

Estimación de Modelos Predictivos para Accidentes Viales en el Eje 5 Oriente

Escuela superior de Física y Matemáticas

Artículo

Estudiante: Angulo Payán Carmen Gabriela

Fecha: Noviembre 20, 2024

Resumen: La Ciudad de México, una de las metrópolis más transitadas del país, enfrenta desafíos constantes en la gestión de su tránsito debido a la interacción diaria entre vehículos de transporte público, de emergencia, motocicletas y automóviles particulares.

Entre las zonas con mayor incidencia de accidentes destaca el Eje 5 Oriente, es por ello que para abordar esta problemática, se utiliza el método de Bayes Empírico como modelo predictivo, con el objetivo de analizar cómo distintos 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 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 son una de las principales causas de lesiones y muertes a nivel mundial, con una incidencia alarmante: aproximadamente 1,19 millones de personas están involucradas en estos eventos cada año, y entre 20 y 50 millones sufren heridas (OMS, 2023). En el contexto metropolitano, aunque existen datos sobre el tipo y frecuencia de estos accidentes, su análisis es complejo debido a la influencia de múltiples factores, como las condiciones climáticas, la infraestructura vial y el comportamiento de los conductores.

A pesar de estos desafíos, se propone el desarrollo de un modelo predictivo que, a partir de las características físicas de las vialidades y los factores climatológicos, permita identificar condiciones de alto riesgo y anticipar zonas con mayor incidencia de accidentes. Este modelo emplea el método bayesiano empírico, una herramienta eficaz para estimar la probabilidad de accidentes en ubicaciones específicas, como intersecciones o tramos de carretera, considerando tanto modelos generales como datos históricos.

En el caso de la Ciudad de México, este enfoque

resulta particularmente relevante debido al registro sistemático de accidentes realizados por la Secretaría de Seguridad Ciudadana en sus diferentes alcaldías. Este análisis no solo podría contribuir a reducir la ocurrencia de accidentes, sino también a mejorar la planificación urbana y las políticas de seguridad vial.

1 Base de datos

Para desarrollar el modelo, se generó una base de datos que comprende información sobre el Eje 5 Oriente durante el periodo de 2020 a 2023.

Las bases de datos se obtuvieron del Portal de Datos Abiertos de la CDMX, por parte de la Secretaría de Seguridad Ciudadana (SSC) se obtuvieron los referentes con los accidentes de tránsito y se complementaron con información climática disponible por la Comisión Nacional de Agua (CONAGUA). En total, se analizaron 418 incidentes, 317 para intersecciones semaforizadas y 101 para intersecciones no semaforizadas.

La información disponible en la base de datos del SSC son: folio, fecha de evento, hora de evento, tipo de evento, tipo de vehículo, latitud, longitud, punto 1, punto 2, alcaldía, colonia, tipo de intersección, clasificación de la vialidad, personas le-

sionadas y personas fallecidas. Y en el caso de la CONAGUA la precipitación pluvial en mm mensual por año.

Las variables utilizadas en el modelo se muestran en la Tabla 1

Table 1: Variables utilizadas en el modelo

Variables		Nomenclatura
Respuesta:	Choques	choques
Explicativas:		
Exposición	Tráfico Promedio Diario Anual	TPDA
Red	Intersecciones tipo Y	Y
	Intersecciones tipo T	T
	Intersecciones tipo cruz	cruz
Clima	Precipitación pluvial en mm	precipitación

2 Metodología

El modelo generalizado tiene la siguiente forma:

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

En este modelo bayesiano, la modelación del fenómeno para contextos interurbanos o rurales permite una variedad de 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. Este ajuste permite manejar discrepancias entre las predicciones del modelo y los valores observados, mejorando así la precisión del análisis.

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 y α, β, γ_i son parámetros que se calibran mediante los datos observados. Al aplicar la transformación logarítmica, este modelo se reescribe como:

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

El modelo combina información histórica específica del lugar con proporciones de resultados modelados, lo que refina la predicción 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.

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:

References

$$\ln(\text{choques}) = \beta_0 + \beta_1 \ln(\text{TPDA}) + \beta_2 \text{int}Y \\ + \beta_3 \text{int}Cruz + \beta_4 \text{int}T + \beta_5 \text{precipitacion}$$

Donde $\ln(\text{choques})$ representa el logaritmo del número de accidentes esperados en una intersección, TPDA es el Tránsito Promedio Diario Anual, $\text{int}Y$, $\text{int}Cruz$, $\text{int}T$ es el número de intersecciones respectivamente, precipitación es la variable climática que representa la precipitación media, β_0 el término independiente y β_i para $n = 1, 2, 3, 4, 5$ los coeficientes a estimar.

Como se menciono anteriormente en la Tabla 1 las variables explicativas fueron las siguientes:

- **log(TPDA):** logaritmo del TPDA
- **Y:** vale 1 es la intersección tiene 3 accesos y 0 en caso contrario
- **T:** vale 1 es la intersección tiene 3 accesos y 0 en caso contrario
- **cruz:** : vale 1 es la intersección tiene 4 accesos y 0 en caso contrario

3 Conclusiones

- [1] Elvik, R. (2007) State-of-the-art approaches to road accident black spot management and safety analysis of road networks. TI Report 883, 2007. The Institute of Transport Economics. Oslo, Noruega.
- [2] Queralt, N., Taratiel, D., de Graaf, C., Caballol, R., Cimiraglia, R., & Angeli, C. (2008). On the applicability of multireference second-order perturbation theory to study weak magnetic coupling in molecular complexes. *Journal of computational chemistry*, 29(6), 994-1003.
- [3] Roca-Sanjuán, D., Aquilante, F., & Lindh, R. (2012). Multiconfiguration second-order perturbation theory approach to strong electron correlation in chemistry and photochemistry. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 2(4), 585-603.
- [4] Jiménez Rosas. C.A.S (2022). Estudio topológico del flujo de corriente eléctrica en compuestos tipo ML_4 [Tesis de licenciatura, Universidad Nacional Autónoma de México]. TESIUNAM.