

تحلیل فضایی تعداد تصادفات رانندگی منجر به فوت در کلان‌شهر تهران

زهرا فضلعلی، کیومرث مترجم^۱
گروه آمار دانشگاه تربیت مدرس

چکیده: در این مقاله با استفاده از داده‌های مربوط به تصادفات رانندگی منجر به فوت در مناطق ۲۲ گانه شهر تهران در سال ۱۳۹۴ به تحلیل فضایی تعداد تصادفات پرداخته شده است. سپس با استفاده از مدل اتورگرسیون فضایی پس از مدل‌بندی کردن داده‌ها و انجام پیشگویی فضایی در قالب نقشه، به بررسی عوامل موثر بر تصادفات رانندگی نیز پرداخته شده است.

واژه‌های کلیدی: آمار فضایی، ترافیک، تصادفات رانندگی، مدل اتورگرسیون فضایی
کد موضوع‌بندی ریاضی (۲۰۱۰): 62M20، 62J12، 62H11.

۱ مقدمه

داده‌های تصادفات جاده‌ای^۱ به داده‌های اطلاق می‌شود که بر حسب زمان و مکان وقوع تصادفات جمع‌آوری می‌شوند. این گونه داده‌ها مورد علاقه فراوان محققان به ویژه متخصصان آمار فضایی هستند زیرا این نوع داده‌ها عموماً بر حسب زمان و مکان رخداد همبستگی فضایی دارند. در این مقاله با در نظر گرفتن همبستگی فضایی مکان وقوع تصادفات به تحلیل این دسته از داده‌ها پرداخته شده است. در بسیاری از موارد که تجزیه و تحلیل رگرسیون برای داده‌های وابسته به مکان اعمال می‌شود، مشکلات مربوط به همبستگی فضایی و سایر اثرات مکانی ممکن است منجر به افزایش خطا در مدل شود. به همین علت، روش‌های استاندارد مدل‌سازی در رگرسیون چند متغیره ناکارآمد می‌شوند زیرا همبستگی فضایی اثرات نامطلوبی بر برآورد ضرایب رگرسیون خواهد گذاشت. صرف نظر از اینکه توزیع آماری متغیر پاسخ نرمال، پواسون یا دوجمله‌ای باشد، تحلیل اکتشافی داده‌ها و استفاده از مدل‌های فضایی، موجب استنباط بهتری از داده‌ها خواهد شد و قدرت جدیدی را در

^۱Road Traffic Crashes

تجزیه و تحلیل داده‌های فضایی به ارمغان می‌آورد. همچنین این مدل‌ها قابلیت لحاظ کردن اثرات پنهان فضایی را نیز دارند. توانایی‌های توصیفی روش‌های آمار فضایی درک بهتری را از داده‌های مشاهده شده را به تصویر می‌کشد و رویکردهای توسعه یافته‌ای را ارائه می‌دهد که به نوبه خود سبب بهبود در برازش مدل و بهبود برآورد پارامترها می‌شود.

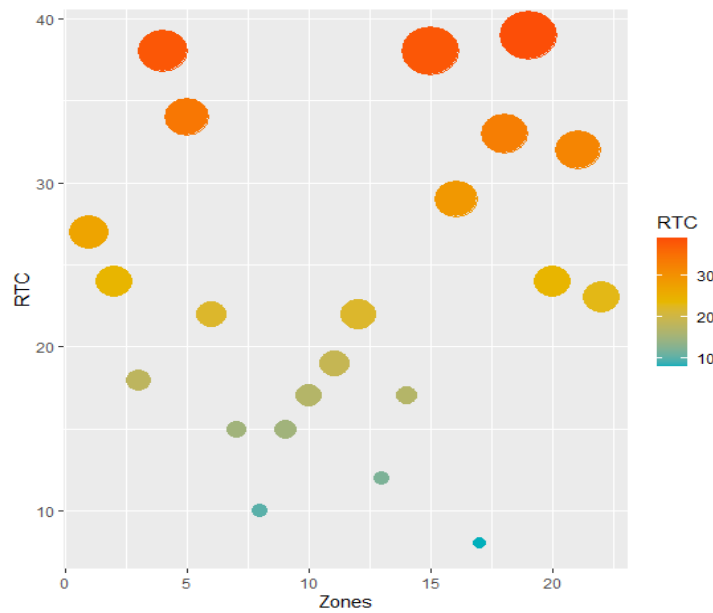
برای ارزیابی عوامل موثر بر تصادفات جاده‌ای مطالعات زیادی در کشورهای مختلف دنیا انجام شده است. در این مطالعات روش‌های مختلف مدل‌سازی برای اهداف تحلیلی و پیش‌بینی استفاده شده است. که از آن جمله می‌توان به مدل‌های مبتنی بر توزیع پواسون، پواسون-گاما و پواسون-لگ نرمال اشاره نمود. **جوانیس و چانگ (۱۹۸۶)**، **جونز و همکاران (۱۹۹۱)** و پس از آن **ماهر و موشن (۱۹۸۸)** سپس **روکه کاردوسو (۲۰۱۴)** از مدل‌های رگرسیون پواسون برای مطالعه تصادفات جاده‌ای استفاده کردند. همچنین **یه و همکاران (۲۰۰۹)** و **یو و عبدل (۲۰۱۳)** بر اساس فراوانی تعداد تصادفات از مدل رگرسیونی چند متغیره پواسن استفاده کردند. **جیانگ و همکاران (۲۰۱۴)** برای شناسایی کانون خطر تصادفات از مدل‌های پواسن لگ-نرمال با اثر تصادفی استفاده کردند. **لویین و همکاران (۱۹۹۵)** با در نظر گرفتن همبستگی فضایی از مدل‌های خود همبسته فضایی در تحلیل داده‌های تصادفات جاده‌ای استفاده کردند. **یو و همکاران (۲۰۱۳)** برای بررسی پویایی اثرات تصادفی، مدل‌های رگرسیون لجستیک را بکار بردند. علاوه بر این مطالعات گذشته **اردوغان (۲۰۰۹)** تصادفات جاده‌ای را با استفاده از مدل رگرسیون وزن جغرافیایی مطالعه کرد. **عبدالله و زامری (۲۰۱۲)** نیز از رویکرد رگرسیون فازی برای تحلیل داده‌های تصادفات استفاده کردند.

در ادامه این مقاله در بخش ۲ به معرفی و تحلیل اکتشافی داده‌های مربوط به تعداد تصادفات رانندگی در کلانشهر تهران بر اساس اطلاعات (سالنامه آماری شهر تهران، ۱۳۹۴). پرداخته می‌شود. در بخش ۳ مدل اتورگرسیو فضایی معرفی خواهد شد و نحوه کاربست آن در تحلیل داده‌های تصادفات رانندگی شهر تهران مورد بررسی قرار خواهد گرفت و در بخش ۴ به بحث و بررسی نتایج پرداخته می‌شود.

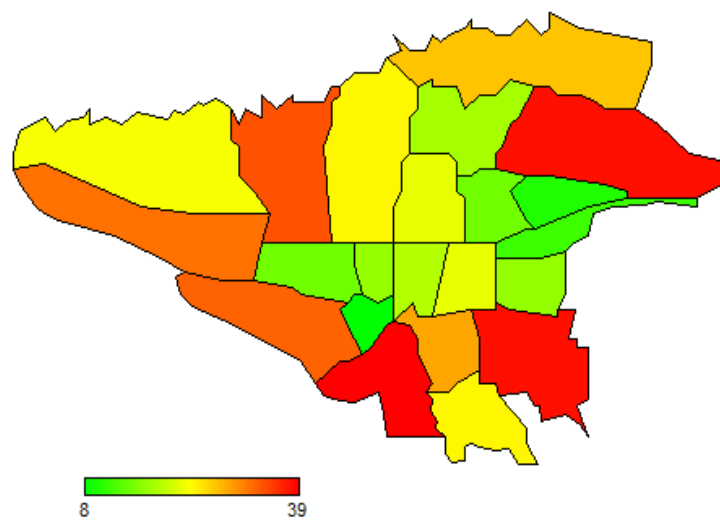
۲ معرفی داده‌های تصادفات رانندگی شهر تهران

در این مطالعه با استفاده از داده‌های تصادفات رانندگی منجر به فوت مربوط به مناطق بیست و دو گانه شهر تهران در سال ۱۳۹۴ به تحلیل فضایی آنها با استفاده از مدل اتورگرسیو فضایی پرداخته شده است. بر اساس داده‌ها بیشترین تعداد تصادفات منجر به فوت در منطقه ۱۹ با ۳۹ مورد فوتی و کمترین تعداد تصادفات در منطقه ۱۷ با ۸ مورد فوت ثبت شده است. میانه تعداد تصادفات ۲۴ است بر این اساس ۱۲ منطقه تعداد تصادفات کمتر از ۲۴ و ۸ منطقه تعداد تصادفات بیشتر از میانه دارند.

آمارها نشان می‌دهد که میانه تعداد تصادفات منجر به فوت در مناطق ۲ و ۲۰ با ۲۴ مورد فوتی رخ داده است یعنی ۵۰ درصد از نواحی دارای تصادفات منجر به فوت کمتر از مناطق ۲ و ۲۰ و ۵۰ درصد دیگر تعداد تصادفات منجر به فوتشان بیش از این مناطق است. همچنین ۲۵ درصد مناطق بیش از ۳۰ مورد تصادفات منجر به فوت داشته‌اند، در حالی که در ۷۵ درصد بقیه بین ۸ تا ۳۰ مورد تصادف رخ داده است. در ۴۱ درصد از مناطق تعداد تصادفات رانندگی منجر به فوت از ۲۰ مورد در سال کمتر است. حدود ۳۱ درصد از نواحی نیز بین ۲۰ تا ۳۰ تصادف منجر به فوت در طی سال ۱۳۹۴ داشته‌اند. نمودار حبابی داده‌ها در شکل شماره ۱ آمده است در این نمودار اندازه هر دایره متناسب با تعداد تصادفات در منطقه مورد نظر است یعنی دایره‌های بزرگتر نشان دهنده تعداد تصادفات بیشتر می‌باشد. شکل شماره ۲ نحوه توزیع داده‌های تصادفات رانندگی را در مناطق شهر تهران نشان می‌دهد این نمودار بیانگر همبستگی فضایی تصادفات در مناطق ۲۲ گانه شهر تهران است. در شکل ۳ میزان پراکندگی داده‌ها با استفاده از نمودار جعبه‌ای مشخص شده است.



شکل ۱: وضعیت تعداد تصادفات رانندگی در مناطق ۲۲ گانه شهر تهران



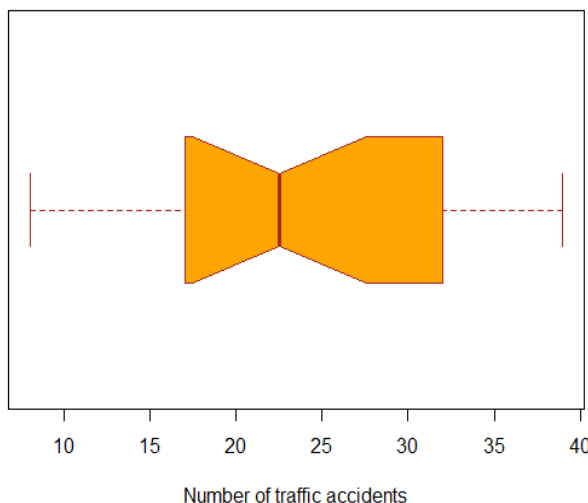
شکل ۲: نقشه فضایی تصادفات رانندگی مناطق ۲۲ گانه شهر تهران در سال ۱۳۹۴

۳ مدل اتورگرسیو فضایی

در این مطالعه برای تجزیه و تحلیل داده‌های مربوط به تعداد تصادفات رانندگی منجر به فوت در مناطق ۲۲ گانه شهر تهران از مدل اتورگرسیو فضایی^۲ استفاده شده که بصورت زیر تعریف می‌شود

$$Y = \rho W(y) + X\beta + \varepsilon \tag{۱.۳}$$

^۲Spatial Autoregressive Models (SAR)



شکل ۳: نمودار جعبه‌ای تصادفات رانندگی در مناطق ۲۲ گانه شهر تهران

که در آن Y بردار متغیر پاسخ $1 \times N$ بعدی که در این مطالعه بیانگر تعداد تصادفات رانندگی در مناطق ۲۲ گانه شهر تهران است. ρ ضریب تأخیر فضایی است. W یک ماتریس وزن فضایی $N \times N$ معلوم است که عناصر روی قطر آن صفر است. در تعریف همسایگی در این مدل دو منطقه که با هم مرز مشترک داشته باشند را دو ناحیه همسایه می‌گوییم. ماتریس وزن‌های فضایی با استفاده از نقشه جغرافیایی و نحوه قرار گرفتن مناطق مختلف در کنار هم تعریف می‌شود بر این اساس اگر دو منطقه مرز مشترک داشته باشند در ماتریس وزن عدد یک و در غیر اینصورت عدد صفر را قرار می‌دهیم. با توجه به تعریف ماتریس وزن‌های فضایی $(I_N - \rho W)$ برای تمامی مقادیر $|\rho| \leq 1$ ناکین است. که در آن I_N یک ماتریس همانی است و $W(y)$ یک ماتریس وزنی از بردار $1 \times N$ بعدی مشاهدات است و X ماتریس مشاهدات است که شامل k متغیر تبیینی است. در این مطالعه، مساحت منطقه بر حسب هکتار و جمعیت منطقه به عنوان متغیرهای تبیینی در مدل استفاده شده است و فرض بر این است که X یک ماتریس رتبه کامل ستونی است. β یک بردار $1 \times k$ بعدی از پارامترها است. ε یک بردار $1 \times N$ بعدی مربوط به خطای تصادفی است که فرض می‌شود دارای توزیع نرمال با میانگین و واریانس ثابت است بطوریکه $E(\varepsilon|X) = 0$ برقرار باشد. عموماً برای برآورد پارامترهای مدل از روش حداقل مربعات معمولی استفاده می‌شود. واضح است که برآورد ماکسیمم درستنمایی برای β تابعی از ρ است بنابراین هنگامی که مقدار ρ تعیین شد، می‌توان مستقیماً برآورد β را بدست آورد. در این مقاله برای برآورد پارامترهای مدل از نرم‌افزار R استفاده شده است.

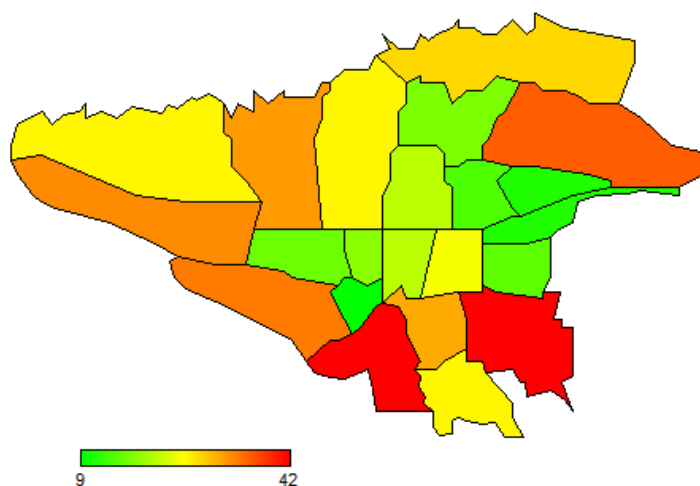
با استفاده از مدل ۱.۳ و با در نظر گرفتن ماتریس همسایگی مرز مشترک پارامترهای مدل اتورگرسیو فضایی برآورد شده است. همچنین نتایج برآورد پارامترهای مدل در جدول ۱ نشان داده شده است. نکته مهم برآورد مقدار ρ است که مثبت و قابل توجه است و نشان‌دهنده وجود همبستگی فضایی در داده‌های مربوط به تصادفات است. به عبارت دیگر در تصادفات رانندگی در شهر تهران همبستگی فضایی جز عوامل مهم و معنی دار تشخیص داده شده است.

برآورد پارامترها برای متغیرهای جمعیت و مساحت مناطق نشان می‌دهد که این دو متغیر در تعداد رخداد تصادفات رانندگی موثرند و البته در این بین سهم مساحت منطقه تا حدودی کمتر و حتی در سطح ۹۵ درصد معنی دار نیست. با این حال برآورد معنی دار تأخیر فضایی نشان از همبستگی فضایی بالایی در داده‌های مربوط به تعداد تصادفات شهر تهران دارد. با در نظر گرفتن علامت پارامترهای برآورد شده در مدل می‌توان گفت جمعیت به طور هم جهت با تعداد تصادفات منجر به فوت ارتباط دارد این تأثیر برای مساحت نیز مثبت بوده و نشان می‌دهد تعداد تصادفات با مساحت هر ناحیه رابطه مستقیم دارد. شکل ۴ نقشه فضایی پیشگویی تصادفات برای نواحی مختلف و همچنین شکل ۵ واریانس پیشگویی فضایی

جدول ۱: برآورد پارامترهای مدل اتورگرسیو فضایی برای تعداد تصادفات رانندگی در مناطق ۲۲ گانه شهر تهران

متغیر	برآورد	انحراف استاندارد	z - مقدار	p - مقدار
تاخیر فضایی	۰/۵۲	۰/۰۸	۳/۶۸	۰/۰۱
عرض از مبدا	۲/۱۱	۰/۴۹	۰/۰۹	۰/۸۳
مساحت منطقه	۰/۲۲	۰/۰۲	۳/۲۱	۰/۰۶
جمعیت منطقه	۰/۶۲	۰/۱۴	۲/۹۸	۰/۰۴

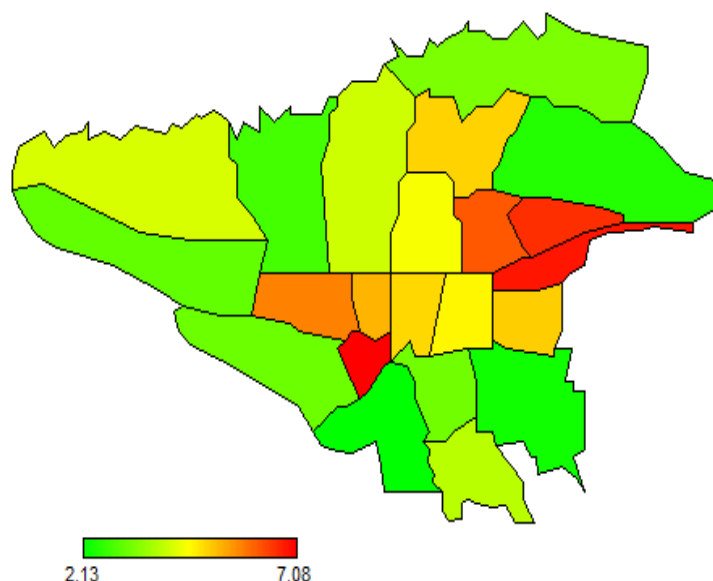
را نشان می‌دهد.



شکل ۴: نقشه فضایی پیشگویی تعداد تصادفات رانندگی در مناطق ۲۲ گانه شهر تهران

۴ نتیجه‌گیری

در این مطالعه با استفاده از داده‌های تصادفات درون شهری تهران در سال ۱۳۹۴ که منجر به فوت شده‌اند به تحلیل فضایی تعداد تصادفات پرداخته شد. نتایج نشان داد که همبستگی فضایی قابل توجهی در تعداد تصادفات رانندگی منجر به فوت در مناطق ۲۲ گانه شهر تهران وجود دارد. همچنین نشان داده شد که مناطقی با مساحت بیشتر، تعداد تصادفات بیشتری دارند و عامل جمعیت نیز موجب افزایش تعداد تصادفات در مناطق مختلف شهر تهران می‌گردد. طبق گزارش سازمان بهداشت جهانی، صدمات ناشی از ترافیک جاده‌ای در سال ۲۰۱۶ باعث حدود ۱/۳۵ میلیون مرگ در سراسر جهان شده است. در مبحث تصادفات جاده‌ای، ایران رتبه دوم را از میان ۱۹۰ کشور در دنیا دارد. میزان تلفات سوانح رانندگی در ایران ۲۵ برابر ژاپن، ۲ برابر ترکیه و تا ۱۰۰ برابر بیشتر از برخی کشورهای دنیا است. با این که تعداد وسایل نقلیه در انگلستان بیش از سه برابر خودروهای موجود در ایران است اما تعداد تصادفات ۳۲ برابر از ایران کمتر است؛ یعنی می‌توان گفت به نسبت خودروهای موجود در ایران، حدود ۱۰۰ برابر بیشتر از انگلستان تصادفات رانندگی رخ می‌دهد. با توجه به اهمیت تصادفات رانندگی در ایران و به ویژه کلانشهر تهران که رخداد تصادفات نه تنها خسارت مستقیم مالی و جانی برای شهروندان ایجاد می‌کند بلکه با ایجاد ترافیک سنگین هزینه‌های پنهان دیگری نیز بر جامعه تحمیل می‌کند. لذا در این مقاله با در نظر



شکل ۵: واریانس پیشگویی تصادفات رانندگی در مناطق ۲۲ گانه شهر تهران

گرفتن داده‌های مربوط به تعداد تصادفات در کلانشهر و با در نظر گرفتن موقعیت مکانی تصادفات به پیشگویی فضایی تعداد تصادفات و عوامل موثر بر آن پرداخته شد. در مطالعات آتی می‌توان آمار تعداد تصادفات را با استفاده از مدل‌های فضایی-زمانی مورد بررسی دقیق‌تر قرار داد همچنین اثر عواملی مانند متوسط تعداد وسیله نقلیه در هر منطقه و همچنین سن و جنس رانندگان حادثه‌ساز را نیز در مدل وارد نمود تا از طریق کنترل این عوامل و اتخاذ سیاست‌های مناسب عبور و مرور زمینه برای کاهش تعداد تصادفات رانندگی در شهر تهران فراهم شود.

تقدیر و تشکر

نویسنده از حمایت قطب علمی تحلیل داده‌های وابسته فضایی و فضایی-زمانی دانشگاه تربیت مدرس کمال تشکر و قدردانی را دارد.

مراجع

سالنامه آماری شهر تهران، سازمان فن‌آوری اطلاعات و ارتباطات شهرداری تهران (۱۳۹۴)، tmicto.tehran.ir.

Abdullah L. Zamri N. (2012)., Road accident models with two threshold levels of fuzzy linear regression., *Journal of Emerging Trends in Computing and Information Sciences*, **3(2)**, 225-230.

Erdogan, S. (2009)., Explorative spatial analysis of traffic accident statistics and road mortality among the provinces of Turkey, *Journal of Safety Research*, **40(5)**, 341-351.

Jiang, X., Abdel-Aty, M. Alamili, S. (2014), Application of Poisson random effect models for highway network screening., *Accident Analysis and Prevention*, **63**, 74-82.

- Jones, B., Janssen, L. Mannering, F. (1991)., Analysis of the frequency and duration of freeway accidents in Seattle., *Accident Analysis and Prevention*, **23(4)**, 239-255.
- Jovanis, P. Chang, H. (1986), Modelling the relationship of accidents to miles traveled., *Transportation Research Record*, **1068**, 42-51.
- Levine, N., Kim, K. E. Nitz, L. H. (1995a), Spatial analysis of Honolulu motor vehicle crashes: I. Spatial patterns., *Accident Analysis and Prevention*, **27(5)**, 663-674.
- Maher, M. J. Mountain, L. J. (1988), The identification of accident blackspots: A comparison of current methods., *Accident Analysis and Prevention*, **20(2)**, 143-151.
- Roque, C. Cardoso, J. L. (2014), Investigating the relationship between run-off the-road crash frequency and traffic flow through different functional forms, *Accident Analysis and Prevention*, **63**, 121-132.
- Ye, X., Pendyala, R. M., Washington, S. P., Konduri, K., Oh, J. (2009), A simultaneous equations model of crash frequency by collision type for rural intersections., *Safety Science*, **47(3)**, 443-452.
- Yu, R. Abdel-Aty, M. (2013), Investigating different approaches to develop information priors in hierarchical Bayesian safety performance functions, *Accident Analysis and Prevention*, **56**, 51-58.
- Yu, R., Abdel-Aty, M., Ahmed, M. (2013), Bayesian random effect models incorporating real-time weather and traffic data to investigate mountainous freeway hazardous factors, *Accident Analysis and Prevention*, **50**, 371-376.