

Estimación de modelos predictivos para accidentes viales en el Eje 5 Oriente de la Ciudad de México

Carmen Gabriela Angulo Payán
Simulación II, ESFM-IPN, Ciudad de México, México
Teléfono: (55) 3136-0729 E-mail: cangulop1500@alumno.ipn.mx

Resumen

La Ciudad de México, una de las metrópolis más transitadas del país, enfrenta desafíos constantes en la gestión del tránsito debido a la interacción diaria entre vehículos de transporte público, emergencias, motocicletas y automóviles particulares. Entre las áreas con mayor incidencia de accidentes se encuentra el Eje 5 Oriente. Para abordar esta problemática, se emplea el método de Bayes Empírico como modelo predictivo, con el fin de analizar cómo diferentes atributos de las intersecciones influyen en la ocurrencia de accidentes. Este análisis incluyó la creación de modelos diferenciados para intersecciones semaforizadas y no semaforizadas, lo que permitió adoptar un enfoque más específico y detallado en la evaluación de riesgos.

Palabras clave: Bayes Empírico, predicción de accidentes, Eje 5 Oriente, intersecciones.

Los accidentes de tránsito representan una de las principales causas de lesiones y muertes a nivel mundial, con cifras alarmantes: aproximadamente 1.19 millones de personas están involucradas en estos eventos cada año, y entre 20 y 50 millones sufren heridas, según datos de la Organización Mundial de la Salud (2023). Este fenómeno adquiere mayor relevancia en entornos metropolitanos, donde la complejidad del análisis se incrementa debido a la interacción de múltiples factores, como las condiciones climáticas, la infraestructura vial y el comportamiento de los conductores.

En este contexto, el Eje 5 Oriente de la Ciudad de México destaca como una de las zonas con mayor incidencia de accidentes de tránsito, según los registros de la Secretaría de Seguridad Ciudadana. Por ello, este estudio se centra en esta vialidad, analizando sus características físicas y factores asociados a los incidentes para identificar condiciones de alto riesgo y anticipar las zonas más vulnerables.

Con este objetivo, se propone un modelo predictivo basado en el método bayesiano empírico, que combina datos históricos y modelos generales para estimar la probabilidad de accidentes en ubicaciones específicas, como intersecciones del Eje 5 Oriente. Este enfoque, además de ser una herramienta clave para la prevención, tiene el potencial de contribuir significativamente a la planificación urbana y al diseño de políticas de seguridad vial más efectivas en la Ciudad de México.

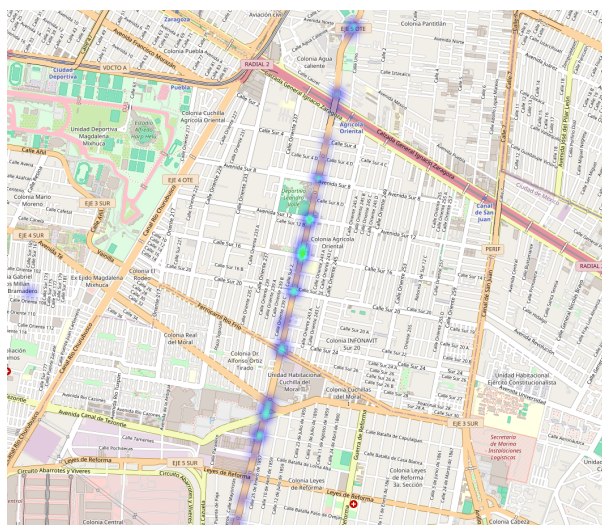


Figura 1: Mapa de calor de incidentes en el Eje 5 Oriente

1. Base de datos

Para desarrollar el modelo, se generó una base de datos que comprende información sobre el Eje 5 Oriente durante el período de 2020 a 2023. Las bases de datos se obtuvieron del Portal de Datos Abiertos de la Ciudad de México, proporcionadas por la Secretaría de Seguridad Ciudadana (SSC), donde se recopilaron registros relacionados con accidentes de tránsito. Esta información se complementó con datos climáticos proporcionados por la Comisión Nacional del Agua (CONAGUA).

La información disponible en la base de datos de la SSC incluye las siguientes variables: folio, fecha del evento, hora del evento, tipo de evento, tipo de vehículo, latitud, longitud, punto 1, punto 2, alcaldía, colonia, tipo de intersección, clasificación de la vialidad, número de personas lesionadas y número de personas fallecidas. Por su parte, los datos de la CONAGUA proporcionaron registros de precipitación pluvial mensual (en mm) por año.

La base de datos original de la SSC contenía aproximadamente 130,000 registros correspondientes a incidentes ocurridos en la Ciudad de México durante el período señalado. Estos datos fueron filtrados para seleccionar únicamente los eventos ocurridos en el Eje 5 Oriente (identificados a través de la variable "punto 1"). Posteriormente, se eligieron las intersecciones con mayor incidencia de accidentes, clasificándolas en tres tipos principales: tipo Y, tipo T y en cruz. El tipo de evento analizado fue choque, omitiendo los registros con valores faltantes (nomenclatura "NA").

La inclusión de la precipitación como variable climática se justifica por estudios del Centro de Experimentación y Seguridad Vial México (CESVI), que estiman que durante la temporada de lluvias los accidentes viales pueden incrementarse entre un 15 % y un 20 %.

En total, se analizaron 418 incidentes, de los cuales 317 correspondieron a intersecciones semaforizadas y 101 a intersecciones no semaforizadas. Las variables utilizadas en el modelo se describen a continuación:

Variable resultado		
Choques		

Variables explicativas		
Grupo	Variable	Nomenclatura
Exposicion	Trafico Promedio Anual Diario	TPDA
Red	Intersección tipo Y	tipo_Y
	Intersección tipo T	tipo_T
	Interseccion tipo cruz	tipo_cruz
Clima	Precipitación	precipitacion

2. Metodología

En este modelo bayesiano, la modelación del fenómeno en contextos interurbanos o rurales permite el uso de diversas formas funcionales. Este enfoque utiliza una distribución binomial negativa para modelar el término de error, incorporando un parámetro de sobredispersión que aborda situaciones en las que la variabilidad de los datos supera el promedio. Sin este parámetro, un modelo Poisson estándar podría subestimar la varianza, comprometiendo la precisión de las estimaciones. Este ajuste es particularmente útil en el análisis de accidentes de tránsito, donde la dispersión de los datos suele ser mayor debido a la influencia de múltiples factores.

El modelo propuesto se aplica al análisis de accidentes en el Eje 5 Oriente, incorporando variables específicas de la vialidad, como tipo de intersección y el volumen de tráfico. Además, combina información histórica específica del lugar con predicciones calculadas por el modelo, lo que refina la estimación final de accidentes.

Se basa en principios ampliamente aceptados en el análisis de seguridad vial, estimando el número esperado de accidentes $E(\lambda)$ en función del volumen de tráfico, las características geométricas de las vialidades y factores externos. Este enfoque sigue la línea de estudios anteriores, como los de Hauer

(1995) e Ivan (2004), quienes emplearon modelos log-lineales para incorporar múltiples factores de riesgo.

El modelo generalizado, tiene la siguiente forma:

$$E(\lambda) = e^{\alpha + \beta \ln(Q) + \sum \gamma_i X_i} \quad (1)$$

Donde:

- $E(\lambda)$: es el número esperado de accidentes
- Q : es el volumen de tráfico
- X_i : son los factores de riesgo adicionales, como características geométricas o climáticas
- α, β, γ_i : son parámetros que se calibran mediante los datos observados

Al aplicar la transformación logarítmica, este modelo se reescribe como:

$$\ln(E(\lambda)) = \alpha + \beta \ln(Q) + \sum_i \gamma_i X_i \quad (2)$$

Este formato log-lineal facilita la interpretación de los coeficientes y la estimación mediante métodos estadísticos como los Modelos Generalizados Lineales (GLMs). En este estudio, se utiliza una variante adaptada para intersecciones urbanas, considerando tanto el volumen de tráfico como factores geométricos y climáticos que influyen en el número de choques registrados. El modelo utilizado es el siguiente:

$$\ln(\text{choques}) = \alpha + \beta \ln(\text{TPDA}) + \gamma_1 \text{tipo_Y} + \gamma_2 \text{tipo_T} + \gamma_3 \text{tipo_cruz} + \gamma_4 \text{precipitacion} \quad (3)$$

3. Resultados

Variable	No semaforizadas		Semaforizadas	
	Coeficiente	P-valor	Coeficiente	P-valor
Intercept	0.3353	0.142	0.3621	1.337
Log(TPDA)	0.9009	1.023	0.8906	0.492
Tipo Y	0.1080	0.114	0.1090	0.710
Tipo T	0.1037	0.124	0.1215	0.547
Tipo cruz	0.1236	0.147	0.1316	0.495
Precipitación	0.0245	0.225	0.0220	0.069
Número de observaciones	101		317	
pseudo R2	0.01374		0.01371	

4. Resultados

Aquí puedes incluir los resultados de tu trabajo.

5. Conclusiones

Finalmente, puedes añadir tus conclusiones.