



# Modelos estructurales marginales para el control de sesgos en estudios observacionales con factores de riesgo y exposición tiempo-dependientes

Código en: 

Seminario Métodos de Investigación en Salud Pública

11 de julio, 2022

Andrés González Santa Cruz

[gonzalez.santacruz.andres@gmail.com](mailto:gonzalez.santacruz.andres@gmail.com) 



# Índice

1. Problemática
2. Causalidad
3. Ajuste confusión
4. Marginal Structural Models
5. Aplicación
6. Desafíos

# Problemática

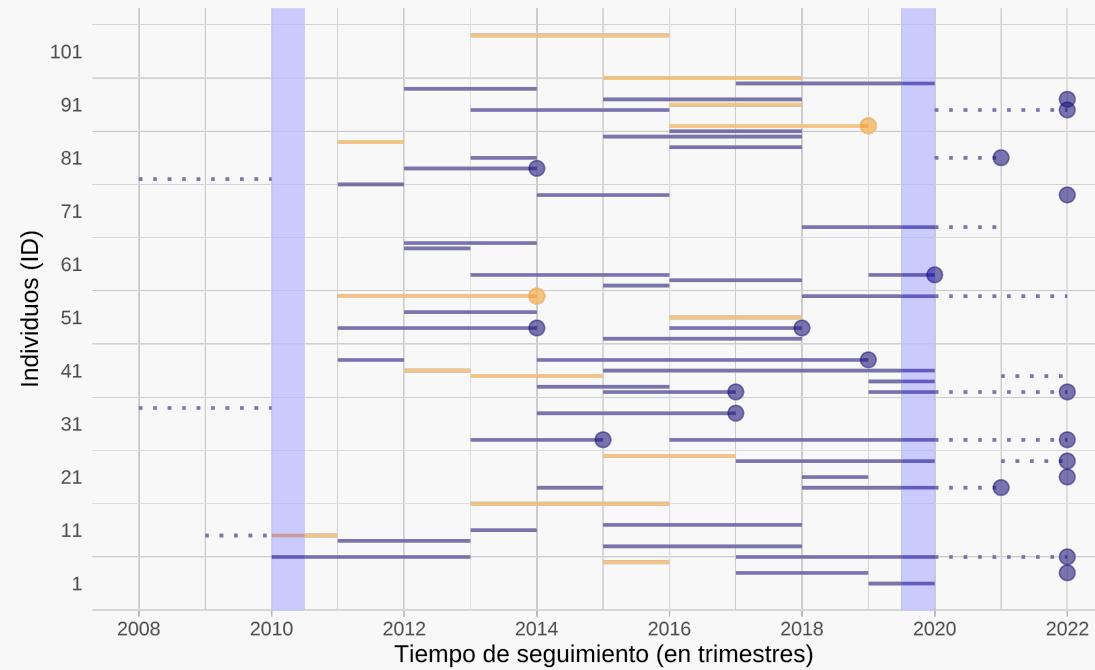
Estructura hipotética más cercana  
a lo real

Primera aproximación

Diagrama Causal Simplificado

Entonces...

► código



Punto= Completa tratamiento;  
Azul= Tratamiento ambulatorio;  
Amarillo= Tratamiento residencial;  
Línea= Tiempo en tratamiento;  
Área sombreada=Inicio y término seguimiento

# Objetivo

Este proyecto apunta a servir como un material introductorio sobre modelos estructurales marginales para el control de sesgos en estudios observacionales con factores de riesgo y/o exposición tiempo-dependientes.

- Se mostrará su estimación paso a paso y mediante un ejercicio reproducible.
- Se utilizará una base de datos simulada para ilustrar el proceso.

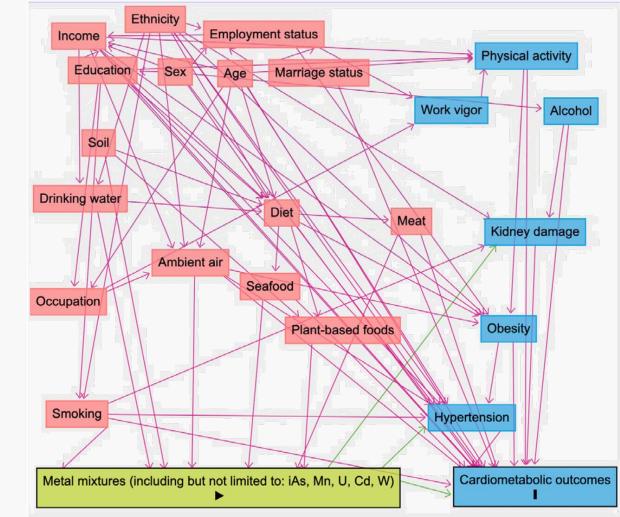
# Causalidad (1)

- Salud pública- epidemiología [3]
- Qué pasaría si... // Qué causa ... [4;5]



- Asociación vs. Causalidad [4;6]
- Inferencia [6]
- RCTs [7]
- Estudios observacionales

*"It's much easier to get a result than it is to get an answer"* [8]



Riseberg,E., Melamed, R., James, K., Alderete, T. & Corlin, L. (2021). Development and application of an evidence-based directed acyclic graph to evaluate the associations between metal mixtures and cardiometabolic outcomes. doi: <https://doi.org/10.1101/2021.03.05.21252993>

# Causalidad (2)

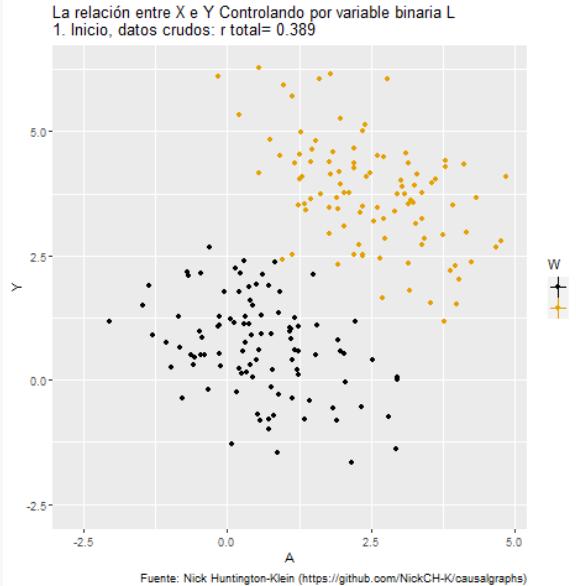
**SUTVA** [9;10;11]



- *Consistencia*
- *Positividad*
- *Intercambiabilidad*

# Ajuste confusión

- **Estratificación / Regresiones**
- **Estandarización** [12;5;13]



# Marginal structural models

- ¿Por qué el nombre?
- Tres principales:
  - IPTW
  - Fórmula G paramétrica
  - Estimación G
- Ventajas y desventajas:

# Aplicación (1)

Se generó una base de datos de 1.000 observaciones con las siguientes características [14] :

- $t_1 = \text{Tratamiento residencial Tiempo } 0 (A_0) \sim \text{Bernoulli}(p=.23)$
- $t_2 = \text{Tratamiento residencial Tiempo } 1 (A_1) \sim \text{Bernoulli}(p=.22)$
- $t_3 = \text{Tratamiento residencial Tiempo } 2 (A_2) \sim \text{Bernoulli}(p=.15)$
- $c_1 = \text{Policonsumo} \sim \text{Bernoulli}(p=.32)$
- $c_2 = \text{Edad de inicio consumo de sustancias} \sim N(3,0.5)$
- $dt_1 = \text{Meses sin readmisión Primer tto.} \sim \text{Poisson}(\lambda=4)$
- $dt_2 = \text{Meses sin readmisión Segundo tto.} \sim \text{Poisson}(\lambda=3)$
- $dt_3 = \text{Meses sin readmisión Tercer tto.} \sim \text{Poisson}(\lambda=2)$
- $v_1 = \text{Confusor tiempo-dependiente (t1)} \sim 5 + 0.4\beta_{T\text{to residencial basal}} + 0.78\beta_{\text{Meses sin readmision 1}} + N(0, \sqrt{0.99})$
- $v_2 = \text{Confusor tiempo-dependiente (t2)} \sim 5 + 0.4\beta_{T\text{to residencial (2}^{\text{do}}\text{)}} + 0.78\beta_{\text{Meses sin readmision 1}} + N(0, \sqrt{0.55})$
- $v_3 = \text{Confusor tiempo-dependiente (t3)} \sim 5 + 0.4\beta_{T\text{to residencial (3}^{\text{er}}\text{)}} + 0.78\beta_{\text{Meses sin readmision 1}} + N(0, \sqrt{0.33})$

## ► código

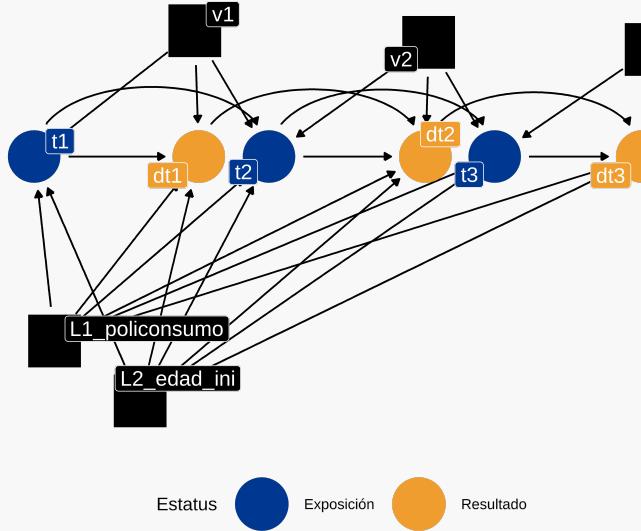
Muestra de la base de datos (Primeras 4 filas)

id	time	policonsumo	edad_ini	t	v	dt
1	1	0	30	0	10.633698	7
1	2	0	30	0	6.549493	3
1	3	0	30	0	7.595415	3
2	1	1	44	0	9.276188	2



# Aplicación (2)

► código



Nota. Ak= Modalidad (Residencial/Ambulatoria); dtk= Meses libre de readmisión;  
vk= Confusor tiempo-dependiente

# Aplicación (3)

- Se estima una ponderación inversa al tratamiento que tiene en cuenta exposición y/o confusores tiempo-dependientes [15]

► código

- Una vez que se genera el modelo GEE sin ajustar por la probabilidad inversa de asignación a tratamiento, se contrasta con el modelo en el que se pondera por dicha probabilidad.

Se seleccionó el modelo con menor QIC.

► código

► código

Comparación resultados

	cc	t	Estimate	Std.err	lwr	upr
Modelo sin IPTW		t	0.99	1.03	0.94	1.05
Modelo con IPTW		t	0.88	1.03	0.83	0.94

# Desafíos

- Structural Nested Models
- Confusores tiempo-modificados (*time-modified confounding*) [16]
- Generar modelos estructurales de análisis de supervivencia (incluir probabilidad inversa de censura)
- Variaciones
- Modelos doble o triplemente robustos
- Machine Learning (bagging, boosting, random forests, neural networks)



# Gracias!

Contacto: [gonzalez.santacruz.andres@gmail.com](mailto:gonzalez.santacruz.andres@gmail.com)



# Referencias

- [1] M. A. Hernán, S. Hernández-Díaz, and J. M. Robins. "A structural approach to selection bias.". Eng. In: *Epidemiology* (Cambridge, Mass.). 15.5 (sept.. 2004), pp. 615-625. ISSN: 1044-3983 (Print). DOI: [10.1097/01.ede.0000135174.63482.43](https://doi.org/10.1097/01.ede.0000135174.63482.43).
- [2] J. Robins. "A new approach to causal inference in mortality studies with a sustained exposure period-application to control of the healthy worker survivor effect". En. In: "*Math. Modelling*" 7 (1986), pp. 1393-1512.
- [3] J. Kaufman. *Conceptos de Causalidad y Sesgos en Estudios Observacionales*. 2020. URL: <https://fivethirtyeight.com/features/science-isnt-broken/#part1> (visited on 07/09/2022).
- [4] A. Gelman and G. Imbens. *Why ask Why? Forward Causal Inference and Reverse Causal Questions*. Working Paper 19614. National Bureau of Economic Research, nov.. 2013. DOI: [10.3386/w19614](https://doi.org/10.3386/w19614). URL: <http://www.nber.org/papers/w19614>.
- [5] M. Á. Hernán and J. M. Robins. *Causal Inference: What If*. Boca Raton: Chapman & Hall/CRC, 2020. URL: [https://cdn1.sph.harvard.edu/wp-content/uploads/sites/1268/2019/10/ci\\_hernanrobins\\_1oct19.pdf](https://cdn1.sph.harvard.edu/wp-content/uploads/sites/1268/2019/10/ci_hernanrobins_1oct19.pdf).
- [6] J. Pearl. "Statistics and Causal Inference: A Review". In: *TEST: An Official Journal of the Spanish Society of Statistics and Operations Research* 12 (feb.. 2003), pp. 281-345. DOI: [10.1007/BF02595718](https://doi.org/10.1007/BF02595718).

# Referencias (2)

[7] A. Muriel, D. Hernández, and V. Abraira. "Modelos estructurales marginales: una herramienta útil que proporciona evidencia a los estudios observacionales". In: *Nefrologia* 2.7 (2011), pp. 7-13. ISSN:

1. DOI: 10.3265/NefrologiaSuplementoExtraordinario.pre2011.Dec.11267. URL: <https://www.revistanefrologia.com/es-modelos-estructurales-marginales-una-herramienta-articulo-X2013757511000636>.

[8] C. Aschwanden. *Science Isn't Broken Its just a hell of a lot harder than we give it credit for.* 2015. URL: <https://fivethirtyeight.com/features/science-isnt-broken/#part1> (visited on 06/09/2022).

[9] D. B. Rubin. "Randomization analysis of experimental data: The Fisher randomization test comment". In: *Journal of the American statistical association* 75.371 (1980), pp. 591-593.

[10] G. W. Imbens and D. B. Rubin. "Causality: The Basic Framework". In: *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. Cambridge University Press, 2015, p. 3â€“22. DOI: 10.1017/CBO9781139025751.002.

[11] D. Tompsett, S. Vansteelandt, O. Dukes, et al. "gestools: General Purpose G-Estimation in R". En. In: *Observational Studies* 8.1 (2022), pp. 1-28. ISSN: 2767-3324. DOI: 10.1353/obs.2022.0003. URL: <https://muse.jhu.edu/article/856403> (visited on 07/10/2022).

[12] S. Vansteelandt and N. Keiding. "Invited commentary: G-computation-lost in translation?" En. In: *Am J Epidemiol* 173.7 (mar.. 2011), pp. 739-742.

# Referencias (3)

- [13] A. I. Naimi, S. R. Cole, and E. H. Kennedy. "An introduction to g methods". In: *International Journal of Epidemiology* 46.2 (dic.. 2016), pp. 756-762. ISSN: 0300-5771. DOI: [10.1093/ije/dyw323](https://doi.org/10.1093/ije/dyw323). eprint: <https://academic.oup.com/ije/article-pdf/46/2/756/25421672/dyw323.pdf>. URL: <https://doi.org/10.1093/ije/dyw323>.
- [14] M. Bounthavong. *Using inverse probability of treatment weights & Marginal structural models to handle time-varying covariates*. 2018. URL: <https://rpubs.com/mbounthavong/IPTW MSM Tutorial> (visited on 06/09/2022).
- [15] W. M. van der Wal and R. B. Geskus. "ipw: An R Package for Inverse Probability Weighting". In: *Journal of Statistical Software* 43.13 (2011), pp. 1-23. URL: <http://www.jstatsoft.org/v43/i13/>.
- [16] R. W. Platt, E. F. Schisterman, and S. R. Cole. "Time-modified Confounding". In: *American Journal of Epidemiology* 170.6 (ago.. 2009), pp. 687-694. ISSN: 0002-9262. DOI: [10.1093/aje/kwp175](https://doi.org/10.1093/aje/kwp175). eprint: <https://academic.oup.com/aje/article-pdf/170/6/687/285709/kwp175.pdf>. URL: <https://doi.org/10.1093/aje/kwp175>.