

Juan Seguro

3Dresyns by Resyner Technologies S.L.

© 2026 Juan Seguro. Todos los derechos reservados.

Correspondencia: info@3Dresyns.com · www.3Dresyns.com

INF-005 · Inferencia Científica y Razonamiento Experimental · Versión v0.2

Resumen

La evidencia científica no se agota en si un resultado cruza un umbral de significación. La evidencia se refiere a cómo las observaciones discriminan entre hipótesis o modelos competidores, a lo sorprendentes que son las observaciones bajo cada alternativa y a cuánto deberían cambiar los datos un estado inferencial dados unos supuestos declarados. Este artículo desarrolla un tratamiento acotado de la fuerza de la evidencia mediante ideas de verosimilitud y factor de Bayes sin convertir INF en un curso general de estadística. Distingue la evidencia contenida en los datos de las odds previas, la creencia posterior y el control de errores a largo plazo. También explica por qué los valores p y los intervalos de confianza responden a preguntas distintas de la razón de verosimilitud entre alternativas especificadas, y por qué un valor p pequeño no mide por sí solo con qué fuerza una explicación científica queda respaldada frente a otra. El artículo no prescribe una escala universal de evidencia. Su propósito es obligar al analista a declarar las afirmaciones rivales, el modelo de datos bajo cada una y la magnitud que se interpreta como peso evidencial.

Palabras clave: fuerza de la evidencia; verosimilitud; razón de verosimilitud; factor de Bayes; valor p ; inferencia científica; hipótesis competidoras

1. La evidencia es comparativa

Un dato adquiere valor evidencial solo en relación con alternativas. El paradigma de verosimilitud de Royall lo hace explícito: la misma observación puede favorecer una hipótesis especificada frente a otra mediante una razón de verosimilitud incluso cuando la probabilidad absoluta de la observación es pequeña bajo ambas [1]. El análisis bayesiano utiliza de forma similar factores de Bayes para comparar lo bien que los datos son predichos por modelos competidores, mientras que las odds previas entran por separado cuando se calculan las odds posteriores [2,3].

2. La evidencia de los datos, la información previa y la decisión son objetos diferentes

- La evidencia de los datos se refiere a cómo los datos observados discriminan entre hipótesis especificadas bajo modelos de muestreo o generativos declarados.
- La información previa se refiere a restricciones o probabilidades asignadas antes de incorporar los datos presentes.
- La creencia posterior combina los datos con la información previa bajo un modelo bayesiano.
- Las reglas de decisión añaden utilidades, pérdidas, costes u objetivos de control de error y, por tanto, no pueden equipararse únicamente con la fuerza de la evidencia.

INF-006 posee el conocimiento previo y las tasas base. INF-005 mantiene por ello conceptualmente separada la cuestión de la evidencia de los datos, incluso cuando un cálculo bayesiano posterior combina ambos componentes.

3. Por qué la significación estadística no es fuerza de la evidencia

Un valor p se define bajo un modelo nulo y un estadístico de prueba; no es la probabilidad posterior de que la hipótesis nula sea verdadera y no compara la nula con una alternativa específica salvo que se introduzca estructura adicional. La bibliografía documenta interpretaciones erróneas persistentes de valores p , intervalos de confianza y potencia [4]. Esto no vuelve inútiles esas herramientas. Significa que su interpretación debe corresponder a la pregunta que responden.

4. Razones de verosimilitud y factores de Bayes

Para dos hipótesis especificadas H_1 y H_0 y datos observados y , la razón de verosimilitud $L(H_1|y)/L(H_0|y)$ describe cuánto más compatibles son los datos observados con una hipótesis que con la otra bajo los modelos utilizados para generar la verosimilitud. Un factor de Bayes integra o promedia la incertidumbre paramétrica según priors específicos de cada modelo y compara el respaldo predictivo marginal de los modelos [2,3]. Ninguna de las dos magnitudes es universal: ambas dependen de cómo se especifiquen los modelos competidores, y los factores de Bayes pueden depender materialmente de las distribuciones previas dentro de los modelos.

5. La evidencia puede ser débil incluso cuando una estimación es precisa

La precisión dentro de un modelo ajustado no equivale a discriminación entre modelos rivales. Un intervalo de confianza o credibilidad estrecho alrededor de un parámetro puede coexistir con incertidumbre estructural sobre qué modelo es apropiado. INF-008 posee esa distinción entre incertidumbre estructural y paramétrica. INF-005 trata por ello la fuerza evidencial como condicional a las alternativas comparadas y no permite que una estimación precisa dentro de una estructura sustituya a la comparación de modelos.

6. No existe un umbral universal de evidencia

Términos como “fuerte”, “moderada” o “decisiva” pueden ser resúmenes útiles solo cuando se declaran la escala subyacente y el contexto. INF no impone un umbral universal de razón de verosimilitud o factor de Bayes. Las consecuencias de una inferencia errónea difieren entre aplicaciones, y un umbral numérico fijo puede crear falsa precisión sobre lo que sigue siendo un juicio dependiente de supuestos. La perspectiva de pruebas severas de Mayo ofrece una advertencia complementaria: la evidencia depende no solo de un estadístico, sino también de la capacidad del procedimiento para revelar el error pertinente [5].

7. Peso evidencial y resultados negativos

No rechazar una hipótesis nula no equivale a evidencia a favor de la nula. Un resultado negativo solo puede favorecer una hipótesis frente a otra cuando las predicciones competidoras son explícitas y el diseño tiene capacidad adecuada. A la inversa, un experimento muy sensible puede producir evidencia contra una hipótesis de efecto grande incluso cuando la estimación observada sea cercana a cero. La distinción entre no detección y evidencia de ausencia sigue dependiendo de la capacidad del método.

8. Disciplina mínima de reporte

- Nombre las hipótesis o modelos competidores.
- Indique si la magnitud informada es un valor p , una razón de verosimilitud, un factor de Bayes, una probabilidad posterior, un intervalo de confianza u otro objeto inferencial.
- No convierta retóricamente una magnitud en otra.
- Indique los supuestos de modelo que determinan la medida de evidencia.
- Separe la evidencia contenida en los datos de la información previa y de los costes de decisión.
- Evite etiquetas verbales universales salvo que la convención cuantitativa subyacente y el contexto sean explícitos.

Referencias

- [1] Royall, R. Statistical Evidence: A Likelihood Paradigm. London: Chapman & Hall, 1997. DOI: 10.1201/9780203738665.
- [2] Goodman, S.N. Toward Evidence-Based Medical Statistics. 2: The Bayes Factor. Annals of Internal Medicine 130(12), 1005-1013 (1999). DOI: 10.7326/0003-4819-130-12-199906150-00019.
- [3] Kass, R.E.; Raftery, A.E. Bayes Factors. Journal of the American Statistical Association 90(430), 773-795 (1995). DOI: 10.1080/01621459.1995.10476572.
- [4] Greenland, S.; Senn, S.J.; Rothman, K.J.; Carlin, J.B.; Poole, C.; Goodman, S.N.; Altman, D.G. Statistical Tests, P Values, Confidence Intervals, and Power: A Guide to Misinterpretations. European Journal of Epidemiology 31, 337-350 (2016). DOI: 10.1007/s10654-016-0149-3.
- [5] Mayo, D.G. Statistical Inference as Severe Testing: How to Get Beyond the Statistics Wars. Cambridge University Press, 2018. DOI: 10.1017/9781107286184.

Declaración de datos y materiales

No se generó ningún conjunto de datos experimental nuevo para este artículo. El artículo es un análisis metodológico basado en la bibliografía citada y en la arquitectura de atribución de contenidos congelada de la serie INF.