متغیر	نماد
—	تعداد واقعی تصادفات در هر ساعت اوج (متغیر وابسته)	numcr	λt
تصادفی	میانگین ساعتی ترافیک روزانه در راه اصلی	ADTmaj	$X1$
تصادفی	میانگین ساعتی ترافیک روزانه در راه فرعی	ADTmin	$X2$
ثابت	تعداد شاخه‌های تقاطع (نماینده ویژگی هندسی)	nleg	$X3$
ثابت	تعداد چراغ راهنما (نماینده وجود گردش به چپ مجاز)	nsign	$X4$
ثابت	متغیر مجازی: ساعت ۶-۷ صبح	mp1	$Y1$
ثابت	متغیر مجازی: ساعت ۷-۸ صبح	mp2	$Y2$
ثابت	متغیر مجازی: ساعت ۱۱-۱۲ ظهر	ap1	$Y3$
ثابت	متغیر مجازی: ساعت ۱۲-۱۳	ap2	$Y4$
ثابت	متغیر مجازی: ساعت ۱۸-۱۹ (۶-۷ عصر)	ep1	$Y5$
ثابت	متغیر مجازی: ساعت ۱۹-۲۰ (۷-۸ عصر، مرجع)	ep2	$Y6$

— داده‌های حجم ترافیک در ساعات اوج برای همان تقاطع‌ها در سال‌های مربوطه.

جدول (۱) پارامترها، متغیرها و اختصارات کوتاه‌شده‌ای را که در این پژوهش برای تحلیل تصادفات در تقاطع‌ها به کار رفته‌اند، خلاصه میکند. با فرض اینکه تقاطع Z دارای متغیرهای $X1$ (حجم ترافیک راه اصلی)، $X2$ (حجم ترافیک راه فرعی)، $X3$ (تعداد شاخه‌ها)، $X4$ (تعداد چراغ راهنما) و متغیرهای مجازی ساعات اوج $Y1$ تا $Y6$ باشد، تعداد تصادفات واقعی λt در هر ساعت اوج t از طریق مدل دوجمله‌ای منفی با پارامترهای تصادفی برآورد می‌شود. هدف نهایی، پیشبینی تعداد تصادفات آینده بر اساس مقادیر این متغیرهاست. بدین منظور، ابتدا داده‌های حجم ترافیک با بازه ۱۵ دقیقه‌ای به صورت ساعتی جمع‌بندی می‌شوند. سپس، تعداد تصادفات مربوط به همان ساعت در هر تقاطع استخراج می‌شود. این فرآیند منجر به ایجاد یک مجموعه داده شامل ۳۴ تقاطع با ۶ مشاهده ساعتی برای هر کدام (مجموعاً ۲۰۴ مشاهده ساعتی) و ۶۹۸ تصادف در کل تقاطع‌ها شد.

از داده‌های ۲۷ تقاطع (۸۰ درصد) برای آموزش مدل‌ها و ۷ تقاطع باقیمانده (۲۰ درصد) برای اعتبارسنجی و آزمون مدل استفاده خواهد شد. خلاصه این داده‌ها در جدول (۲) نشان داده شده است.

همانطور که در جدول (۲) مشاهده می‌شود، متغیر numcr (تعداد تصادفات) دارای واریانس (۱۴/۸) بزرگتر از میانگین (۳/۴۲) است که وجود بیش‌پراکندگی را تأیید می‌کند. این ویژگی استفاده از مدل

جدول ۲. خلاصه آماره‌های متغیرهای جمع‌آوری شده

متغیر	تعداد	میانگین	انحراف معیار	حداقل	چارک اول	میان	چارک سوم	حداکثر
numcr	۲۰۴	۳/۴۲	۳/۸۵	۰	۱	۲	۵	۲۱
ADTmaj	۲۰۴	۱۲۵۰	۷۱۰	۷۰	۷۲۰	۱۲۰۰	۱۷۰۰	۴۴۰۰
ADTmin	۲۰۴	۴۷۰	۳۶۵	۰	۲۰۰	۴۰۰	۶۱۰	۲۱۰۰
nleg	۲۰۴	۳/۸۸	۰/۳۳	۳	۴	۴	۴	۴
nsign	۲۰۴	۳/۵۵	۰/۵۰	۳	۳	۴	۴	۵

۳- روش تحقیق

راهکاری مؤثر شناخته شده است (Geedipally et al., 2012). با توجه به اینکه داده‌های این پژوهش حدنصاب صفرهای زیاد را ندارد (درصد صفرها کمتر از آستانه ۲۵ درصد و حدود ۴/۵ درصد است)، در این پژوهش از مدل «دوجمله‌ای منفی پارامترهای تصادفی» برای رسیدگی به ناهمگونی مشاهد‌نشده استفاده شده است. با فرض اینکه y_{it} نشان‌دهنده تعداد تصادفات در تقاطع i در دوره زمانی t باشد و y_{it} از توزیع دوجمله‌ای منفی پیروی کند، احتمال اینکه تقاطع i در دوره زمانی t دارای تصادف باشد از رابطه زیر به دست می‌آید (Washington et al., 2020).

$$P(y_{it}| \lambda_{it}, \phi) = \frac{\Gamma(\phi + y_{it})}{\Gamma(\phi) \Gamma(y_{it} + 1)} \left(\frac{\phi}{\phi + \lambda_{it}} \right)^\phi \left(\frac{\lambda_{it}}{\phi + \lambda_{it}} \right)^{y_{it}} \quad (1)$$

که λ_{it} نشان‌دهنده میانگین، $\Gamma(\cdot)$ تابع گاما و ϕ ضریب بیش‌پراکندگی است. سپس حال می‌توان λ_{it} را که بیانگر میانگین مورد انتظار تصادفات تقاطع است، به صورت تابعی از متغیرهای توضیح‌دهنده و با استفاده از یک رابطه لگاریتمی-خطی (Log-linear) بیان کرد (Mitra & Washington, 2007):

$$\lambda_{it} = \alpha_{0it} F_{it}^{\alpha_{1it}} \exp(b_{it} X'_{it}) \quad (2)$$

در این مدل، F_{it} نشان‌دهنده حجم ترافیک (به عنوان معیاری از مواجهه) در تقاطع i در دوره زمانی t است. همچنین X'_{it} برداری از متغیرهای توضیح‌دهنده و $\alpha_{0it}, \alpha_{1it}$ و b_{it} پارامترهای رگرسیون را تشکیل می‌دهند. در این ساختار، پارامترها اجازه دارند تا در بین مشاهدات مختلف متفاوت باشند تا ناهمگونی مشاهد‌نشده مورد توجه قرار گیرد.

مدل‌های رگرسیون شمارشی برای بررسی رابطه بین خطر تصادف و عوامل مؤثر بر آن در پژوهش‌های ایمنی راه کاربرد فراوانی دارند (Lord & Mannering, 2010). این مدل‌ها بر این فرض استوارند که تصادفات خروجی یک فرآیند پواسون هستند و تعداد آن‌ها از توزیع پواسون پیروی می‌کند (Lord et al., 2005). با این حال، شواهد پژوهشی نشان می‌دهد که داده‌های مربوط به تعداد تصادفات اغلب دچار «بیش‌پراکندگی» هستند، به این معنا که واریانس داده‌ها از میانگین آن‌ها بزرگتر است. در چنین شرایطی، مدل دوجمله‌ای منفی عملکرد بهتری نسبت به مدل پواسون دارد (Poch & Mannering, 1996). با وجود این مزیت، مدل‌های دوجمله‌ای منفی ممکن است کارایی کافی نداشته باشند.

این موضوع دو دلیل اصلی دارد:

ناهمگونی مشاهد‌نشده: تصادفات نتیجه تعامل پیچیده عوامل متعددی هستند. حذف هر یک از این عوامل یا تعاملات آن‌ها از تحلیل، می‌تواند بر اثر متغیرهای باقی‌مانده تأثیر بگذارد. این پدیده که «ناهمگونی مشاهد‌نشده» نامیده می‌شود (Mannering et al., 2016)، اگر در نظر گرفته نشود، منجر به تخمین‌های ضعیف و ناکارآمد از پارامترهای مدل خواهد شد. یکی از راهکارهای مقابله با این مشکل، استفاده از مدل‌سازی با پارامترهای تصادفی است (Mannering et al., 2016). وجود صفرهای زیاد: فرضیه‌های توزیع در مدل‌های دوجمله‌ای منفی زمانی نقض می‌شوند که داده‌های تصادف حاوی تعداد زیادی صفر باشند. در چنین حالتی، پارامترهای مدل به درستی تخمین زده نمی‌شوند. برای رفع این مشکل، استفاده از مدل‌های ترکیبی مانند «دوجمله‌ای منفی لیندلی»

(۳)

$$b_{it} = b + \omega_{it}$$

۴- نتایج و یافته‌ها

۴-۱- تخمین مدل

به منظور بررسی ضرورت استفاده از مدل دوجمله‌ای منفی با اثرات تصادفی (RENB)، عملکرد آن با مدل دوجمله‌ای منفی کلاسیک (NB) مقایسه شد (جدول ۳). مقایسه براساس مقادیر AIC، BIC و آزمون نسبت درستنمایی انجام گرفت. مقادیر کمتر AIC و BIC برای مدل RENB نشانه برازش بهتر آن است. همچنین مقدار معنادار آزمون درستنمایی نشان می‌دهد که لحاظ کردن اثرات تصادفی موجب بهبود معنادار عملکرد مدل نسبت به مدل NB شده است. بنابراین استفاده از ساختار اثرات تصادفی برای داده‌های این پژوهش مناسب تشخیص داده شد.

۴-۲- تخمین مدل

نتایج برآورد مدل دوجمله‌ای منفی با اثرات تصادفی برای داده‌های ۳۴ تقاطع شهر بروجرد در جدول (۴) ارائه شده است. پارامتر پراکندگی (θ) معادل ۱/۲ برآورد شد که تأییدکننده وجود بیش‌پراکندگی در داده‌ها و مناسب بودن مدل دوجمله‌ای منفی است. هر چه θ کوچکتر باشد، میزان بیش‌پراکندگی بیشتر است؛ مقدار ۱/۲ نشان‌دهنده پراکندگی قابل توجه است.

که در آن b نشان‌دهنده میانگین ثابت پارامترهای تصادفی است و ω_{it} جمله‌ای تصادفی با توزیع مشخص است (در این پژوهش فرض شد که پارامترهای تصادفی از توزیع نرمال پیروی می‌کنند. انتخاب توزیع نرمال به دلیل کاربرد گسترده آن در مالعات ایمنی و توانایی مناسب در مدل‌سازی ناهمگونی مشاهده‌نشده صورت گرفت). نکته حائز اهمیت این است که این پارامترها امکان تغییر بین تقاطع‌های مختلف را دارند، اما برای هر تقاطع در طول دوره‌های زمانی ثابت باقی می‌مانند. به کارگیری این نوع پارامترهای تصادفی گروهی (Oviedo-Trespalcacios et al., 2020) امکان در نظر گرفتن ماهیت پنبلی داده‌ها (شامل چندین مشاهده از یک تقاطع در سال‌های متفاوت) را فراهم می‌سازد. بر این اساس، تابع توزیع احتمال برای y_{it} به صورت زیر بیان می‌شود.

(۴)

$$P(y_{it}|\lambda_{it}, \phi, \sigma) = \int \frac{\Gamma(\phi + y_{it})}{\Gamma\phi + y_{it}!} \left(\frac{\phi}{\phi + \lambda_{it}}\right)^\phi \left(\frac{\lambda_{it}}{\phi + \lambda_{it}}\right)^{y_{it}} g(\omega_{it}) d\omega_{it}$$

که در آن در اینجا $g(\omega_{it})$ نشان‌دهنده تابع چگالی احتمال جمله تصادفی ω_{it} است. این مدل در ادبیات پژوهشی به عنوان «مدل دوجمله‌ای منفی با اثرات تصادفی (RENB) شناخته می‌شود (Anastasopoulos & Mannering, 2009).

در این پژوهش کلیه مراحل مدل‌سازی از طریق بسته نرم‌افزاری R انجام شد.

جدول ۳. مقایسه عملکرد مدل‌های دوجمله‌ای منفی (NB) و دوجمله‌ای منفی با اثرات تصادفی (RENB)

شاخص عملکرد	مدل NB	مدل RENB
لگاریتم درستنمایی	-۴۲۱/۸	-۳۹۸/۶
AIC	۸۶۵/۶	۸۲۱/۲
BIC	۹۰۱/۴	۸۶۷/۹
پارامتر پراکندگی (θ)	۱/۳۷	۱/۲۰
R^2 حاشیه‌ای	۰/۴۳	۰/۴۹
MAD	۲/۷۱	۲/۴۵
MAPE (درصد)	۷۹/۴	۷۱/۶
همبستگی پیرسون	۰/۵۸	۰/۶۷
همبستگی اسپیرمن	۰/۵۴	۰/۶۴

جدول ۴. نتایج تخمین مدل دوجمله ای منفی

متغیر	ضریب تخمین	خطای استاندارد	آماره Z	مقدار p	نوع اثر
ثابت مدل	-۲/۶۱۲۵	۰/۵۸۱۳	-۴/۴۹	۰/۰۰۱>	تصادفی
ADTmaj	۰/۰۰۰۶	۰/۰۰۰۱	۶/۰۰	۰/۰۰۱>	تصادفی
ADTmin	۰/۰۰۰۹	۰/۰۰۰۱	۹/۰۰	۰/۰۰۱>	تصادفی
nleg	۰/۵۶۵۸	۰/۱۴۲۴	۳/۹۷	۰/۰۰۱>	ثابت
nsign	۰/۰۶۷۲	۰/۰۶۶۰	۱/۰۲	۰/۳۰۸	ثابت
mp1	-۰/۲۸۵۰	۰/۱۰۱۷	-۲/۸۰	۰/۰۰۵	ثابت
mp2	-۰/۰۸۰۵	۰/۰۹۸۳	-۰/۸۲	۰/۴۱۳	ثابت
ap1	-۰/۱۵۰۶	۰/۱۰۱۲	-۱/۴۹	۰/۱۳۶	ثابت
ap2	۰/۰۹۴۷	۰/۰۹۴۰	۱/۰۱	۰/۳۱۴	ثابت
ep1	۰/۰۱۳۲	۰/۰۸۹۹	۰/۱۵	۰/۸۸۳	ثابت
ep2	مرجع	-	-	-	-

پارامتر پراکندگی (θ): ۱/۲
 انحراف معیار اثرات تصادفی:
 ثابت مدل = ۰/۴۵
 ADTmaj = ۰/۰۰۰۲
 ADTmin = ۰/۰۰۰۳

اثرات ثابت

ثابت مدل (-۲/۶۱۲۵): مقدار منفی و معنادار ثابت مدل (نسبت خطای استاندارد ۰/۵۸۱۳) نشان می‌دهد که در شرایط پایه (زمانی که تمام متغیرهای کمی صفر بوده و ساعت مرجع یعنی ۱۹-۲۰ در نظر گرفته شود)، لگاریتم مورد انتظار تعداد تصادفات -۲/۶۱۲۵ است که معادل ۰/۰۷۳ تصادف در هر ساعت اوج می‌باشد. این مقدار پایه بسیار کم، قابل توجه است زیرا در غیاب ترافیک و با حداقل شاخه‌ها و چراغ‌ها، تصادفات نادر هستند. حجم ترافیک راه اصلی (با تخمین ۰/۰۰۰۶ و خطای ۰/۰۰۰۱): این ضریب مثبت و معنادار است. به ازای هر واحد افزایش در ADTmaj (یک خودرو در ساعت)، لگاریتم تعداد تصادفات به میزان ۰/۰۰۰۶ افزایش می‌یابد. اگر چه مقدار عددی کوچک است، اما با توجه به دامنه وسیع متغیر (۷۰ تا ۴۴۰۰)، تأثیر آن در عمل قابل ملاحظه است. برای مثال، افزایش ۱۰۰۰ خودرو در ساعت در راه اصلی، لگاریتم نرخ مورد انتظار تصادف را حدود ۸۲ درصد افزایش می‌دهد.

حجم ترافیک راه فرعی (۰/۰۰۰۹ با خطای استاندارد ۰/۰۰۰۱): ضریب مثبت و معنادار. تأثیر ADTmin نسبت به ADTmaj قدری بیشتر است. این یافته نشان می‌دهد که افزایش ترافیک در راه فرعی به ازای هر واحد، تأثیر بیشتری بر افزایش تصادفات دارد. برای مثال، افزایش ۱۰۰۰ خودرو در ساعت در راه فرعی، لگاریتم نرخ مورد انتظار تصادف را حدود ۱۴۶ درصد افزایش می‌دهد. این امر می‌تواند ناشی از تداخل جریان‌های ترافیکی در تقاطع‌ها باشد. تعداد شاخه‌ها (۰/۵۶۵۸ با خطای استاندارد ۰/۱۴۲۴): ضریب مثبت و معنادار. افزایش تعداد شاخه‌های تقاطع از سه به چهار شاخه، نرخ مورد انتظار تصادفات را حدود ۷۶ درصد افزایش می‌دهد. تعداد چراغ راهنما (۰/۰۶۷۲ با خطای استاندارد ۰/۰۶۶۰): ضریب مثبت اما از نظر آماری معنادار نیست (نسبت ضریب به خطای استاندارد حدود ۱/۰۲). این نشان می‌دهد که تعداد چراغ راهنما اثر معناداری بر تعداد تصادفات ندارد و ممکن است با سایر متغیرها همبسته باشد. متغیرهای مجازی ساعتی: ساعت مرجع (۱۹-۲۰) در نظر گرفته شده است.

هر تقاطع دقت متوسطی دارد ($MAD=2/45$ و $MAPE=71$).
با این حال مقدار R^2 شرطی (عملکرد مدل با احتساب اثرات تصادفی) برابر $0/71$ نشان از عملکرد خوب مدل دارد و مقدار R^2 حاشیه‌ای (عملکرد مدل با پارامترهای ثابت) برابر $0/49$ نشان از عملکرد متوسط مدل دارد که نتیجه‌ای متداول در ادبیات مدل‌سازی تصادفات است و مقداری قابل قبول است (کوهی و شعبانی، ۱۳۹۹). تفسیر نتیجه هر یک از سنجه‌های ارائه‌شده در جدول (۴) به شرح زیر است:

MAD - (میانگین قدر مطلق انحرافات): متوسط (میانگین خطای مطلق $2/45$ تصادف است یعنی برای یک تقاطع معمولی با میانگین ۴ تا ۵ تصادف، خطای نسبی حدود ۵۰ درصد است).

$MSPE$ - (میانگین مربعات خطای پیش‌بینی): متوسط (مجذور خطای هر پیش‌بینی $7/2$ واحد است. با توجه به میانگین داده‌ها که حدود $3/5$ تصادف در هر بازه است و پراکندگی بالا که برابر $1/2$ است، این مقدار نشان می‌دهد که مدل برای پیش‌بینی نقطه‌ای دقت متوسطی دارد).

$MAPE$ - (میانگین درصد خطای مطلق): طبیعی برای داده‌های شمارشی با مقادیر کم (تعداد تصادفات اغلب ۵-۰ است).
 MSE (میانگین مربعات خطا): قابل قبول (حتی بدون اثرات تصادفی، متغیرهای ثابت ۴۹ درصد از واریانس را توضیح می‌دهند. اکثر مدل‌های خطی تعمیم‌یافته در این حوزه R^2 حاشیه‌ای بین $0/20$ تا $0/50$ دارند).

R^2 - شرطی: مناسب (مدل با احتساب اثرات تصادفی، ۷۱ درصد از واریانس را توضیح می‌دهد. در ادبیات ایمنی، R^2 بیشتر از $0/60$ اغلب به عنوان برازش قوی تلقی می‌شود).

R^2 - حاشیه‌ای: متوسط (حتی بدون اثرات تصادفی، متغیرهای ثابت ۴۹ درصد از واریانس را توضیح می‌دهند).

-همبستگی پیرسون (مشاهده در برابر پیش‌بینی): همبستگی قوی بین پیش‌بینی و واقعیت (برای داده‌های شمارشی مناسب است).

-همبستگی اسپیرمن (رتبه‌ای): خوب (همبستگی رتبه‌ای $0/64$ نشان‌دهنده تطابق خوب رتبه‌بندی تقاطع‌ها توسط مدل با واقعیت است. برای اولویت‌بندی و تخصیص منابع اقدامات ایمنسازی مانند نصب چراغ، این سطح همبستگی عملیاتی است).

-ساعت ۶-۷ (با ضریب $0/2850$ - و خطای استاندارد $0/1017$): ضریب منفی و معنادار. در این ساعت، تصادفات ۲۵ درصد نسبت به ساعت مرجع کمتر است. ترافیک در ساعات اولیه صبح معمولاً کمتر است.

-ساعت ۷-۸ (با ضریب $0/0805$ - و خطای استاندارد $0/0983$): ضریب منفی اما معنادار نیست. تفاوت معناداری با مرجع ندارد.
-ساعت ۱۱-۱۲ (با ضریب $0/1506$ - و خطای استاندارد $0/1012$): ضریب منفی و در مرز معناداری (نسبت $1/49$). ممکن است کاهش نسبی تصادفات در این ساعت وجود داشته باشد.

-ساعت ۱۲-۱۳ (با ضریب $0/0947$ و خطای استاندارد $0/0940$): ضریب مثبت اما معنادار نیست. نشانه‌ای از افزایش جزئی.
-ساعت ۱۸-۱۹ (با ضریب $0/0132$ و خطای استاندارد $0/0899$): ضریب مثبت و بسیار کوچک؛ همچنین معنادار نیست.

در مجموع، تنها ساعت (۶-۷) کاهش معنادار نسبت به ساعت مرجع را نشان می‌دهد؛ سایر هنگامه‌ها تفاوت معناداری ندارند. این الگو می‌تواند ناشی از توزیع تصادفات در طول روز و تفاوت در ترکیب ترافیک باشد.

اثرات تصادفی

انحراف معیار اثر تصادفی ثابت مدل معادل $0/45$ برآورد شد. این مقدار نشان‌دهنده وجود ناهمگنی قابل توجه بین تقاطع‌ها است. یعنی پس از کنترل متغیرهای توضیحی، هنوز تفاوت‌های تصادفی سیستماتیک بین تقاطع‌ها باقی می‌ماند (برای مثال تفاوت در هندسه، دید، روشنایی و غیره که اندازه‌گیری نشده‌اند). انحراف معیار اثرات تصادفی برای ضرایب ADT_{maj} و ADT_{min} به ترتیب $0/0002$ و $0/0003$ است. این مقادیر بسیار کوچک‌اند که نشان می‌دهد اثر حجم ترافیک بر تصادفات در بین تقاطع‌ها نسبتاً همگن است و نیازی به مدل‌سازی همراه با اثرات تصادفی نبوده است. با این حال از آنجاکه از این روش استفاده شده با اطمینان بیشتری می‌توان نتایج را ارائه داد.

۳-۴- عملکرد مدل

مقادیر سنجه‌های ارزیابی عملکرد مدل در جدول (۵) ارائه شده است. این شاخص‌ها براساس داده‌های آزمون (۲۰ درصد داده‌های اولیه) محاسبه شدند تا از بیش‌برازش مدل جلوگیری شود. نتایج نشان می‌دهد که مدل در پیش‌بینی نقطه‌ای تعداد تصادفات برای

جدول ۵. نتایج ارزیابی عملکرد مدل

ردیف	سنجه	مقدار
۱	MAD	۲/۴۵
۲	MSPE	۷/۲۰
۳	MAPE	٪ ۷۱/۶
۴	MSE	۱۴/۵۲
۵	R ² شرطی	۰/۷۱
۶	R ² حاشیه‌ای	۰/۴۹
۷	همبستگی پیرسون	۰/۶۷
۸	همبستگی اسپیرمن	۰/۶۴

۴-۴- مدل نهایی

مدل نهایی نشان می‌دهد که حجم ترافیک راه اصلی و فرعی و تعداد شاخه‌ها مهمترین پیشبینی کننده‌های تصادفات در تقاطع‌ها هستند و ساعات اولیه صبح (۶-۷) نسبت به عصر تصادفات کمتری دارد. با توجه به نتایج تخمین مدل که در جدول (۳) نشان داده شده است، مدل نهایی پیشبینی تصادفات ساعات اوج در تقاطع‌های درون‌شهری مطابق معادله (۵) است که در آن متغیرهای X_i و Y_i

مطابق جدول (۶) است. نکته مهم برای استفاده از این مدل این است که در هر بازه زمانی فقط یکی از متغیرهای ساعتی (Y_1 تا Y_5) مقدار ۱ دارد و بقیه ۰ هستند (متغیر Y_6 به دلیل اینکه ساعت مرجع است در مدل وجود ندارد). اگر ساعت مورد نظر جزو هیچ‌کدام از این بازه‌ها نباشد، تمام Y_i ها صفر هستند. برای مثال اگر بخواهیم تعداد تصادفات را برای بازه زمانی (۶-۷ صبح) پیشبینی کنیم، مقدار Y_1 در مدل ۱ است و بقیه Y ها صفر هستند.

$$\ln(\mu) = -2.6125 + 0.0006 \cdot X_1 + 0.0009 \cdot X_2 + 0.5658 \cdot X_3 + 0.0672 \cdot X_4 - 0.2850 \cdot Y_1 - 0.0805 \cdot Y_2 - 0.1506 \cdot Y_3 + 0.0947 \cdot Y_4 + 0.0132 \cdot Y_5 \quad (5)$$

جدول ۶. متغیرها در مدل نهایی

نماد	متغیر	شرح
X1	ADTmaj	میانگین ساعتی ترافیک روزانه در راه اصلی
X2	ADTmin	میانگین ساعتی ترافیک روزانه در راه فرعی
X3	nleg	تعداد شاخه‌های تقاطع
X4	nsign	تعداد چراغ راهنما
Y1	mp1	ساعت ۶-۷
Y2	mp2	ساعت ۷-۸
Y3	ap1	ساعت ۱۱-۱۲
Y4	ap2	ساعت ۱۲-۱۳
Y5	ep1	ساعت ۱۸-۱۹

۵- نتیجه گیری

به عنوان ابزاری برای غربالگری و اولویت بندی اقدامات ایمنی مورد استفاده قرار گیرد.

با وجود نتایج مطلوب، پژوهش حاضر دارای محدودیت هایی نیز هست. نبود برخی متغیرهای هندسی و محیطی نظیر زاویه تقاطع، فاصله دید، شرایط روشنایی و وضعیت روسازی، می تواند بخشی از ناهمگونی مشاهده نشده را توضیح دهد. همچنین داده های این پژوهش محدود به یک شهر بوده و تعمیم نتایج به سایر شهرها نیازمند مطالعات تکمیلی است. پیشنهاد می شود در مطالعات آینده، از داده های چندین شهر، متغیرهای رفتاری و محیطی و همچنین مدل های مکانی-زمانی و روش های یادگیری ماشین برای مقایسه عملکرد مدل ها استفاده شود.

در مجموع، نتایج این پژوهش نشان می دهد که مدل دوجمله ای منفی با اثرات تصادفی (RENB)، ابزار مناسبی برای تحلیل تصادفات ساعات اوج در تقاطع های شهری بوده و می تواند در توسعه سامانه های تصمیم یار ایمنی و مدیریت هوشمند ترافیک شهری مورد استفاده قرار گیرد.

۶- پانویس

1. Negative Binomial
2. Random Effects Negative Binomial

۷- مراجع

-کوهی، محمد و شعبانی، شاهین (۱۳۹۹). مدلسازی فراوانی تصادفات در تقاطع های سه شاخه شهری. فصلنامه جاده، دوره ۲۸، شماره ۱۰۳. ۵۰-۳۷.

-Abdel-Aty, M. A., & Radwan, A. E. (2000). Modeling traffic accident occurrence and involvement. *Accident Analysis & Prevention*, 32(5), 633-642.

-Almutairi, O., (2024), A nested grouped random parameter negative binomial model for modeling segment-level crash counts, *Heliyon*, 10, e28900.

-Alqatawna, A., Rivas Álvarez, A. M., & García-Moreno, S. S.-C. (2021). Comparison of Multivariate Regression Models and Artificial Neural Networks for Prediction Highway Traffic

در این پژوهش، مدل دوجمله ای منفی با اثرات تصادفی (RENB) به منظور تحلیل و پیش بینی فراوانی تصادفات ساعات اوج روزانه در ۳۴ تقاطع شهری شهر بروجرد توسعه داده شد. نتایج نشان داد که داده های تصادف دارای بیش پراکندگی معنادار بوده و استفاده از مدل های شمارشی پیشرفته نسبت به مدل های خطی و پواسون، قابلیت بیشتری در تحلیل این نوع داده ها دارد. همچنین مقایسه مدل RENB با مدل دوجمله ای منفی کلاسیک نشان داد که لحاظ کردن ناهمگونی مشاهده نشده بین تقاطع ها موجب بهبود معنادار برازش مدل شده است. مقدار بالای R^2 شرطی و نتایج آزمون نسبت درست نمایی، کارایی مناسب ساختار اثرات تصادفی را تأیید نمود. یافته های پژوهش نشان داد که حجم ترافیک راه فرعی نسبت به راه اصلی اثر قوی تری بر افزایش تصادفات دارد. این نتیجه بیانگر آن است که در ساعات اوج، تداخل جریان های ترافیکی ورودی از راه های فرعی نقش مهمی در افزایش ریسک تصادف ایفا می کند. بنابراین تقاطع هایی که دارای حجم بالای ترافیک در راه فرعی هستند، نیازمند توجه ویژه در طراحی و کنترل ترافیکی می باشند. همچنین نتایج نشان داد که افزایش تعداد شاخه های تقاطع، به طور قابل توجهی احتمال وقوع تصادفات را افزایش می دهد؛ موضوعی که اهمیت ساده سازی هندسی و کاهش نقاط تعارض ترافیکی را آشکار می سازد.

بررسی متغیرهای زمانی نیز نشان داد که تنها بازه زمانی ۶ تا ۷ صبح دارای کاهش معنادار تصادفات نسبت به ساعت مرجع ۱۹ تا ۲۰ است، در حالی که سایر ساعات اوج تفاوت معناداری با ساعت مرجع نداشتند. این یافته بیانگر آن است که ساعات عصرگاهی، به ویژه بازه ۱۹ تا ۲۰، بحرانی ترین دوره از نظر خطر وقوع تصادف در تقاطع های شهری محسوب می شوند. از منظر مدیریتی، این نتایج می تواند در برنامه ریزی زمان بندی چراغ های راهنمایی، تخصیص نیروهای کنترل ترافیک و زمان بندی عملیات اجرایی و تعمیراتی مورد استفاده قرار گیرد.

ارزیابی عملکرد مدل نشان داد که مدل پیشنهادی با وجود ماهیت گسسته و تصادفی داده های تصادف، از توان پیش بینی قابل قبولی برخوردار است. مقدار همبستگی اسپیرمن نشان داد که مدل توانایی مناسبی در رتبه بندی تقاطع ها از نظر خطر تصادف دارد و می تواند

- application using crash data. *Accid. Anal. Prev.* 45, 258–265.
- Gong, H., Wang, F., Zhou, B., Dent, Sh., (2020). Application of random effects negative binomial model with clustered dataset for vehicle crash frequency analysis, *International Journal of Transportation Science and Technology*, 9, 183–194.
- Gu, X., Yan, X., Ma, L., & Liu, X. (2020). Modeling the service-route-based crash frequency by a spatiotemporal-random-effect zero-inflated negative binomial model: An empirical analysis for bus-involved crashes. *Accident Analysis & Prevention*, 144, 105674.
- Guo, M., Zhao, X., Yao, Y., Yan, P., Su, Y., Bi, C., & Wu, D. (2021). A study of freeway crash risk prediction and interpretation based on risky driving behavior and traffic flow data. *Accident Analysis & Prevention*, 160, 106328.
- Islam, A. S. M. M., Shirazi, M., & Lord, D. (2022). Finite mixture Negative Binomial Lindley for modeling heterogeneous crash data with many zero observations. *Accident Analysis & Prevention*, 175, 106765.
- Joshua, S. C., & Garber, N. J. (1990). Estimating truck accident rate and involvements using linear and Poisson regression models. *Transportation Planning and Technology*, 15(1), 41–58.
- Jovanis, P. P., & Chang, H. L. (1986). Modeling relationship of accidents to miles traveled. *Transportation Research Record Journal of the Transportation Research Board* 1068.
- Lord, D., Manar, A., Vizioli, A., (2005). Modeling crash-flow-density and crash-flow-V/C ratio relationships for rural and urban freeway segments. *Accid. Anal. Prev.* 37 (1), 185–199.
- Lord, D., Mannering, F., (2010). The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives. *Transp. Res. A Policy Pract.* 44 (5), 291–305.
- Mannering, F.L., Shankar, V., Bhat, C.R., (2016). Unobserved heterogeneity and the statistical analysis of highway accident data. *Anal. Methods Accid. Res.* 11, 1–16.
- Mathew, S., Pulugurtha, S. S., & Duvvuri, S. (2022). Exploring the effect of road network, demographic, and land use characteristics on teen crash frequency using geographically weighted negative binomial regression. *Accident Analysis & Prevention*, 168, 106615.
- Accidents in Spain: A Case Study. *Transportation Research Procedia*, 58, 277–284.
- Anastasopoulos, P.C., Mannering, F.L., (2009). A note on modeling vehicle accident frequencies with random-parameters count models. *Accid. Anal. Prev.* 41 (1), 153–159.
- Anastasopoulos, P.Ch., Mannering, F.L., (2009). A note on modeling vehicle accident frequencies with random-parameters count models. *Accid. Anal. Prev.* 41, 153–159.
- Azmeri Khan, Sh., Afghari, A. P., Yasmin, Sh., -Haque, M., (2023). Effects of design consistency on run-off-road crashes: An application of a Random Parameters Negative Binomial Lindley model, *Accident Analysis and Prevention*, 186, 107042.
- Basso, F., Basso, L. J., Bravo, F., & Pezoa, R. (2018). Real-time crash prediction in an urban expressway using disaggregated data. *Transportation Research Part C: Emerging Technologies*, 86, 202–219.
- Cai, Q., Abdel-Aty, M., Yuan, J., Lee, J., & Wu, Y. (2020). Real-time crash prediction on expressways using deep generative models. *Transportation Research Part C: Emerging Technologies*, 117, 102697.
- Champahom, T., Jomnonkwao, S., Karoonsoontawong, A., & Ratanavaraha, V. (2022). Spatial zero-inflated negative binomial regression models: Application for estimating frequencies of rear-end crashes on Thai highways. *Journal of Transportation Safety & Security*, 14(3), 523–540.
- Chen, E., Tarko, A.P., (2014). Modeling safety of highway work zones with random parameters and random effects models. *Anal. Methods Accid. Res.* 1, 86–95.
- Chin, H.C., Quddus, M.A., (2003). Applying the random effect negative binomial model to examine traffic accident occurrence at signalized intersections. *Accid. Anal. Prev.* 35, 253–259.
- Fu, X., Liu, J., Jones, S., Barnett, T., & Khattak, A. J. (2022). From the past to the future: Modeling the temporal instability of safety performance functions. *Accident Analysis & Prevention*, 167, 106592.
- Geedipally, S.R., Lord, D., Dhavala, S.S., (2012). The negative binomial-Lindley generalized linear model: Characteristics and

- Tavakoli Kashani, A., Mirhashemi, A., Amirifar, S., (2022), A review on Intersection Safety Studies with Bibliometric Methods, *Amirkabir Journal of Civil Engineering*, 54(5), 363-366.
- Washington, S., Karlaftis, M.G., Mannering, F., Anastasopoulos, P., (2020). Statistical and econometric methods for transportation data analysis. *CRC Press*.
- Xu, P., & Huang, H. (2015). Modeling crash spatial heterogeneity: Random parameter versus geographically weighting. *Accident Analysis & Prevention*, 75, 16–25.
- Young, J., & Park, P. Y. (2013). Benefits of small municipalities using jurisdiction-specific safety performance functions rather than the Highway Safety Manual's calibrated or uncalibrated safety performance functions. *Can. J. Civ. Eng.*, 40(6), 517–527.
- Yuan, J., Abdel-Aty, M., Fu, J., Wu, Y., Yue, L., & Eluru, N. (2021). Developing safety performance functions for freeways at different aggregation levels using multi-state microscopic traffic detector data. *Accident Analysis & Prevention*, 151, 105984.
- Yuan, J., Abdel-Aty, M., Fu, J., Wu, Y., Yue, L., & Eluru, N. (2021). Developing safety performance functions for freeways at different aggregation levels using multi-state microscopic traffic detector data. *Accident Analysis & Prevention*, 151, 105984.
- Zegeer, C., Stewart, R., Reinfurt, D., Neuman, T., Hamilton, E., Miller, T., & Hunter, W. (1991). Cost-effective geometric improvements for safety upgrading of horizontal curves. *TRID - TRB*, 237.
- Miaou, S.-P. (1994). The relationship between truck accidents and geometric design of road sections: Poisson versus negative binomial regressions. *Accident Analysis & Prevention*, 26(4), 471–482.
- Mitra, S., Washington, S., 2007. On the nature of over-dispersion in motor vehicle crash prediction models. *Accid. Anal. Prev.* 39 (3), 459–468.
- Moomen, M., Rezapour, M., Raja, M., Ksaibati, K., (2020). Predicting downgrade crash frequency with the random-parameters negative binomial model: Insights into the impacts of geometric variables on downgrade crashes in Wyoming, *LATSS Research*, 44, 94–102.
- Oviedo-Trespalacios, O., Afghari, A.P., Haque, M.M., (2020). A hierarchical Bayesian multivariate ordered model of distracted drivers' decision to initiate risk- compensating behaviour. *Anal. Methods Accid. Res.* 26, 100121.
- Poch, M., Mannering, F., (1996). Negative binomial analysis of intersection-accident frequencies. *J. Transp. Eng.* 122 (2), 105–113.
- Retallack, A. E., & Ostendorf, B. (2020). Relationship between Traffic Volume and Accident Frequency at Intersections. *IJERPH*, 17(4), 1393.
- Shankar, V., Mannering, F., & Barfield, W. (1995). Effect of roadway geometrics and environmental factors on rural freeway accident frequencies. *Accident Analysis & Prevention*, 27(3), 371–389.

Analysis and Prediction of Accident Frequency during Peak Hours at Urban Intersections Using the Random-Effects Negative Binomial (RENB) Model

*Shahin Shabani, Department of Civil Engineering,
Payame Noor University (PNU), Tehran, Iran.*

*Amir Bahador Azimi, Ph.D., Student, Department of Civil Engineering,
Payame Noor University (PNU), Tehran, Iran.*

E-mail: shabani@pnu.ac.ir

Received: March 2026- Accepted: August 2026

ABSTRACT

This study aimed to predict crash frequency during daily peak hours at urban intersections using a Random-Effects Negative Binomial (RENB) model. The dataset comprised crash records from 34 urban intersections in Borujerd, Iran, covering a five-year period from 2021 to 2025 (corresponding to the Iranian calendar years 1400–1404). After aggregating the data on an hourly basis, the final dataset consisted of 204 observations and 698 reported crashes. The explanatory variables included hourly traffic volumes on major and minor roads, the number of intersection approaches, the number of traffic signals, and dummy variables representing peak-hour periods. Statistical analysis confirmed the presence of significant overdispersion in the crash data. Moreover, the relatively low proportion of zero-count observations indicated that the Negative Binomial model was more appropriate than zero-inflated alternatives. A comparison between the conventional Negative Binomial (NB) model and the Random-Effects Negative Binomial (RENB) model, based on the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and Likelihood Ratio (LR) test, demonstrated that incorporating random effects significantly improved model performance. The results revealed that traffic volume on minor roads had a greater impact on crash frequency than traffic volume on major roads. Specifically, an increase of 1,000 vehicles per hour on the minor road was associated with an approximately 146% increase in the expected crash frequency. Furthermore, increasing the number of intersection approaches from three to four resulted in a 76% increase in the expected crash frequency.

Keywords: Accident Modeling, Peak Hours, Random-Effects Negative Binomial Model, Urban Intersections