

Juan Seguro

3Dresyns by Resyner Technologies S.L.

© 2026 Juan Seguro. Todos los derechos reservados.

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

INF-007 · Inferencia Científica y Razonamiento Experimental · Versión v0.2

Resumen

Un experimento es más informativo cuando explicaciones plausibles hacen predicciones diferentes bajo las condiciones realmente ensayadas. Repetir simplemente una medición, aumentar el tamaño muestral o perturbar un sistema no garantiza discriminación si todas las hipótesis activas predicen el mismo resultado. Este artículo desarrolla la lógica general de los experimentos discriminantes. Parte del conjunto de rivales establecido en INF-004, identifica predicciones que divergen entre alternativas y pregunta si la medición o intervención propuesta tiene capacidad suficiente para revelar esas divergencias. El artículo utiliza las ideas de pruebas severas e inferencia fuerte como principios de diseño, rechazando a la vez una visión simplista según la cual una sola predicción fallida falsaría mecánicamente una teoría. Deben hacerse explícitas las hipótesis auxiliares, los límites de medición, las condiciones de contorno y la dependencia del estado. El resultado de un experimento discriminante puede ser la eliminación de un rival, la retención de varios o información insuficiente. El diseño de experimentos diagnósticos específico de VPP permanece en ER-184.

Palabras clave: experimento discriminante; prueba severa; falsación; hipótesis rivales; diseño experimental; capacidad de prueba; intervención

1. Los experimentos deben separar alternativas, no limitarse a perturbar sistemas

El programa de inferencia fuerte de Platt destacó las hipótesis múltiples y los experimentos capaces de excluir alternativas [1]. La idea central sigue siendo útil: un experimento añade poco valor inferencial si todas las explicaciones activas predicen el mismo resultado. Un experimento discriminante elige, en cambio, una perturbación o medición para la que los rivales plausibles divergen.

2. Empezar por el conjunto de rivales

INF-004 posee la decisión sobre qué explicaciones permanecen activas. INF-007 toma ese conjunto como entrada. Cada rival debe vincularse a una predicción condicionada a supuestos auxiliares declarados. El objetivo de diseño no es “recoger más datos” en abstracto, sino obtener observaciones cuya distribución o patrón cualitativo difiera lo suficiente entre rivales como para cambiar el estado inferencial.

3. Severidad y capacidad

El marco de pruebas severas de Mayo subraya que un resultado aporta evidencia contra un error o discrepancia solo si el procedimiento tenía alta capacidad para revelar esa discrepancia en caso de existir [2]. INF utiliza esto como principio operativo de diseño, no como doctrina estadística universal. Un resultado negativo de un ensayo insensible discrimina débilmente; un resultado negativo de una prueba validada y de alta capacidad puede ser muy informativo respecto de hipótesis que exigían un efecto detectable.

4. La intervención puede crear contraste

La observación por sí sola puede dejar rivales observacionalmente equivalentes. La intervención puede cambiar la estructura causal o el estado del sistema de una forma que genere predicciones distintas. La inferencia causal moderna formaliza esta distinción entre observar una variable e intervenir sobre ella [3,4]. Sin embargo, la intervención no es automáticamente decisiva: incumplimiento, vías compensatorias, efectos fuera del objetivo y limitaciones de medición pueden preservar la ambigüedad.

5. Los supuestos auxiliares deben quedar expuestos

Una diferencia predicha depende de algo más que de la hipótesis focal. Puede requerir calibración correcta, intervalo de exposición adecuado, estado estable del espécimen, condiciones de contorno correctas y momento de medición válido. Por ello, una predicción fallida cuestiona la conjunción de supuestos focales y auxiliares. El diseño debe identificar qué auxiliares se están utilizando y, cuando sea posible, incluir controles que separen el fallo focal del fallo de medición o implementación.

6. Matriz de experimento discriminante

Rival	Respuesta predicha	Capacidad requerida	Interpretación si la respuesta está ausente
A	Cambio direccional grande bajo la perturbación P	Resolver el cambio por encima de la incertidumbre del método	A se debilita solo si se verifican el estado y la capacidad del método
B	Sin cambio bajo P pero con	Q debe aislar el mecanismo	La ausencia bajo P y Q

	respuesta bajo Q	exclusivo de B	puede dejar ambigüedad de implementación
C	La respuesta sigue a la variable molesta N	Medir N de forma independiente y perturbarla si es posible	C se debilita si la respuesta persiste cuando N está controlada

7. La falsación es operativa y acotada

INF no utiliza “falsado” para significar que un único punto discrepante destruye permanentemente una teoría. La falsación operativa significa que, bajo supuestos especificados y una capacidad de prueba adecuada, un resultado observado es incompatible con la predicción de un rival en grado suficiente para eliminar o debilitar materialmente a ese rival respecto de la pregunta inferencial presente. Una revisión teórica más amplia puede requerir evidencia adicional.

8. Eficiencia del diseño frente a discriminación

Un diseño estadísticamente eficiente para estimar un parámetro no es necesariamente el mejor para distinguir estructuras de modelo. La discriminación puede requerir muestrear en condiciones donde las predicciones rivales divergen con mayor fuerza, en lugar de donde el modelo preferido resulta más fácil de estimar. Esta diferencia es central cuando la incertidumbre estructural domina la incertidumbre paramétrica.

9. Disciplina mínima de reporte

- Indique las explicaciones rivales activas antes del experimento.
- Para cada rival, indique la observación predicha bajo la intervención o medición propuesta.
- Cuantifique o justifique de otro modo la capacidad de la prueba para la diferencia predicha.
- Enumere los supuestos auxiliares necesarios para la predicción.
- Incluya controles que diagnostiquen fallos de implementación o medición cuando sea factible.
- Admita “discriminación insuficiente” como desenlace aceptable.
- No generalice una prueba discriminante satisfactoria más allá del estado y dominio realmente ejercitados.

Referencias

- [1] Platt, J.R. Strong Inference. Science 146(3642), 347-353 (1964). DOI: 10.1126/science.146.3642.347.
- [2] Mayo, D.G. Statistical Inference as Severe Testing: How to Get Beyond the Statistics Wars. Cambridge University Press, 2018. DOI: 10.1017/9781107286184.
- [3] Pearl, J. Causality: Models, Reasoning, and Inference. 2nd ed. Cambridge University Press, 2009. DOI: 10.1017/CBO9780511803161.
- [4] Hernán, M.A.; Robins, J.M. Causal Inference: What If. Boca Raton: Chapman & Hall/CRC, 2020. DOI: 10.1201/9781315374932.

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.