

KNO-015 · Incertidumbre y probabilidad

KNO · Formas de conocer

Juan Segurola

v0.7 · ES

Todos los derechos reservados

Formas de conocer

Perfil epistémico: Interpretación científica actual **Licencia al publicar:** Todos los derechos reservados

Resumen del artículo

Los resultados científicos suelen venir acompañados de intervalos, probabilidades, barras de error o análisis de sensibilidad, pero ¿qué clases de incertidumbre representan realmente esas herramientas? Este artículo pregunta en qué se diferencian la incertidumbre de medida, la variación muestral, la incertidumbre paramétrica, la incertidumbre de modelo, la incertidumbre estructural y el juicio experto, y qué ocurre cuando se comprimen en un solo número. Examina intervalos de confianza, intervalos creíbles, valores p, propagación, predicción y análisis de sensibilidad preguntando qué afirmaciones dependen del marco estadístico elegido y qué límites permanecen fuera de él. También pregunta cómo comunicar incertidumbre sin convertir ignorancia en falsa precisión ni tratar todo desconocimiento como cuantificable. El problema central es metodológico: qué representación de la incertidumbre corresponde a la afirmación y qué queda sin representar aun después del cálculo.

En 60 segundos

Un número sin declaración de incertidumbre puede parecer más exacto de lo que permite la evidencia. Sin embargo, añadir «±» no basta. El intervalo puede describir el comportamiento en muestreos repetidos, una distribución de probabilidad asignada a un mensurando, una distribución posterior bayesiana, una distribución predictiva o una convención basada en un factor de cobertura. Son objetos que responden a preguntas distintas.

La disciplina práctica es, por tanto: identificar la magnitud de interés; enumerar las fuentes materiales de incertidumbre; decidir cuáles pueden representarse probabilísticamente; especificar el marco inferencial; propagar la incertidumbre a través del modelo de medida o estadístico; examinar la sensibilidad a los supuestos; y explicar qué significa realmente el intervalo o la probabilidad resultante. La probabilidad puede organizar la incertidumbre. No puede borrar los supuestos, las omisiones desconocidas ni la mala especificación del modelo. La idea práctica es separar el resultado observado de los supuestos que lo conectan con la interpretación. Cada herramienta reduce algunas ambigüedades y deja otras abiertas; por eso el análisis debe declarar qué puede identificar, qué depende de condiciones adicionales y qué observación podría contradecir la conclusión. La precisión formal no elimina por sí sola esos límites.

Por qué importa esta pregunta

Las decisiones científicas suelen tomarse bajo incertidumbre, pero incertidumbres distintas exigen respuestas distintas. Un intervalo estrecho puede coexistir con un modelo mal especificado; un instrumento preciso puede medir la cantidad equivocada; una probabilidad puede estar bien calculada y excluir el desconocimiento más importante. Tratar todas las incertidumbres como intercambiables produce falsa precisión. Distinguir qué se mide, qué se muestrea, qué se modeliza y qué queda fuera del modelo es esencial para interpretar la evidencia y comunicar la fuerza de una conclusión.

1. La incertidumbre no es una sola sustancia

La palabra cubre mecanismos distintos que no deben colapsarse en una categoría residual única.

[C01] La incertidumbre de medida se refiere al conocimiento incompleto del valor de un mensurando dados el procedimiento de medida, las entradas, las correcciones y el modelo. El marco GUM establece reglas generales para evaluar y expresar esa incertidumbre, y documentos posteriores del JCGM amplían el marco a la propagación de distribuciones y a los modelos de medida. [S01–S03]

[C02] La incertidumbre muestral aparece porque una muestra finita es solo una realización de un proceso generador de datos más amplio. Su tratamiento estadístico depende del mecanismo de muestreo o asignación y del estimador estudiado. Es conceptualmente distinta de la incertidumbre de calibración de un instrumento o de la incertidumbre estructural del modelo. [S01, S08]

[C03] La incertidumbre del modelo incluye la forma funcional, la estructura omitida, las aproximaciones y las especificaciones alternativas plausibles. La incertidumbre estructural no siempre puede reducirse a un único error estándar calculado dentro de un modelo elegido. [S07, S09]

Estas categorías pueden interactuar. Una constante de calibración puede entrar en un modelo estadístico; un modelo puede transformar entradas medidas; un diseño muestral puede determinar qué parámetros son estimables. La exigencia metodológica no consiste en aislar artificialmente cada fuente, sino en evitar que un componente cuantificado se confunda con toda la incertidumbre.

En qué punto estamos

La palabra cubre mecanismos distintos que no deben colapsarse en una categoría residual única. Las incertidumbres de medida, muestral y del modelo son conceptualmente distintas, pero pueden interactuar, como cuando una constante de calibración entra en un modelo estadístico. Lo decisivo es que ningún componente cuantificado se confunda con toda la incertidumbre. La sección siguiente examina la idea de que un resultado de medida no son solo sus dígitos mostrados.

2. Un resultado de medida no son solo sus dígitos mostrados

Un instrumento digital puede mostrar 12,345678 sin establecer conocimiento hasta seis decimales.

[C04] En el enfoque GUM/NIST, los componentes individuales de incertidumbre estándar se combinan, incluida la covarianza cuando corresponde, para obtener una incertidumbre estándar combinada. Cuando procede esa forma de comunicación, la incertidumbre expandida se obtiene multiplicando la incertidumbre estándar combinada por un factor de cobertura. [S01, S04]

[C05] Un factor de cobertura como $k = 2$ no garantiza mecánicamente una cobertura del 95 % en todos los casos. La interpretación probabilística depende de supuestos distribucionales y de otro tipo, y un número pequeño de grados de libertad efectivos puede exigir factores diferentes. [S04, S05]

La lección es general. La precisión que imprime un programa es formato. La precisión evidencial procede del modelo de incertidumbre y de la justificación de sus supuestos.

En qué punto estamos

Un instrumento digital puede mostrar 12,345678 sin establecer conocimiento hasta seis decimales. La precisión que imprime un programa es formato. La precisión evidencial procede del modelo de incertidumbre y de la justificación de sus supuestos. La sección siguiente examina la idea de que la probabilidad es una herramienta de representación, no un disolvente de la ignorancia.

3. La probabilidad es una herramienta de representación, no un disolvente de la ignorancia

La probabilidad entra en la ciencia con más de una función. En modelos de muestreo repetido caracteriza el comportamiento de variables aleatorias bajo repeticiones. En análisis bayesiano también se utiliza para representar incertidumbre sobre cantidades desconocidas condicionada a un modelo, una información previa y unos datos observados. En metrología, se emplean distribuciones de probabilidad para representar valores razonablemente atribuibles a cantidades inciertas y propagar esas distribuciones a través de modelos de medida. [S02, S03]

[C06] La misma probabilidad numérica puede, por tanto, proceder de construcciones inferenciales distintas. Un informe científicamente adecuado debe indicar a qué objeto aleatorio o cantidad incierta se refiere y bajo qué supuestos. [S02, S06]

KNO-015 no decide el significado filosófico de la probabilidad. Eso corresponde a IDE-008. Su preocupación es operacional: no pasar de una interpretación a otra sin declararlo.

En qué punto estamos

La probabilidad entra en la ciencia con más de una función. [S02, S03] KNO-015 no decide el significado filosófico de la probabilidad. En análisis bayesiano también se utiliza para representar incertidumbre sobre cantidades desconocidas condicionada a un modelo, una información previa y unos datos observados. La sección siguiente examina la idea de que los intervalos de confianza y los intervalos creíbles responden a preguntas distintas.

4. Los intervalos de confianza y los intervalos creíbles responden a preguntas distintas

La expresión «intervalo del 95 %» está incompleta.

[C07] Un procedimiento frecuentista de confianza se calibra mediante sus propiedades en muestreos repetidos bajo el modelo supuesto. La interpretación habitual del 95 % de confianza se refiere a la cobertura a largo plazo del procedimiento, no a una probabilidad posterior de que el parámetro fijo se encuentre en el intervalo concreto observado. [S10]

[C08] Un intervalo creíble bayesiano es una declaración probabilística sobre un parámetro o cantidad condicionada al modelo bayesiano especificado, incluidas la distribución previa y la verosimilitud. Por tanto, su interpretación no es intercambiable con la de un intervalo de confianza frecuentista, aunque sus extremos numéricos sean similares. [S10]

Ninguna etiqueta hace desaparecer los supuestos. La pregunta relevante es qué datos, modelo y construcción inferencial justifican el intervalo.

En qué punto estamos

La expresión «intervalo del 95 %» está incompleta. Ninguna etiqueta hace desaparecer los supuestos. La pregunta relevante es qué datos, modelo y construcción inferencial justifican el intervalo. Lo importante aquí es la relación declarada entre evidencia, método e inferencia, no una promesa genérica de certeza. La sección siguiente examina la idea de que un valor p no es la probabilidad de que una hipótesis sea verdadera.

5. Un valor p no es la probabilidad de que una hipótesis sea verdadera

Este error es tan frecuente que merece una frontera propia.

[C09] La American Statistical Association señala que los valores p pueden indicar incompatibilidad entre los datos y un modelo estadístico especificado, pero no miden la probabilidad de que la hipótesis estudiada sea verdadera, no miden el tamaño ni la importancia científica del efecto y no deberían determinar por sí solos conclusiones científicas o decisiones de política. [S06]

Un valor p se calcula bajo un modelo especificado, normalmente incluyendo una hipótesis nula y otros supuestos. No es una probabilidad posterior de esa hipótesis salvo que se construya adicionalmente un modelo bayesiano; en tal caso, la probabilidad posterior resultante es otro objeto.

En qué punto estamos

Este error es tan frecuente que merece una frontera propia. No es una probabilidad posterior de esa hipótesis salvo que se construya adicionalmente un modelo bayesiano; en tal caso, la probabilidad posterior resultante es otro objeto. La sección siguiente examina la idea de que la incertidumbre se propaga a través de modelos.

6. La incertidumbre se propaga a través de modelos

Si una salida depende de entradas inciertas, la incertidumbre de esas entradas se propaga hacia la salida.

[C10] La GUM proporciona reglas analíticas de propagación bajo sus condiciones de aplicabilidad, mientras JCGM 101 especifica la propagación de distribuciones de probabilidad a través de un modelo de medida mediante métodos de Monte Carlo. [S01, S02]

La propagación Monte Carlo no es magia. Hereda las distribuciones de entrada, la estructura de dependencia y el modelo de medida suministrados. Si son deficientes, una simulación enorme puede producir una respuesta muy precisa para un modelo de incertidumbre equivocado.

En qué punto estamos

Si una salida depende de entradas inciertas, la incertidumbre de esas entradas se propaga hacia la salida. La propagación Monte Carlo no es magia. Hereda las distribuciones de entrada, la estructura de dependencia y el modelo de medida suministrados. La elección del modelo determina también qué incertidumbre se transmite. La sección siguiente examina la idea de que la incertidumbre estructural necesita visibilidad propia.

7. La incertidumbre estructural necesita visibilidad propia

Algunas incertidumbres quedan fuera de la matriz de covarianza de parámetros del modelo seleccionado: mecanismos alternativos, condiciones de contorno, retroalimentaciones no representadas, dependencia desconocida o desacuerdo experto no resuelto.

[C11] La guía de incertidumbre del IPCC separa explícitamente la evaluación de evidencia y acuerdo, confianza, verosimilitud o probabilidad, y pide declarar supuestos y evaluar las incertidumbres estructurales cuando proceda. [S07]

Esta separación evita un error de categoría. Una distribución probabilística puede resumir incertidumbre condicionada a un modelo mientras la incertidumbre sobre el propio modelo permanece solo parcialmente cuantificada.

En qué punto estamos

Algunas incertidumbres quedan fuera de la matriz de covarianza de parámetros del modelo seleccionado: mecanismos alternativos, condiciones de contorno, retroalimentaciones no representadas, dependencia desconocida o desacuerdo experto no resuelto. Una distribución probabilística puede resumir incertidumbre condicionada a un modelo mientras la incertidumbre sobre el propio modelo permanece solo parcialmente cuantificada. La sección siguiente examina la idea de que el juicio experto puede disciplinarse sin fingir que es un dato observado.

8. El juicio experto puede disciplinarse sin fingir que es un dato observado

En algunos dominios hay pocas frecuencias directas, los modelos discrepan o las condiciones futuras no son repeticiones de un proceso histórico estable. El juicio experto puede entonces contribuir a la evaluación de incertidumbre.

[C12] La guía del IPCC utiliza lenguaje calibrado y una evaluación trazable de evidencia y acuerdo para disciplinar esos juicios, en vez de presentar toda declaración de confianza como si fuera una frecuencia observada en datos. [S07]

El juicio experto no es automáticamente más débil ni más fuerte que una salida estadística. Su función evidencial depende de transparencia, calibración, independencia, diseño de la elicitación y evidencia utilizada por los expertos. La regla crítica es la procedencia: el lector debe saber cuándo un número procede de observaciones, un conjunto de modelos, una elicitación formal o una combinación.

En qué punto estamos

En algunos dominios hay pocas frecuencias directas, los modelos discrepan o las condiciones futuras no son repeticiones de un proceso histórico estable. El juicio experto puede entonces contribuir a la evaluación de incertidumbre. La sección siguiente examina la idea de que los desconocidos desconocidos no están dentro de todos los intervalos.

9. Los desconocidos desconocidos no están dentro de todos los intervalos

Un presupuesto de incertidumbre solo puede incluir fuentes representadas en el análisis.

[C13] Los informes científicos deben distinguir, por tanto, la incertidumbre cuantificada de las limitaciones reconocidas pero no cuantificadas y del error genuinamente no reconocido. Las National Academies recomiendan comunicar incertidumbres reconocidas y fuentes potenciales en dimensiones estocásticas, de medida, computacionales, de conocimiento, de modelización y de análisis, en lugar de tratar un único intervalo estadístico como exhaustivo. [S08]

Eso no vuelve inútiles los intervalos cuantificados. Define su jurisdicción.

En qué punto estamos

Un presupuesto de incertidumbre solo puede incluir fuentes representadas en el análisis. Eso no vuelve inútiles los intervalos cuantificados. Su valor está en hacer inspeccionable este paso inferencial concreto, no en tratar el método como autojustificado. La sección siguiente examina la idea de que el análisis de sensibilidad prueba cuánto dependen las conclusiones de los supuestos.

10. El análisis de sensibilidad prueba cuánto dependen las conclusiones de los supuestos

Si una conclusión cambia drásticamente bajo especificaciones alternativas plausibles, la incertidumbre no queda capturada citando solo el error estándar de una especificación.

Un análisis de sensibilidad puede variar distribuciones previas, formas funcionales, parámetros auxiliares, supuestos sobre datos ausentes, entradas de calibración o elecciones estructurales plausibles. No es una cura universal: el analista todavía debe decidir qué alternativas son plausibles. Pero muestra dónde dependen las conclusiones de elecciones que de otro modo permanecerían ocultas.

En qué punto estamos

Si una conclusión cambia drásticamente bajo especificaciones alternativas plausibles, la incertidumbre no queda capturada citando solo el error estándar de una especificación. Un análisis de sensibilidad puede variar distribuciones previas, formas funcionales, parámetros auxiliares, supuestos sobre datos ausentes, entradas de calibración o elecciones estructurales plausibles. La sección siguiente examina la idea de que la incertidumbre predictiva no es la incertidumbre de parámetros.

11. La incertidumbre predictiva no es la incertidumbre de parámetros

Estimar una respuesta media y predecir una observación nueva son tareas distintas. Un parámetro del modelo puede estar bien estimado mientras los resultados futuros individuales siguen siendo muy variables.

[C14] Una declaración completa de incertidumbre debe identificar la cantidad objetivo: un parámetro, un estado latente, un mensurando, una observación futura, una media poblacional o un resultado relevante para una decisión. Objetivos distintos requieren en general construcciones de incertidumbre distintas. [S01, S03, S10]

Por eso «la incertidumbre es $\pm x$ » suele estar insuficientemente especificado.

En qué punto estamos

Estimar una respuesta media y predecir una observación nueva son tareas distintas. Un parámetro del modelo puede estar bien estimado mientras los resultados futuros individuales siguen siendo muy variables. Por eso «la incertidumbre es $\pm x$ » suele estar insuficientemente especificado. Esa diferencia cambia qué intervalo debe interpretarse. La sección siguiente examina un flujo de trabajo práctico para la incertidumbre.

12. Un flujo de trabajo práctico para la incertidumbre

Para el trabajo científico, la siguiente secuencia es más robusta que añadir un intervalo al final:

1. Definir la cantidad objetivo y el estimando o mensurando.
2. Identificar el proceso generador de datos o de medida.
3. Enumerar las fuentes materiales de incertidumbre y sus dependencias.
4. Separar variabilidad estocástica y conocimiento incompleto cuando la distinción sea relevante.
5. Especificar el modelo y el marco inferencial.
6. Cuantificar los componentes que puedan justificarse cuantitativamente.
7. Propagarlos a través del modelo, analítica o computacionalmente según proceda.
8. Probar la sensibilidad frente a supuestos alternativos plausibles.
9. Registrar limitaciones importantes que permanezcan sin cuantificar.
10. Explicar el significado de cada intervalo, probabilidad y declaración de cobertura.

El flujo no garantiza verdad. Hace auditable el camino desde la evidencia hasta la afirmación de incertidumbre.

En qué punto estamos

En lugar de añadir un intervalo al final, el flujo define la cantidad objetivo y su proceso de datos o medida, enumera fuentes de incertidumbre, separa variabilidad y conocimiento incompleto, especifica modelo y marco, cuantifica y propaga componentes justificados, prueba la sensibilidad, registra lo no cuantificado y explica qué significa cada intervalo. Hace auditable el camino sin garantizar verdad. La sección siguiente examina lo que la incertidumbre no autoriza.

13. Lo que la incertidumbre no autoriza

La incertidumbre no permite afirmar que todas las hipótesis sean igualmente plausibles. Tampoco un intervalo estrecho demuestra que un modelo sea correcto. Un resultado puede ser estadísticamente preciso y estar sistemáticamente equivocado. Un intervalo ancho puede reflejar honestamente información débil en vez de un fallo metodológico.

La mejor declaración de incertidumbre no es necesariamente el número más pequeño. Es aquella cuyo alcance, supuestos y construcción resisten la inspección.

En qué punto estamos

La incertidumbre no permite afirmar que todas las hipótesis sean igualmente plausibles. Un resultado puede ser estadísticamente preciso y estar sistemáticamente equivocado. Un intervalo ancho puede reflejar honestamente información débil en vez de un fallo metodológico. Esto delimita el alcance del artículo y prepara la síntesis final sin ampliar la afirmación más allá de lo establecido en las secciones anteriores.

Nota de frontera

KNO-015 trata **los métodos mediante los cuales el trabajo científico identifica, representa, propaga y comunica incertidumbre**. No pretende resolver qué es fundamentalmente la probabilidad; esa cuestión conceptual corresponde a **IDE-008 · Probability / Probabilidad**. Tampoco trata la incertidumbre como concepto ontológico o epistemológico general; eso corresponde a **IDE-017 · Uncertainty / Incertidumbre**. Se apoya directamente en **KNO-011 · Measuring Before Theorising**, **KNO-012 · Experiments and Variable Control**, **KNO-013 · Models, Maps and Reality** y **KNO-014 · Causality and Correlation**.

Enunciados de encuadre declarados no representados como afirmaciones numeradas

Esta lista es una ayuda de control, **no una garantía de exhaustividad semántica**. Registra enunciados sustantivos de encuadre que se han dejado deliberadamente fuera del registro numerado de afirmaciones y explica el motivo.

- «La incertidumbre no permite afirmar que todas las hipótesis sean igualmente plausibles». Encuadre normativo de la incertidumbre calibrada, no una afirmación empírica distinta.
- «Un resultado puede ser estadísticamente preciso y estar sistemáticamente equivocado». Advertencia metodológica condensada, ya delimitada por las secciones sobre incertidumbre sistemática y de modelo.
- «Un intervalo ancho puede reflejar honestamente información débil en vez de un fallo metodológico». Encuadre interpretativo de la amplitud del intervalo, no una regla cuantitativa independiente.
- «La mejor declaración de incertidumbre no es necesariamente el número más pequeño». Síntesis de las distinciones de incertidumbre anteriores, no una nueva proposición atribuida a una fuente.

Lo que sabemos y cómo lo sabemos

La incertidumbre puede proceder de la medida, el muestreo, los parámetros, la predicción, la estructura del modelo y el juicio, y esas fuentes no son intercambiables. Los intervalos y probabilidades estadísticas adquieren significado dentro de su marco inferencial, mientras el análisis de sensibilidad expone la dependencia de los supuestos.

Lo que todavía no sabemos

Ningún intervalo finito contiene automáticamente todo desconocimiento relevante. Omisiones estructurales, mecanismos no modelizados y posibilidades realmente imprevistas pueden permanecer fuera de un modelo probabilístico aunque la aritmética dentro de él sea correcta.

Qué podría hacernos cambiar de opinión

Una conclusión debe revisarse si cambia la calibración, nuevos datos amplían o reducen la incertidumbre muestral, estructuras de modelo alternativas ajustan de forma comparable, los análisis de sensibilidad revelan supuestos frágiles o mecanismos ignorados pasan a ser plausibles y relevantes.

Si solo recuerdas una cosa

Representa la incertidumbre en la forma que corresponde a la afirmación y declara qué desconocimientos importantes quedan fuera del cálculo.

Artículos relacionados

- KNO-011 · Measuring Before Theorising
- KNO-014 · Causality and Correlation
- KNO-016 · How Evidence Is Built: Putting the Tools Together
- KNO-021 · Open Science and Reproducibility

Referencias

[S01] Joint Committee for Guides in Metrology (JCGM). *Evaluation of measurement data — Guide to the expression of uncertainty in measurement*. JCGM 100:2008(E). DOI: 10.59161/JCGM100-2008E.

[S02] Joint Committee for Guides in Metrology (JCGM). *Evaluation of measurement data — Supplement 1 to the “Guide to the expression of uncertainty in measurement” — Propagation of distributions using a Monte Carlo method*. JCGM 101:2008. DOI: 10.59161/JCGM101-2008.

[S03] Joint Committee for Guides in Metrology (JCGM). *Guide to the expression of uncertainty in measurement — Part 6: Developing and using measurement models*. JCGM GUM-6:2020. DOI: 10.59161/JCGMGUM-6-2020.

[S04] Taylor, B. N.; Kuyatt, C. E. *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*. NIST Technical Note 1297. Manifestación web de NIST consultada el 5 de septiembre de 2026; página actualizada el 6 de mayo de 2026. <https://www.nist.gov/pml/nist-technical-note-1297>

[S05] National Institute of Standards and Technology. *NIST TN 1297: Appendix B. Coverage Factors*. Página web de NIST, actualizada el 11 de septiembre de 2025; consultada el 5 de septiembre de 2026. <https://www.nist.gov/pml/nist-technical-note-1297/nist-tn-1297-appendix-b-coverage-factors>

[S06] Wasserstein, Ronald L.; Lazar, Nicole A. “The ASA’s Statement on p-Values: Context, Process, and Purpose.” *The American Statistician* 70(2), 2016, 129–133. DOI: 10.1080/00031305.2016.1154108.

[S07] Mastrandrea, Michael D.; Field, Christopher B.; Stocker, Thomas F.; Edenhofer, Ottmar; Ebi, Kristie L.; Frame, David J.; Held, Hermann; Kriegler, Elmar; Mach, Katharine J.; Matschoss, Patrick R.; Plattner, Gian-Kasper; Yohe, Gary W.; Zwiers, Francis W. *Guidance Note for Lead Authors of the IPCC Fifth Assessment Report on Consistent Treatment of Uncertainties*. IPCC, 2010.

[S08] National Academies of Sciences, Engineering, and Medicine. *Reproducibility and Replicability in Science*. Washington, DC: The National Academies Press, 2019. DOI: 10.17226/25303.

[S09] Draper, David. “Assessment and Propagation of Model Uncertainty.” *Journal of the Royal Statistical Society: Series B* 57(1), 1995, 45–70. DOI: 10.1111/j.2517-6161.1995.tb02015.x.

[S10] Possolo, Antonio; Gelabert, Maria Victoria; Hibbert, David; Stohner, Jürgen; Bodnar, Olha; Meija, Juris. *Guía Breve sobre la Incertidumbre de Medición*. NIST Technical Note 2330, 2025. DOI: 10.6028/NIST.TN.2330.

[S11] Joint Committee for Guides in Metrology (JCGM). *Evaluation of measurement data — Guide to the expression of uncertainty in measurement. Amendment 1: Nonlinearity in measurement models*. JCGM 100:2008/Amd.1:2026. DOI: 10.59161/PPDI3267.