

Estimación de actividad vehicular a partir de registros de odómetro en centros de diagnóstico automotor

Diego Arias

Universidad del Valle, Escuela de Estadística



Marzo 22, 2023

Estimación de actividad vehicular a partir de registros de odómetro en centros de diagnóstico automotor



1 Contexto

2 Metodología

3 Resultados

Estimación de actividad vehicular a partir de registros de odómetro en centros de diagnóstico automotor



1 Contexto

2 Metodología

3 Resultados



#aunachonlu

CENTRAL VILLAGE

บขส.14

204

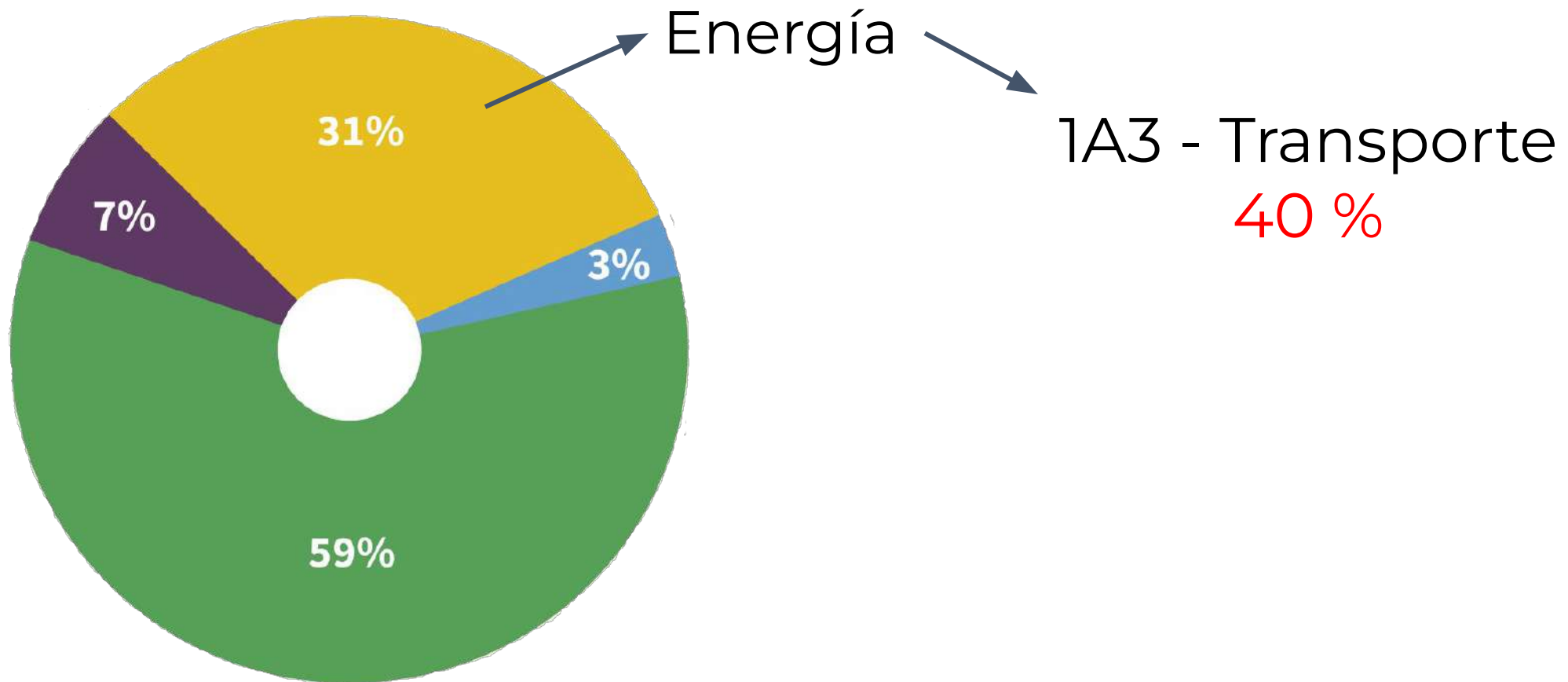
730

บขส.14

COCA-COLA

TAXI

El **transporte**, principal emisor de gases de efecto invernadero (GEI) del sector Energía - INGEI 2018



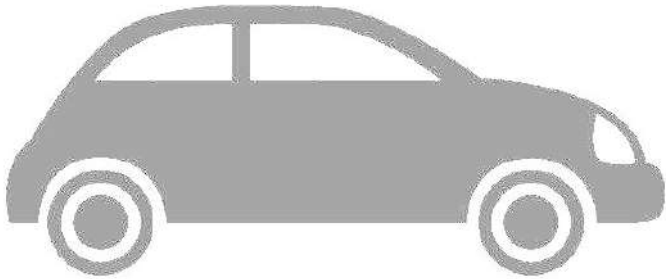
¿Cómo se cuantifican las emisiones
de GEI?

Emission \propto *Actividad*



Por medio del
consumo de
combustibles o ...

... con los **kilómetros recorridos**
(VKT) de los vehículos



$$VKI_t = v_t \times t$$

Los registros de odómetro de los CDA, un indicador de actividad vehicular



Estimación de actividad vehicular a partir de registros de odómetro en centros de diagnóstico automotor



1 Contexto

2 Metodología

3 Resultados

Se analizó la información de 35 CDAs de Palmira, Yumbo y Cali (~ 400.000 registros)

1

Limpieza



Dominios de las variables **categóricas** (placas, tipo de combustible y vehículos), **numéricas** (modelo y kilometraje) y de tipo **fecha**.

2

Validación



Registros excesivos de placas por CDA.

Preservación del registro de kilometraje más antiguo (duplicidad).

Diferencia entre kilometrajes de revisiones consecutivas > 0 .

Definimos la actividad vehicular por hora teniendo en cuenta dos escenarios

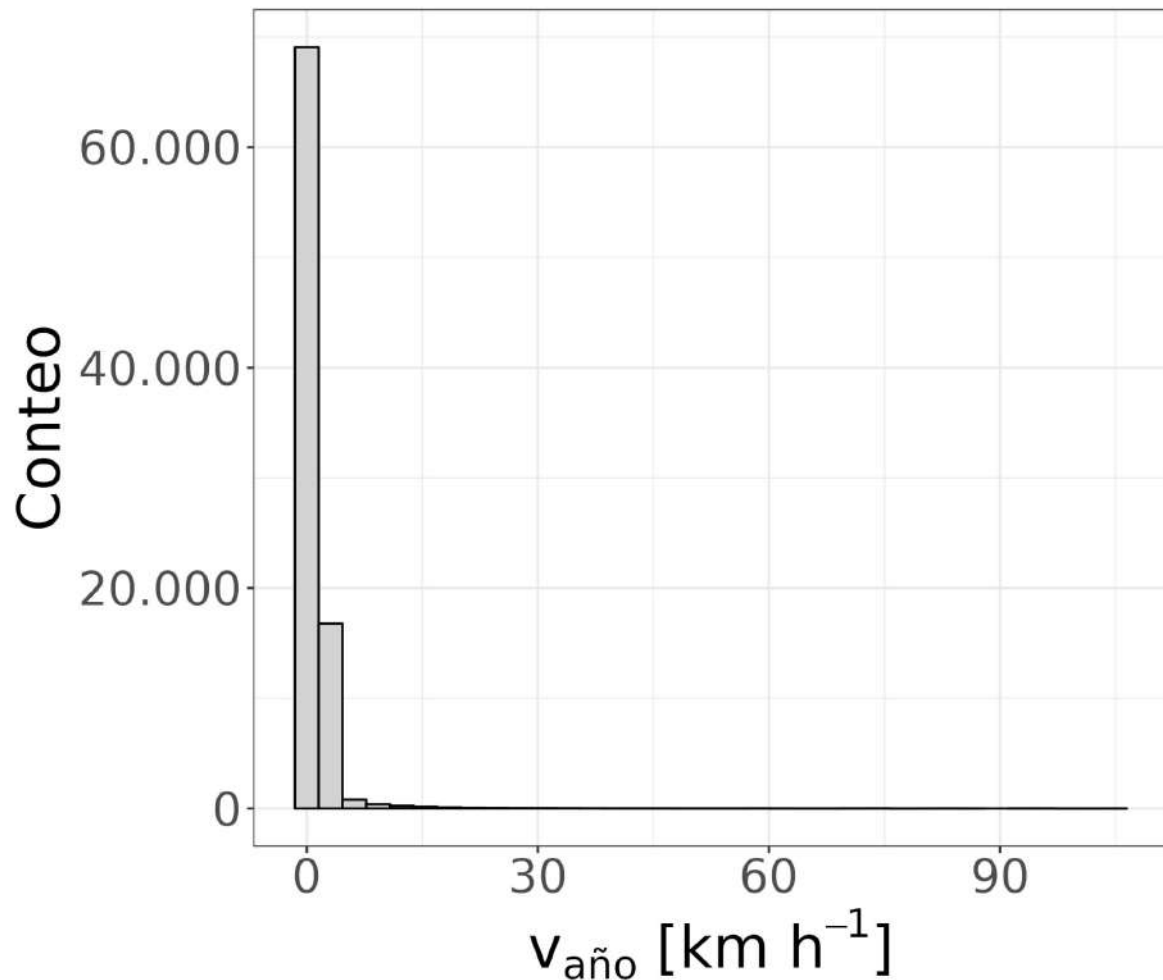
1 Múltiples registros

$$\longrightarrow v_{\text{año}} = \frac{km_{\text{año}} - km_{\text{año}-1}}{(fecha\ revisión_{\text{año}} - fecha\ revisión_{\text{año}-1})_{\text{horas}}} \left[\frac{\text{km}}{\text{h}} \right]$$

2 Único registro

$$\longrightarrow v_{\text{año}} = \frac{km_{\text{año}}}{(fecha\ revisión_{\text{año}} - modelo)_{\text{horas}}} \left[\frac{\text{km}}{\text{h}} \right]$$

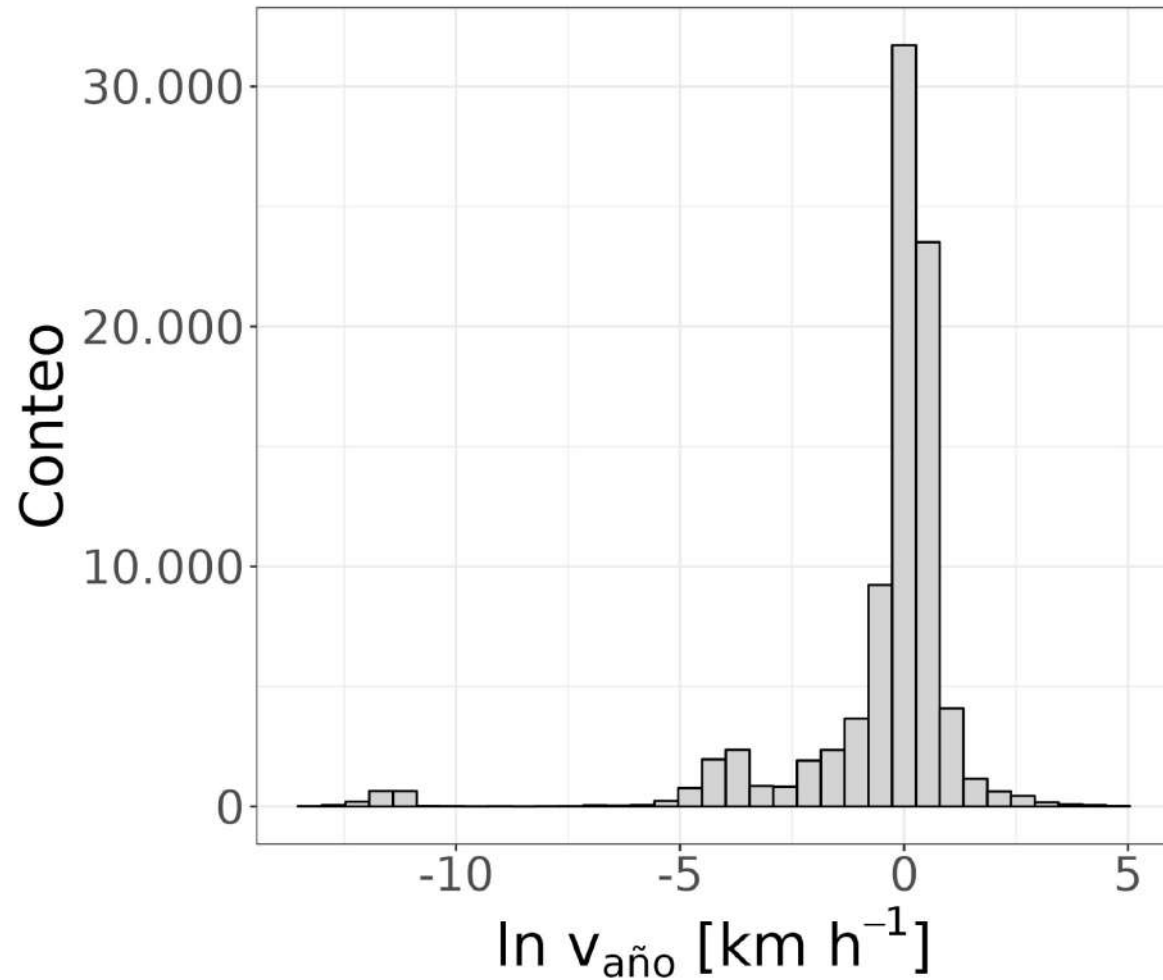
Distribución de la actividad por hora para vehículos particulares en escala **natural**



Asimetría ~ 10

Curtois ~ 150

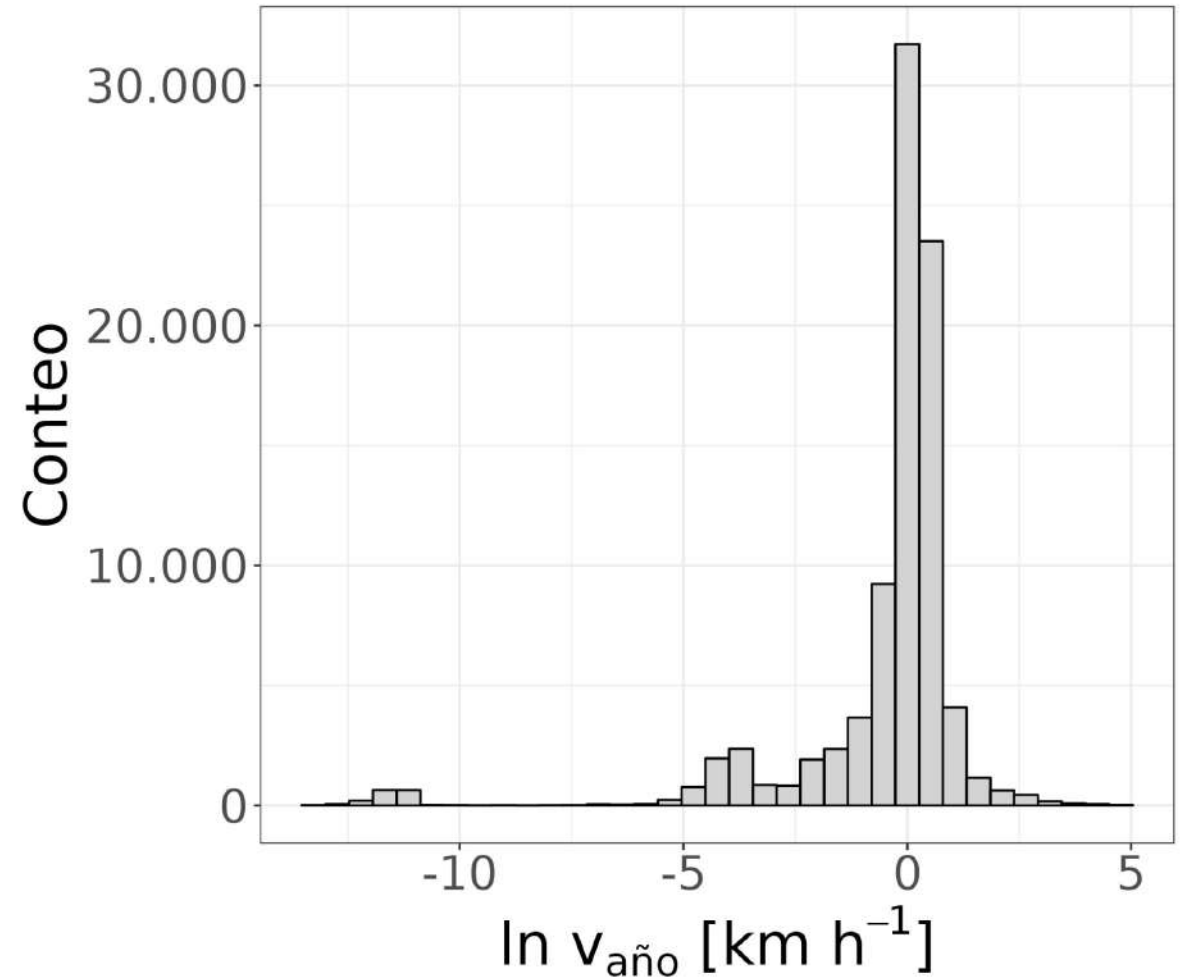
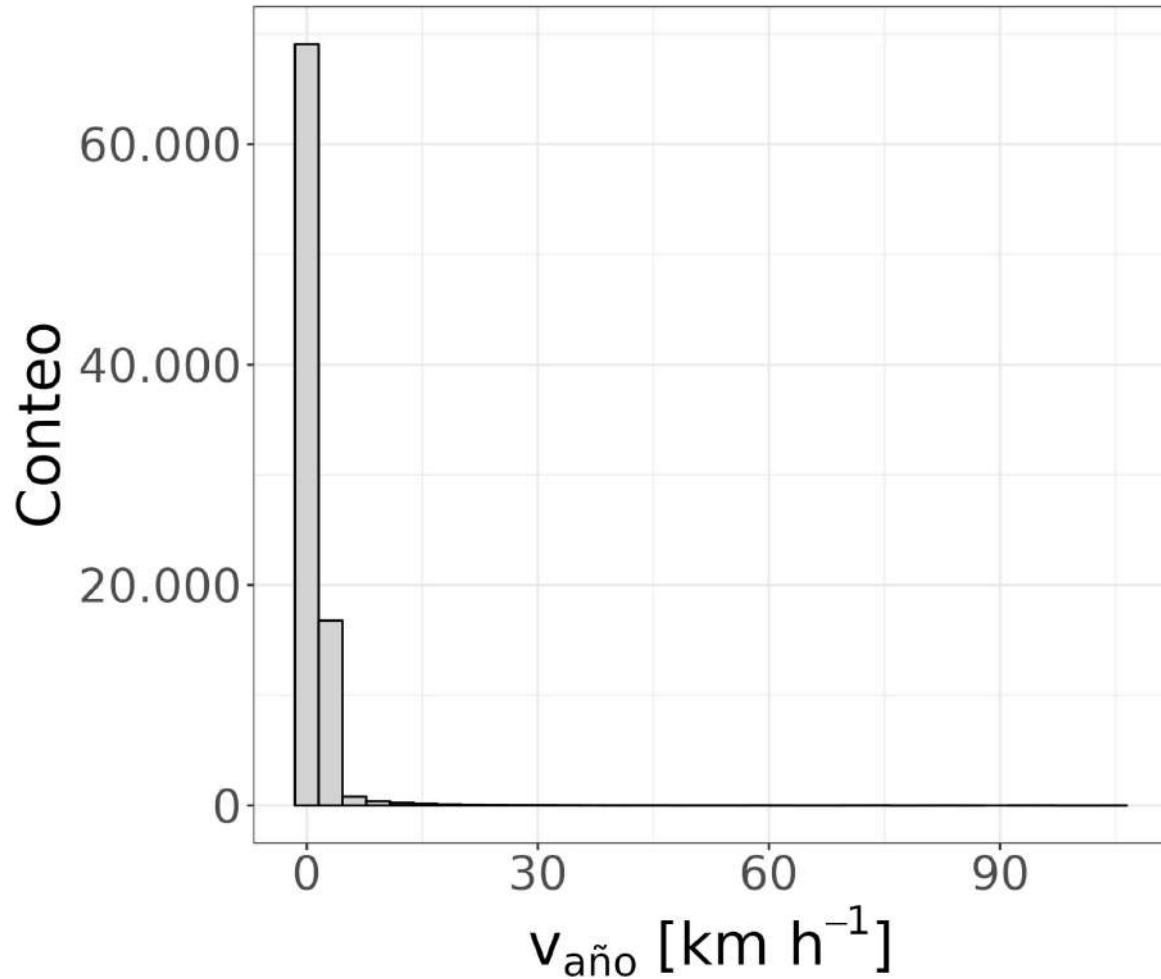
Distribución de la actividad por hora para vehículos particulares en escala **logarítmica**



Asimetría ~ -3

Curtosis ~ 15

La transformación logarítmica disminuye la asimetría y curtosis de la distribución



¿Cuál es la mejor estimación de
la **actividad por hora**
representativa para cada
tipología vehicular?

1 Modelar la distribución de probabilidad

$$f(v_t | \theta_1, \theta_2, \dots; w_1, w_2, \dots) = \sum_{k=1}^{\infty} w_k \phi(v_t | \theta_k)$$

Mezcla infinita de densidades Gaussianas (MIDG)

2 Calcular el valor esperado ponderado

$$\mu^P = \sum_{k=1}^K w_k \mu_k$$

Estimación de actividad vehicular a partir de registros de odómetro en centros de diagnóstico automotor

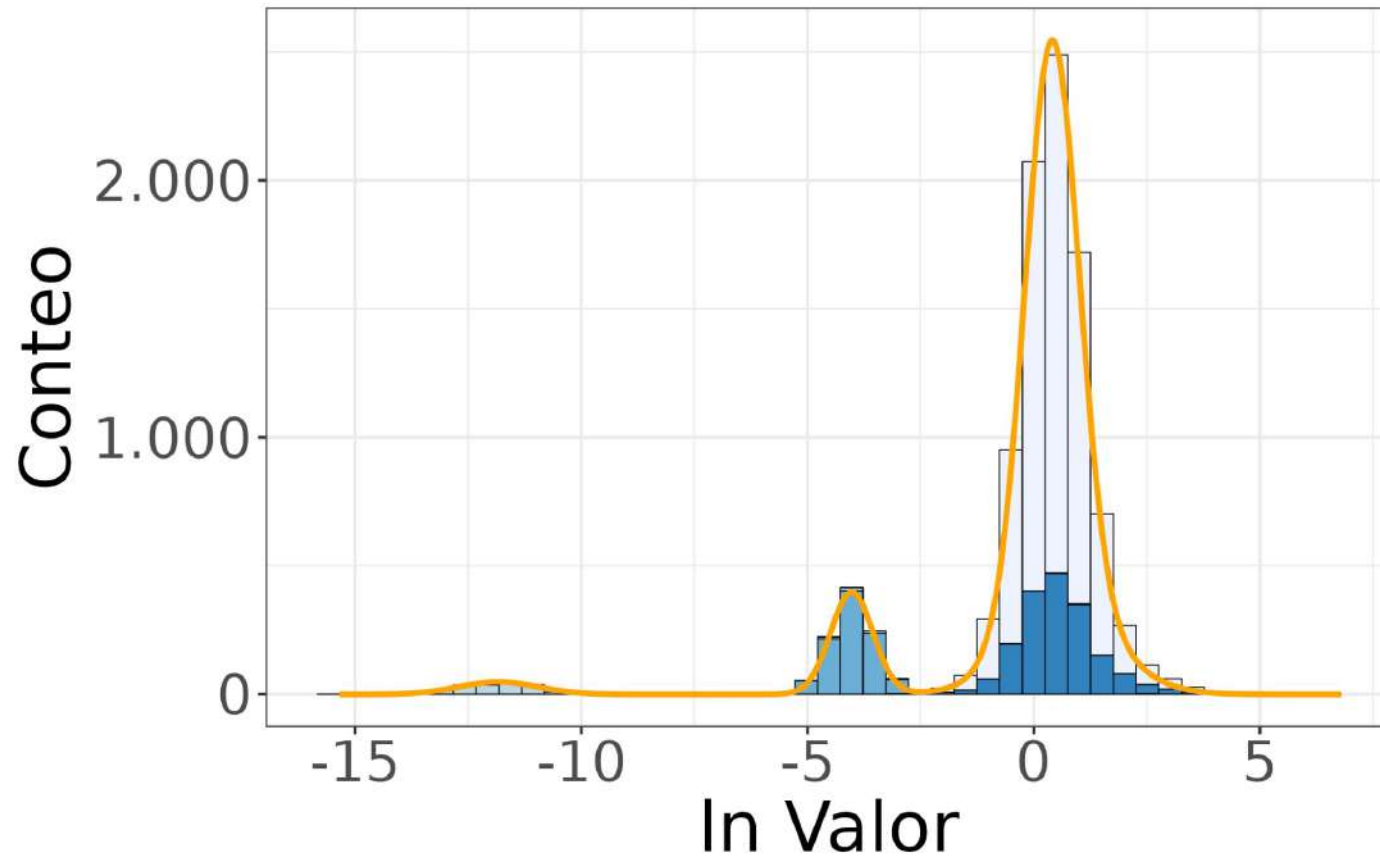


1 Contexto

2 Metodología

3 Resultados

Simulación: El modelo MIDG es capaz de detectar el número de componentes y parámetros en mezclas de distribuciones

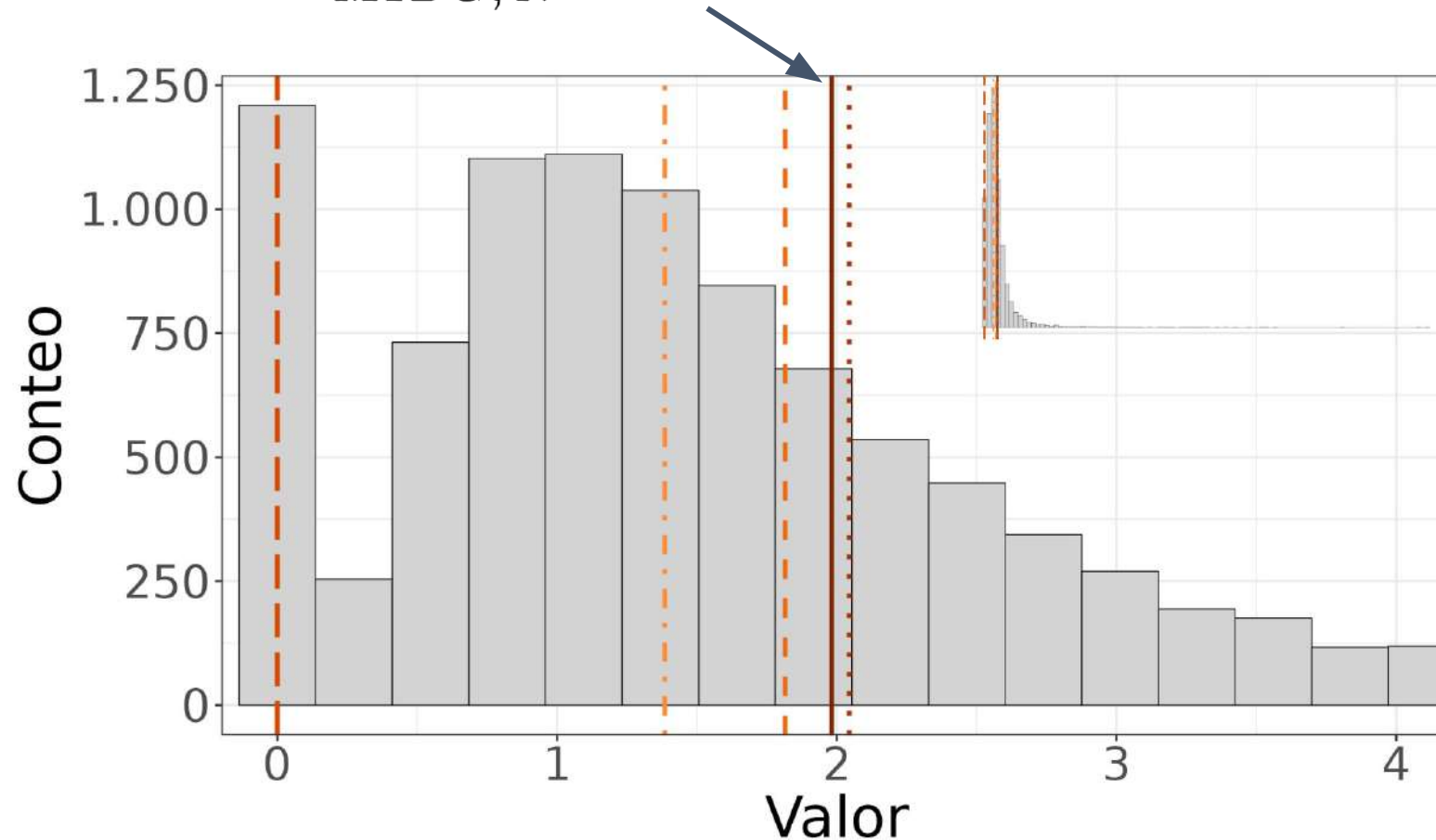


— Ajuste por MIDG □ 1 □ 2 □ 3 □ 4

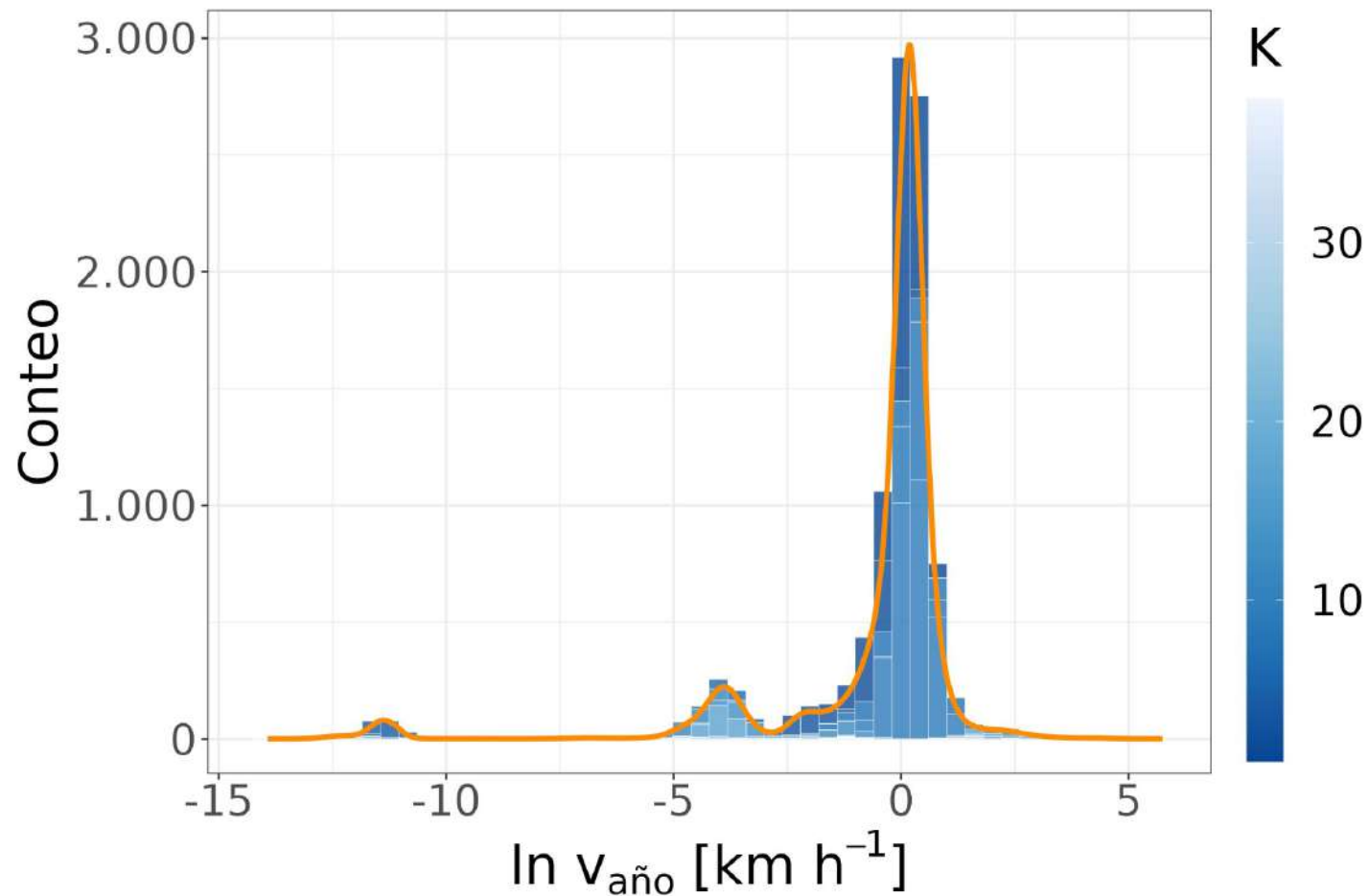
$$\mu_{MIDG, N}^P = \sum_{k=1}^4 w_k e^{\bar{\mu}_k + \frac{1}{2} \bar{\sigma}_k^2}$$

El valor verdadero se encuentra **acotado** por el valor esperado por el modelo MIDG y la media aritmética

$$\text{moda} < M < \mu_{MIDG, N}^P < \mu < \bar{X}$$



Caso de estudio: Para Cali, los kilómetros recorridos por un vehículo particular son ~ 11.400 km



$$\mu_{MIDG, N}^P = 1,2 \left[\frac{\text{km}}{\text{h}} \right]$$

$$\bar{X} = 1,34 \left[\frac{\text{km}}{\text{h}} \right]$$

Conclusiones

- 1** Los **registros de odómetro** de los CDAs constituyen una valiosa fuente de información para estimar la **actividad vehicular**.
- 2** La aplicación de la transformación logarítmica **suaviza** el comportamiento estadístico de la distribución de la actividad horaria.
- 3** El valor real de la actividad horaria podría estar **acotado** por la media estimada por MIDG y la media aritmética.

Agradecimientos

Adriana Pereira¹, Ph.D.

Cesar Ojeda¹, Ph.D.

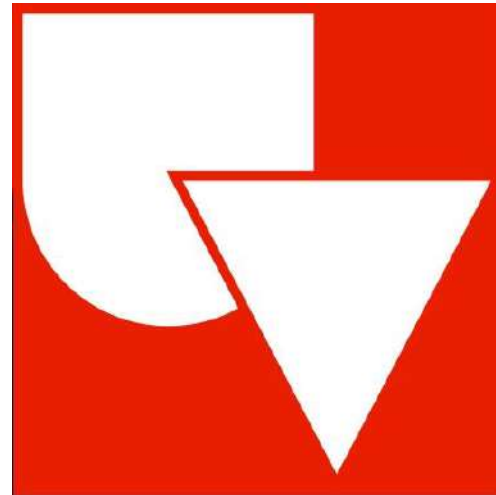
Rodrigo Jiménez², Ph.D.

Sonia Mangones², Ph.D.

Germán Rueda², Ph.D.

¹Universidad del Valle

²Universidad Nacional de Colombia



**Universidad
del Valle**



UNIVERSIDAD
NACIONAL
DE COLOMBIA

Referencias

Alejandro Jara (2007). Applied Bayesian Non- and Semi-parametric Inference Using DPpackage. R News 7(3), 17-26. URL <https://CRAN.R-project.org/doc/Rnews/>

Alejandro Jara, Timothy Hanson, Fernando Quintana, Peter Mueller, Gary Rosner (2011). DPpackage: Bayesian Semi- and Nonparametric Modeling in R. Journal of Statistical Software, 40(5), 1-30. URL <http://www.jstatsoft.org/v40/i05/>.

Beautiful Free Images & Pictures | Unsplash. Unsplash. <https://unsplash.com/>

Christensen, R., Johnson, W., Branscum, A., & Hanson, T.E. (2010). Bayesian Ideas and Data Analysis: An Introduction for Scientists and Statisticians (1st ed.). CRC Press. <https://doi.org/10.1201/9781439894798>

Freepik | Graphic Resources for everyone. <https://www.freepik.com/>

Frühwirth-Schnatter, S. (2006). Finite mixture and Markov switching models. Berlin: Springer.

Moszoro M. & Soto M. (2022). Road Quality and Mean Speed Score. IMF Working Papers 2022/095 (Washington: International Monetary Fund).

MacEachern, S. N. and Muller, P. (1998) Estimating mixture of Dirichlet Process Models. Journal of Computational and Graphical Statistics, 7 (2): 223-338.

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.



Muchas gracias

Más información



<https://casap.science/>



casap@casap.science