

Experimentos de simulación para estimar parámetros de calidad de un test diagnóstico con diferentes mecanismos de verificación parcial



Universidad de Jaén



UNIVERSIDAD
DE GRANADA

VI Jornada de Estadística como
Herramienta Científica
Jaén, 4 al 5 de diciembre de 2025

Elena Cruz Guardia¹, Andrea Ruiz Vega², Miguel Á. Montero Alonso³, Juan de Dios Luna del Castillo⁴

¹ Estadística e Investigación Operativa. Universidad de Granada.

elenacruz@correo.ugr.es

² Estadística e Investigación Operativa. Universidad de Granada.

arvega@ugr.es

³ Estadística e Investigación Operativa. Universidad de Granada.

mmontero@ugr.es

⁴ Estadística e Investigación Operativa. Universidad de Granada.

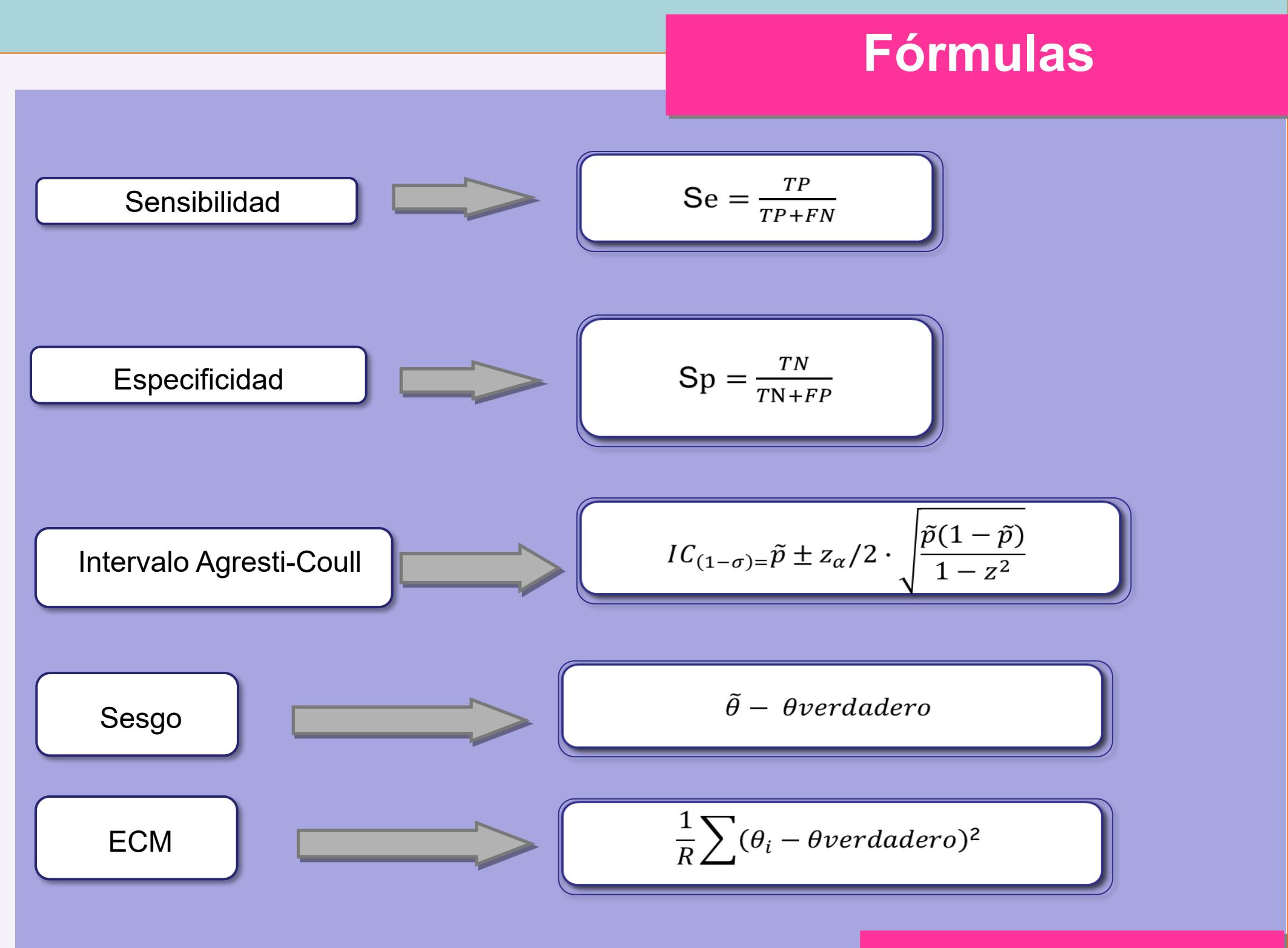
jdluna@ugr.es

En estudios diagnósticos, la estimación de la sensibilidad y la especificidad puede verse afectada cuando el gold standard solo se aplica a una parte de los pacientes, situación conocida como verificación parcial. Este sesgo depende del mecanismo que determina qué sujetos se verifican y puede comprometer la validez de los resultados si no se corrige adecuadamente. En este trabajo se realiza un experimento de simulación para evaluar el comportamiento de las estimaciones basadas en casos completos bajo distintos mecanismos de verificación y diferentes configuraciones de sensibilidad, especificidad y tamaño muestral.

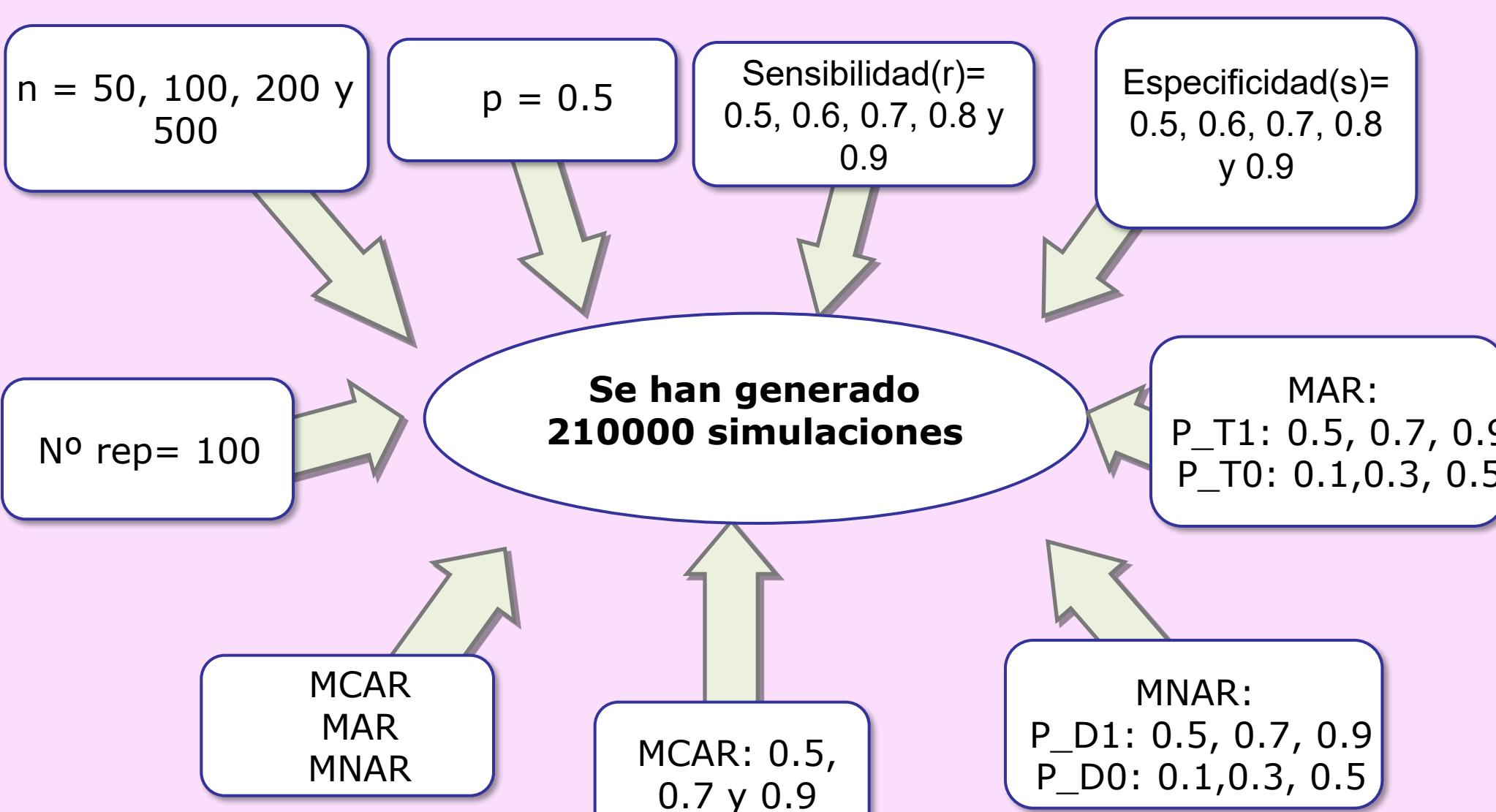
Introducción

La tabla resume las frecuencias teóricas de la tabla 2x2 que relaciona el resultado del test T(1 = positivo, 0 = negativo) y el estado real de enfermedad D(1 = enfermo, 0 = sano). A partir de esta estructura se definen la sensibilidad, la especificidad y la prevalencia. Además, recoge las probabilidades de la distribución multinomial con la que se generan las muestras simuladas en el estudio. Esta formulación permite incorporar posteriormente los mecanismos de verificación (MCAR, MAR y MNAR) que generan sesgo en las estimaciones cuando solo una parte de los pacientes es verificada.

	T = 1	T = 0
D = 1	$P(D=1 \cap T=1) = P(D=1) P(T=1 D=1) = pr$	$P(D=1 \cap T=0) = P(D=1) P(T=0 D=1) = p(1-r)$
D = 0	$P(D=0 \cap T=1) = P(D=0) P(T=1 D=0) = (1-p)(1-s)$	$P(D=0 \cap T=0) = P(D=0) P(T=0 D=0) = (1-p)s$



Experimentos de Simulación



- MCAR:** el análisis de casos completos (CCA) funciona adecuadamente: las estimaciones de Se y Sp son insesgadas y estables, incluso con pocos datos.
- MAR:** el método CCA produce estimaciones sistemáticamente sesgadas, especialmente en la especificidad, y este sesgo no se corrige ni aumentando el tamaño muestral ni el porcentaje de verificación.
- MNAR:** el comportamiento fue sorprendentemente bueno: las estimaciones obtenidas por CCA fueron muy similares a las del escenario MCAR. Este resultado, puede explicarse por la estructura de verificación simulada: Se asignó una alta probabilidad de verificación a los enfermos ($D = 1$). En cambio, los sanos ($D = 0$) tenían muchas menos probabilidades de ser verificados. Este desequilibrio en la muestra verificada enriqueció la información sobre los enfermos y, al aplicar CCA, se redujo el error en las estimaciones.

Resultados gráficos

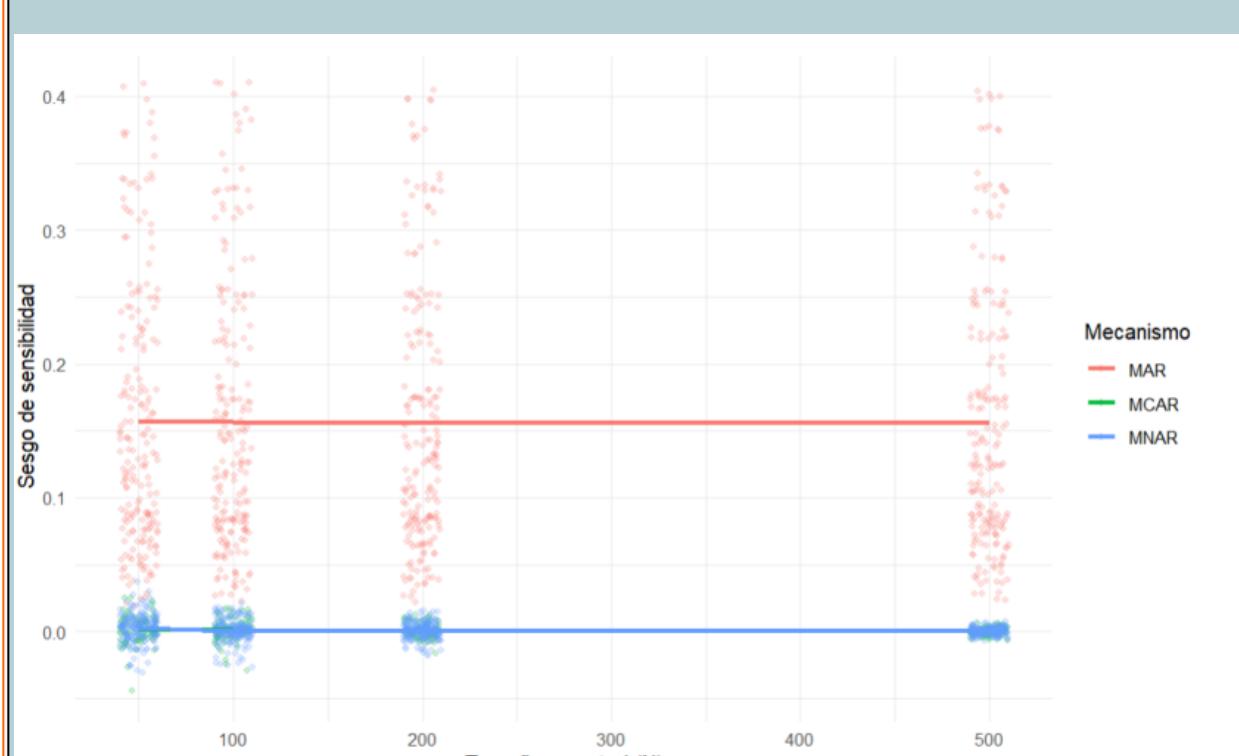


Figura 1: Sesgo sensibilidad según n y mecanismo

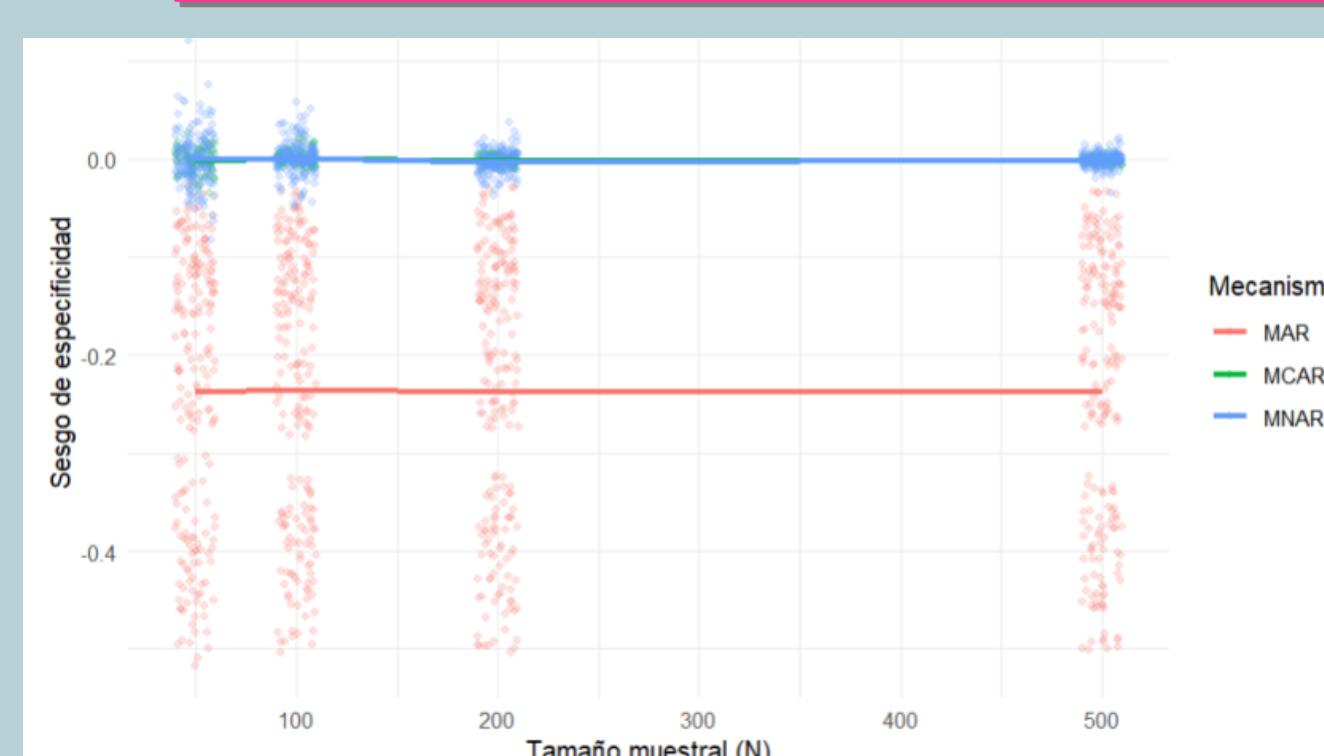


Figura 2: Sesgo especificidad según n y mecanismo

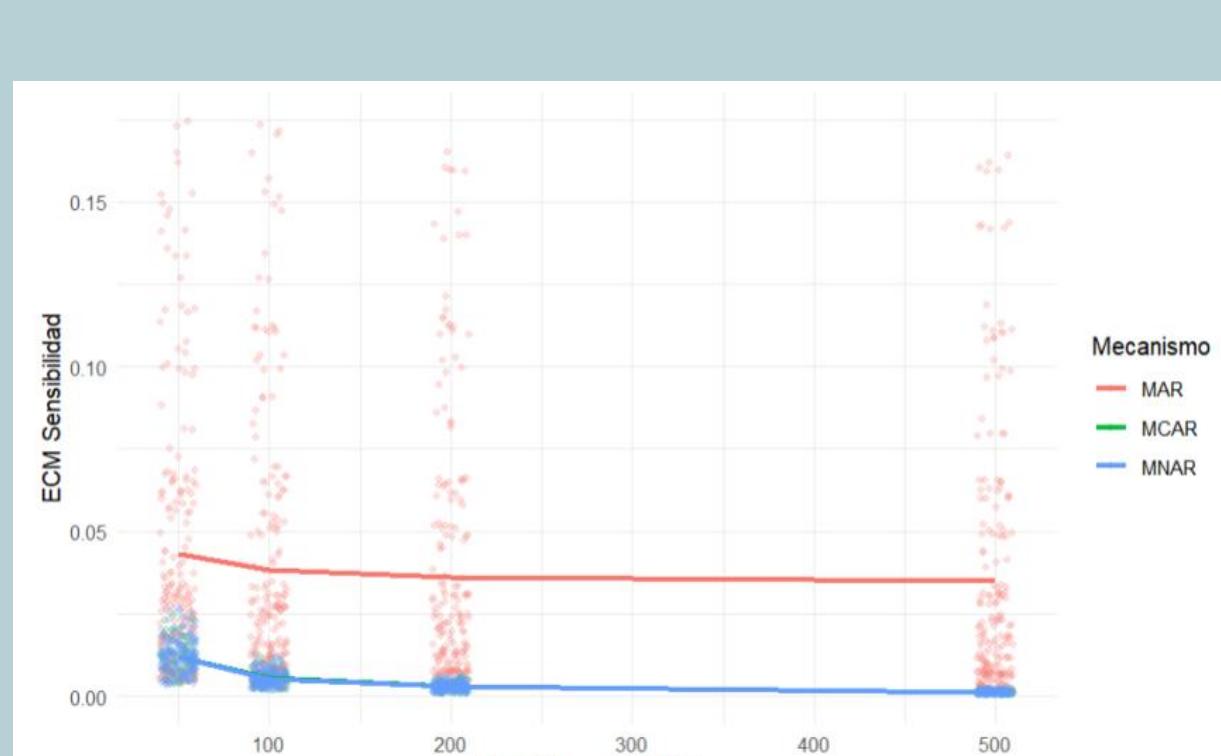


Figura 3: ECM sensibilidad según n y mecanismo

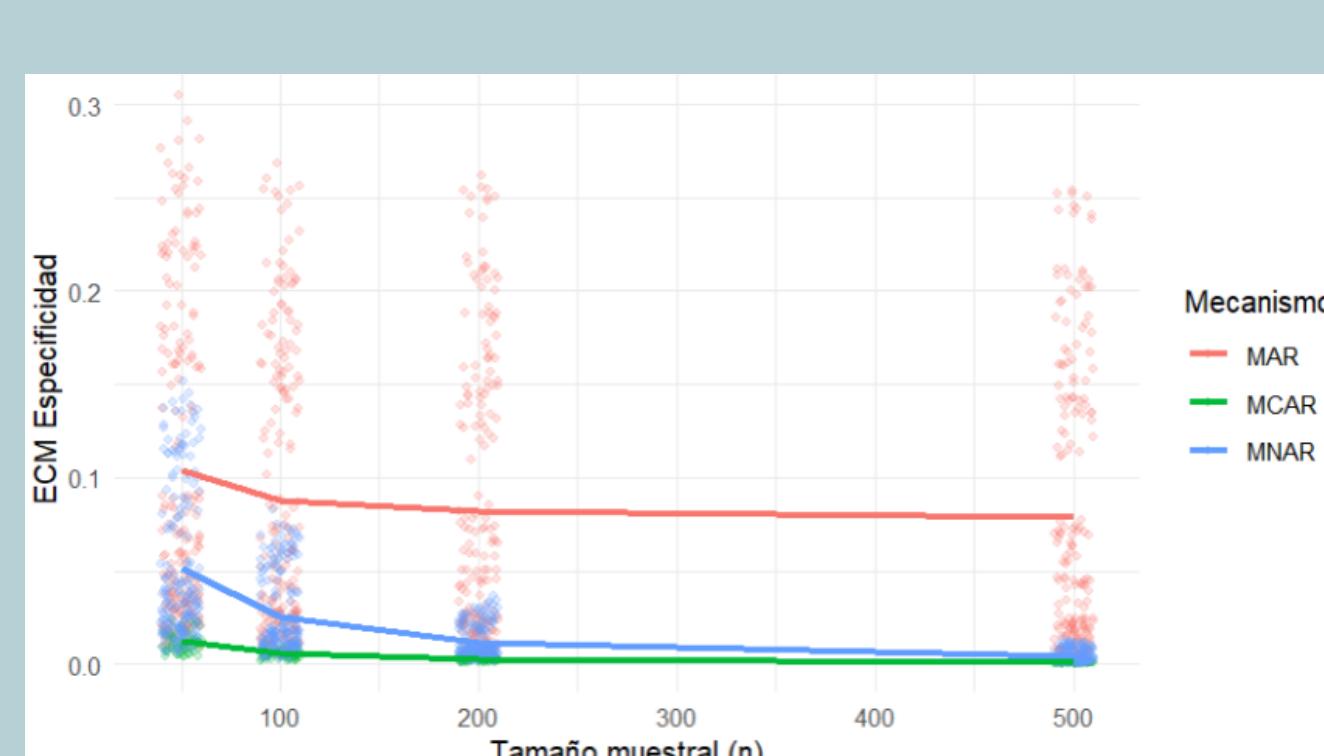


Figura 4: ECM especificidad según n y mecanismo

Bibliografía

1. Morris, T. P., White, I. R., & Crowther, M. J. (2019). Using simulation studies to evaluate statistical methods. *Statistics in Medicine*, 38(11), 2074–2102. <https://doi.org/10.1002/sim.8086>
2. Pepe, M. S. (2003). The statistical evaluation of medical tests for classification and prediction. Oxford University Press.
3. Arifin, W. N., & Yusof, U. K. (2022). Correcting for partial verification bias in diagnostic accuracy studies: A tutorial using R. *Statistics in Medicine*, 41(9), 1709–1727. <https://doi.org/10.1002/sim.9311>