

---

# Análisis probabilístico de la movilidad urbana en la ZMVM a partir de la Encuesta Origen–Destino 2017 mediante redes bayesianas multinomiales

---

Joshua A. Chaidez Ochoa<sup>1</sup>, Luis C. Marrufo Padilla<sup>1</sup>, Ángel Esparza Enríquez<sup>1</sup> and Gustavo A. Aguilar Torreblanca<sup>1</sup>

<sup>1</sup> Tecnológico de Monterrey, Escuela de Ingeniería y Ciencias, Guadalajara, Jalisco

---

**Abstract**—En general las ciudades enfrentan un mismo problema, la movilidad, y la ZMVM no es una excepción. Esta zona presenta múltiples problemas en infraestructura y planeación urbana, que se traducen en problemas como tiempos de viaje prolongados y un transporte público ineficiente. En este estudio implementamos el uso de redes bayesianas multinomiales alimentadas con datos del INEGI para modelar la relación entre las variables ligadas al transporte y los tiempos de viaje, a través de grafos acíclicos dirigidos tanto planteados como contruidos mediante un algoritmo de hill-climbing. Como resultado, una red bayesiana que describe las relaciones entre las variables del viaje y a la cual es posible hacerle consultas de interés como la probabilidad de que el tiempo de viaje dependiendo factores como el medio de transporte o el estrato sociodemográfico de las personas. Estos resultados son de utilidad al momento de implementar políticas públicas o tomar decisiones sobre la planeación urbana a futuro.

**Keywords**—Redes bayesianas, Análisis multinomial, Modelos probabilísticos, DAG, Inferencia causal, Movilidad urbana, ZMVM, Encuesta Origen-Destino 2017, INEGI

---

## I. INTRODUCCIÓN

La movilidad urbana en la Zona Metropolitana del Valle de México (ZMVM) presenta ser un desafío importante, debido al uso intensivo del automóvil y a la inadecuada planificación urbana que ocasionan prolongados tiempos de traslado [1]. Para profundizar en este fenómeno, la Encuesta Origen–Destino en Hogares de la Zona Metropolitana del Valle de México 2017 (EOD-2017) del INEGI proporciona información relevante sobre motivos de viaje, modos de transporte y características socioeconómicas de los hogares.

Este estudio aplica redes bayesianas multinomiales como herramienta para modelar dependencias probabilísticas entre variables diferentes de transporte y estimar probabilidades en escenarios específicos. Se utilizarán grafos acíclicos no dirigidos (DAG) para responder preguntas clave relacionadas con los tiempos de viaje, motivos de viaje, medios de transporte, origen y destino del viaje y el estrato socioeconómico. Este trabajo permite cuantificar los patrones de movilidad de la Zona Metropolitana del Valle de México (ZMVM), lo que puede resultar en una mejor planeación de políticas e infraestructura de transporte.

## II. METODOLOGIA

El análisis se realizó partiendo los datos de la EOD-2017 [2]. Tras una exploración de la base de datos se seleccionaron las variables necesarias para contestar las siguientes cuatro preguntas objetivo:

- ¿Cuál es la probabilidad de que el tiempo de viaje sea menor a 20 minutos, dado que su medio de transporte fue el automóvil?
- ¿Cuál es la probabilidad de que los viajes en Ciudad de México se hagan en medios no motorizados (caminar o bicicleta)?
- ¿Cuál es la probabilidad de que un hogar de estrato medio bajo, utilicen más de 2 medios de transporte diferentes?
- ¿Cuál es la probabilidad de que el viaje del hogar a la escuela dure más de 60 minutos si la persona viaja en coche?

### a. Construcción de las DAGs

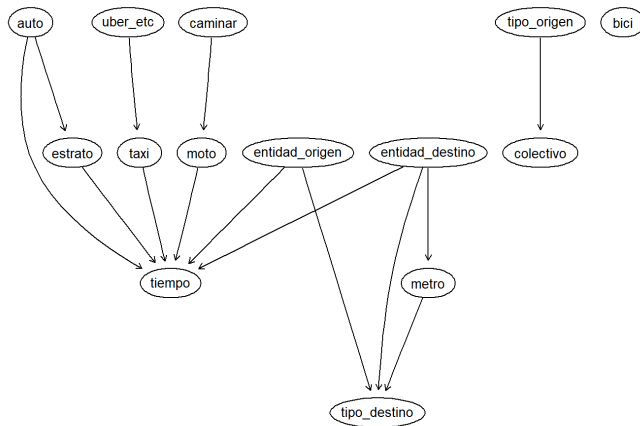
La página de la EOD-2017 contiene múltiples bases de datos para la ZMVM, pero a lo largo de este estudio se utilizará la base de datos referente a viajes realizados, base que contiene todas las variables necesarias para contestar nuestras preguntas objetivo. Antes de generar la red, se limpió la base de datos, eliminando las variables no relevantes para nuestras consultas y los ID's de cada registro, rellorando información faltante con 0 y transformando los tipos de datos para que fueran manejables. Se construyó una nueva variable para el tiempo de viaje, ya que en la base de datos original solo contiene las horas de inicio y fin; finalmente, se discretizó la variable construida de tiempo para que fuera más manejable para `bnlearn`.

Después de preparar la base de datos, planteamos tres posibles DAGs, utilizando la biblioteca `bnlearn` en R para generarlas. También optamos por generar una red adicional mediante el algoritmo hill climbing para contar con más posibilidades. Comparamos resultados BIC para encontrar la mejor red generada en base a ese criterio.

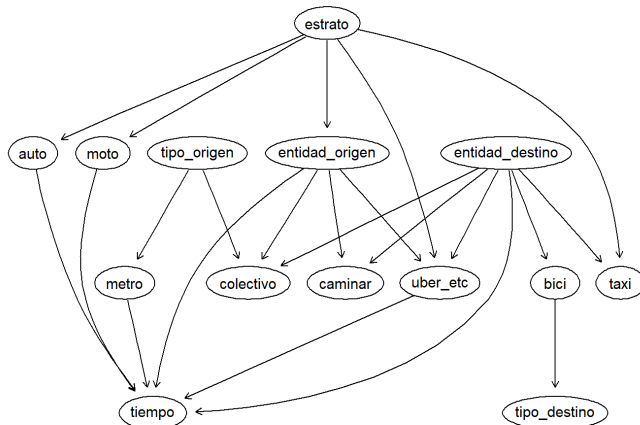
Una vez generada la red, se empleará la función de `cpquery` de la biblioteca de `bnlearn` para estimar la probabilidad de cada consulta, utilizando un número de simulaciones  $n = 10^6$ . Con el fin de reproducibilidad de este estudio, también se empleará una semilla fija `set.seed(1809)`.

### b. Selección de la DAG

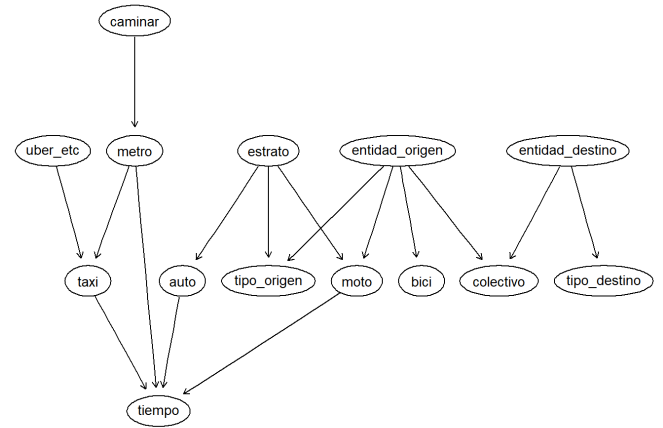
A continuación se muestran las tres estructuras propuestas para este estudio. Se estarán comparando entre ellas y con una generada empleando el algoritmo de hill-climbing.



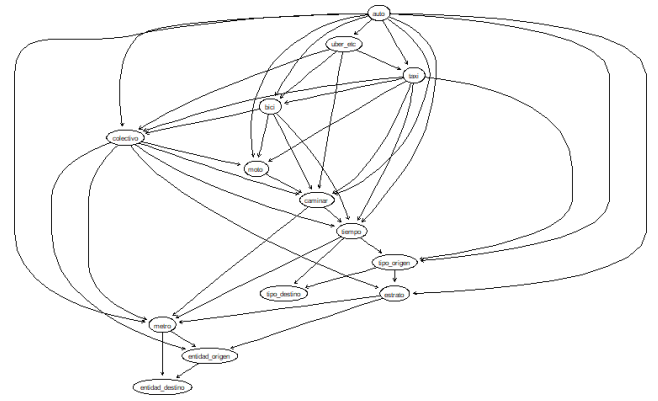
**Fig. 1:** Estructura planteada para la DAG "A".



**Fig. 2:** Estructura planteada para la DAG "B".



**Fig. 3:** Estructura planteada para la DAG "C".



**Fig. 4:** Estructura DAG generada utilizando un algoritmo hill-climbing.

DAG	Puntaje BIC	Puntaje AIC
A	-6916796	-5722651
B	-6223229	-5602247
C	-5544626	-5528259
HC	-4653180	-4620061

**Fig. 5:** Puntajes BIC y AIC para cada estructura DAG planteada.

Al final seleccionamos la DAG que maximiza el puntaje BIC, aquella realizada utilizando hill-climbing, con la cual brindamos solución a las consultas ya mencionadas.

## III. RESULTADOS

Al hacer las consultas en la red correspondiente, obtenemos los siguientes resultados:

- La probabilidad estimada de que el tiempo de viaje sea menor a 20 minutos dado que el medio de transporte fue un automóvil es igual a  $P = 0.3195917$ .
- La probabilidad estimada de que un viaje en la Ciudad de México sea caminando o en bicicleta es de  $P = 0.6171179$ .



- La probabilidad de que un hogar de estrato medio bajo utilice más de dos medios de transporte diferentes es de  $P = 0.07379346$ .
- La probabilidad de que un viaje del hogar a la escuela dure más de 60 minutos dado que la persona viaja en auto es de  $P = 0.05340155$ .

#### IV. CONCLUSIONES

Las redes bayesianas multinomiales son útiles para modelar, entender y evaluar la relación entre los diferentes factores de diversas situaciones, en el caso de este proyecto ofrecen una manera de entender cómo los tiempos de viaje, las decisiones de transporte y las circunstancias de la población se relacionan entre sí. La DAG resultante permite obtener de manera directa consultas a el estado de la vialidad de la zona como los tiempos de trayecto, el uso de diferentes medios de transporte, los destinos de la población y en general darnos acceso a una imagen completa de como y porque se transportan las personas dentro de la Zona Metropolitana del Valle de México (ZMVM).

En conjunto estos datos vistos a través de la red proporcionan una herramienta enorme al momento de tomar decisiones en temas de planeación urbana ya que nos permiten explorar diferentes escenarios y saber cómo estos pueden afectar otra parte completamente diferente del sistema de transporte. Por ejemplo, podemos saber si es plausible hacer cambios para mejorar el transporte público en ciertas zonas de la ciudad donde el trayecto ya es muy tardado, o mejorar vías específicas como el viaje por motocicleta o bici.

Este estudio no identifica directamente la causalidad y la información que provee debe tomarse con precaución y estar acompañado de un análisis más extenso en el tema de urbanismo, además no estamos tomando en cuenta varios aspectos como las calles transitadas o la temporalidad de los datos. Estas son algunas de las consideraciones que se tienen que tener al replicar este estudio o aplicarlo en casos de vida real.

#### REFERENCES

- [1] Comisión Ambiental de la Megalópolis. (2018) Afrontando el reto de la movilidad sustentable en la zmv. [Online]. Available: <https://www.gob.mx/comisionambiental/articulos/afrontando-el-reto-de-la-movilidad-sustentable-en-la-zmv> ? idiom = es
- [2] Instituto Nacional de Estadística y Geografía (INEGI), “Encuesta Origen–Destino en Hogares de la Zona Metropolitana del Valle de México 2017,” <https://www.inegi.org.mx/programas/eod/2017/>, 2017, consultado el 28 de agosto de 2025.