

**Juan Seguro**

3Dresyns by Resyner Technologies S.L.

© 2026 Juan Seguro. Todos los derechos reservados.

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

INF-009 · Inferencia Científica y Razonamiento Experimental · Versión v0.2

## Resumen

Una conclusión es robusta solo respecto de un conjunto declarado de supuestos, modelos, decisiones de procesamiento de datos y especificaciones analíticas plausibles. La robustez no es, por tanto, sinónimo de verdad, replicación ni significación estadística. Este artículo define la robustez como la estabilidad de una conclusión científica a través de una familia de alternativas defendibles declarada prospectivamente o de forma transparente. Distingue sensibilidad paramétrica, sensibilidad de modelo, análisis de curva de especificaciones y análisis multiverso, y explica por qué la flexibilidad oculta del investigador pertenece, en cambio, a INF-010. El requisito central es que las especificaciones alternativas sean científicamente admisibles y no se generen simplemente para hacer que un resultado parezca estable. La robustez puede reforzar la confianza cuando convergen especificaciones materialmente distintas pero justificadas; la inestabilidad puede revelar que la conclusión científica depende de elecciones cuya incertidumbre debe informarse. Ninguno de los dos resultados establece por sí mismo el mecanismo o la estructura causal.

Palabras clave: robustez; análisis de sensibilidad; curva de especificaciones; análisis multiverso; dependencia del modelo; supuestos; inferencia científica

## 1. La robustez es estabilidad condicional

Toda conclusión empírica depende de supuestos: cómo se definen las variables, qué observaciones se incluyen, qué forma funcional se ajusta, qué covariables se controlan, cómo se tratan los valores faltantes y qué clase de modelo se considera plausible. Un análisis de robustez pregunta si la conclusión científica sobrevive a una familia declarada de alternativas, cada una defendible para la afirmación formulada. No pregunta si cualquier análisis imaginable produce la misma respuesta.

## 2. Robustez no significa verdad

Una conclusión puede ser estable a través de muchos modelos relacionados y seguir siendo errónea si toda la familia comparte el mismo defecto estructural. A la inversa, un efecto científicamente real puede parecer inestable cuando los datos disponibles son débiles, el error de medición es grande o modelos razonables apuntan a estimandos distintos. INF prohíbe por ello el colapso sinónimo robustez = verdad. La robustez es evidencia sobre dependencia de supuestos, no una garantía ontológica.

## 3. Cuatro preguntas distintas de robustez

- Sensibilidad paramétrica: ¿cambia la conclusión cuando los valores inciertos de parámetros o entradas molestas varían dentro de intervalos científicamente plausibles?
- Sensibilidad de modelo: ¿cambia la conclusión entre estructuras de modelo materialmente distintas y cada una compatible con la pregunta científica?
- Sensibilidad de especificación: ¿cambia la conclusión entre elecciones defendibles de codificación, transformación, ajuste, inclusión o forma funcional?
- Sensibilidad al objetivo: ¿cambia la conclusión cuando cambian el estimando, la población objetivo o el horizonte temporal? Un cambio aquí puede reflejar una pregunta distinta y no fragilidad.

## 4. Curvas de especificaciones y análisis multiverso

El análisis de curva de especificaciones hace visible un conjunto preespecificado o transparentemente justificado de especificaciones no redundantes y evalúa cómo varían los resultados entre ellas [1]. El análisis multiverso expone de forma similar las consecuencias de alternativas de procesamiento de datos y análisis realizando análisis sobre un conjunto de escenarios razonables [2]. Estas herramientas son valiosas porque revelan dependencias que un único análisis preferido oculta. Su validez sigue dependiendo de cómo se defina el conjunto de especificaciones admisibles.

## 5. Qué pertenece al conjunto admisible

Una alternativa pertenece a un análisis de robustez cuando es científicamente defendible antes de observar si respalda la conclusión preferida. Las alternativas que cambian la afirmación, el estimando o la población objetivo deben etiquetarse como preguntas distintas. Las alternativas que se sabe que violan supuestos no deberían contarse solo para ampliar la curva. El analista debe indicar por qué cada rama es admisible y qué supuestos difieren entre ramas.

## 6. Distinción respecto de INF-010

INF-009 posee los análisis transparentes de robustez en los que el analista expone deliberadamente la dependencia respecto de alternativas declaradas. INF-010 posee la multiplicidad analítica cuando existen muchos caminos posibles, pero solo un subconjunto se selecciona, enfatiza o informa después de ver los resultados. El mismo conjunto de

decisiones analíticas puede, por tanto, utilizarse constructivamente en INF-009 o convertirse en un mecanismo de sesgo en INF-010 según la multiplicidad esté controlada prospectivamente o se revele con transparencia.

## 7. Disciplina de interpretación

- La convergencia entre especificaciones plausibles refuerza la afirmación solo frente a las elecciones concretas que se variaron.
- La inestabilidad identifica una dependencia que pertenece a la declaración de incertidumbre; no invalida automáticamente todas las especificaciones.
- Un análisis de robustez debe informar la distribución y el patrón de los resultados, no solo la proporción que cruza un umbral de significación [3].
- La robustez frente a la especificación estadística no establece identificación causal, mecanismo ni generalización.
- Si todas las especificaciones comparten el mismo defecto de medición o supuesto estructural, la robustez no puede diagnosticar ese fallo común.

## 8. Reporte mínimo

- Indique la afirmación científica y el estimando que permanecen fijos.
- Defina la familia de especificaciones admisibles antes de interpretar sus resultados.
- Explique por qué cada rama principal es científicamente plausible.
- Informe qué decisiones impulsan cambios materiales.
- Separe los cambios causados por estimandos diferentes de los cambios dentro de un mismo estimando.
- Indique explícitamente qué supuestos no se variaron.

## Referencias

[1] Simonsohn, U.; Simmons, J.P.; Nelson, L.D. Specification curve analysis. *Nature Human Behaviour* 4, 1208-1214 (2020). DOI: 10.1038/s41562-020-0912-z.

[2] Steegen, S.; Tuerlinckx, F.; Gelman, A.; Vanpaemel, W. Increasing Transparency Through a Multiverse Analysis. *Perspectives on Psychological Science* 11(5), 702-712 (2016). DOI: 10.1177/1745691616658637.

[3] 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.

## 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.