



NIST Series Technical Note NIST TN 2330

Guía Breve sobre la Incertidumbre de Medición

Maria Victoria Gelabert
Antonio Possolo
David Brynn Hibbert
Jürgen Stohner
Olha Bodnar
Juris Meija

This publication is available free of charge from:
<https://doi.org/10.6028/NIST.TN.2330>

NIST Series Technical Note

NIST TN 2330

Guía Breve sobre la Incertidumbre de Medición

Maria Victoria Gelabert

Laboratorio Tecnológico del Uruguay, Montevideo, Uruguay

Antonio Possolo

Statistical Engineering Division, Information Technology Laboratory

National Institute of Standards and Technology

David Brynn Hibbert

School of Chemistry, UNSW, Sydney, Australia

Jürgen Stohner

Institute of Chemistry and Biotechnology

Zürich Univ. of Applied Sciences, Wädenswil, Switzerland

Olha Bodnar

Örebro Univ. School of Business, Örebro, Sweden

Juris Meija

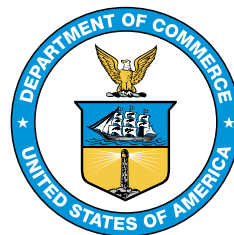
Metrology Research Centre

National Research Council Canada, Ottawa, Canada

This publication is available free of charge from:

<https://doi.org/10.6028/NIST.TN.2330>

January 2025



U.S. Department of Commerce

Gina M. Raimondo, Secretary

National Institute of Standards and Technology

Laurie E. Locascio, NIST Director and Under Secretary of Commerce for Standards and Technology

En este documento se identifican ciertos equipos, instrumentos, software o materiales, comerciales o no comerciales, para especificar adecuadamente procedimientos experimentales. Dicha identificación no implica la recomendación ni el respaldo de ningún producto o servicio por parte del NIST, ni implica que los materiales o equipos identificados sean necesariamente los mejores disponibles para el propósito.

NIST Technical Series Policies

[Copyright, Use, and Licensing Statements](#)

[NIST Technical Series Publication Identifier Syntax](#)

Publication History

Approved by the NIST Editorial Review Board on 2025-02-07

How to cite this NIST Technical Series Publication:

Gelabert MV, Possolo A, Hibbert DB, Stohner J, Bodnar O, Meija J (2025) Guía Breve sobre la Incertidumbre de Medición. (National Institute of Standards and Technology, Gaithersburg, MD), NIST TN 2330. <https://doi.org/10.6028/NIST.TN.2330>

Author ORCID iDs

Maria Victoria Gelabert: 0009-0000-5444-2902

Antonio Possolo: 0000-0002-8691-4190

David Brynn Hibbert: 0000-0001-9210-2941

Jürgen Stohner: 0000-0003-0079-7083

Olha Bodnar: 0000-0003-1359-3311

Juris Meija: 0000-0002-3349-5535

Contact Information

antonio.possolo@nist.gov

Traducción de un informe técnico de la IUPAC con material complementario. Publicación original en inglés © 2023 IUPAC & De Gruyter, licenciada bajo la licencia Creative Commons Atribución-NoComercial-SinDerivadas 4.0 Internacional, publicada en *Pure and Applied Chemistry* 2024; 96(1): 113–134. DOI <https://doi.org/10.1515/pac-2022-1203>

Resumen

Esta *Guía Breve* reintroduce a los lectores en los principales conceptos y herramientas técnicas utilizadas para la evaluación y expresión de la incertidumbre de medición, incluyendo métodos estadísticos clásicos y bayesianos. El enfoque general es el mismo que el adoptado por la guía *Guide to the Expression of Uncertainty in Measurement* (GUM): las cantidades cuyos valores están rodeados de incertidumbre se modelan como variables aleatorias, lo que permite la aplicación de una amplia variedad de técnicas de probabilidad y estadística a la hora de la evaluación de la incertidumbre de medición. Todos los métodos presentados se ilustran con ejemplos que involucran resultados de mediciones reales de una amplia gama de campos de la química y ciencias afines, que van desde la práctica de la química analítica clásica de inicios del siglo XX, hasta estudios contemporáneos de composiciones isotópicas de elementos y ensayos clínicos. El material suplementario ofrece códigos de programación detallados exhaustivamente que permiten a los lectores reproducir todos los cálculos subyacentes a los resultados presentados en los ejemplos.

Palabras clave

Fórmula de Gauss; Incertidumbre de medición; Métodos bayesianos; Métodos de Monte Carlo; Propagación de la incertidumbre.

Índice

1. Introducción	1
1.1. Notación	1
2. Medición	2
3. Modelos de Medición	4
3.1. Modelos de medición determinísticos	4
3.2. Modelos de medición estadística	5
4. Evaluación de la incertidumbre de medición	7
4.1. Introducción	7
4.2. Fórmula de Gauss para la propagación de incertidumbre	7
4.3. Distribuciones de probabilidad	10
4.4. Estimación de máxima verosimilitud	11
4.5. Métodos de Monte Carlo	15
4.6. Métodos no paramétricos	20
4.7. Métodos bayesianos	22
5. Reporte de incertidumbre de medición	27
Referencias	31
Apéndice: Material complementario	38
Ejemplo 3E (Masa molar del iridio — modelo de efectos de laboratorio) .	38
Ejemplo 4B (Isótopos de silicio — Incertidumbre multivariable)	41
Ejemplo 4D (Modelando el peso de las colas)	41
Ejemplo 4E (SiO ₂ en piedra caliza — función de verosimilitud)	42
Ejemplo 4F (Comparación interlaboratorio)	43
Ejemplo 4G (Molnupiravir — Estimación de la eficacia del tratamiento) . .	44
Ejemplo 4H (Estándar de calibración de cadmio — Monte Carlo)	46
Ejemplo 4I (Titulación argentométrica)	47
Ejemplo 4J (Michaelis-Menten — Monte Carlo)	48
Ejemplo 4K (Basaltos del <i>Monte Submarino Verde</i>)	50
Ejemplo 4N (Masa molar del iridio — Métodos de Monte Carlo)	51
Ejemplo 4O (Nitrato en agua de mar — Precisión de medición)	54

Ejemplo 4P (Molnupiravir — Reconsideración escéptica)	56
---	----

Índice de tablas

Tabla 1. Resultados de medición de la fracción de masa de níquel obtenidos en la comparación clave CCQM-K145.	13
Tabla 2. Resultados de medición utilizados para estimar los parámetros del modelo de Michaelis-Menten para la hidrólisis de ONPG.	19

Índice de figuras

Fig. 1. Densidades de probabilidad de v_{MAX} y K_M estimadas por regresión no lineal (NL) y por regresión lineal ponderada (WLB)	20
Fig. 2. Evaluar los resultados de los ensayos clínicos con modelos bayesianos escépticos puede conducir a conclusiones más confiables que los métodos clásicos de análisis estadístico.	27

Agradecimientos

Los autores expresan su profundo agradecimiento a la Dra. Fabienne Meyers por aprobar, en representación de la Unión Internacional de Química Pura y Aplicada (IUPAC), la publicación de esta traducción de un informe técnico de la IUPAC el cual fue publicado originalmente en inglés en *Pure and Applied Chemistry*. También agradecen especialmente a la Dra. Masha Dorogova por conceder dicha aprobación en representación de De Gruyter, la editorial del artículo original. La traducción se publica bajo la misma licencia que aplica al original: *Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International*.

Asimismo, los autores expresan su gratitud a la Profesora Dra. Mariela Pistón (Facultad de Química, Universidad de la República, Montevideo, Uruguay) y al Dr. David Newton (Statistical Engineering Division, NIST) por revisar un primer borrador de esta Nota Técnica y proporcionar valiosas sugerencias para su mejora.

María Victoria Gelabert expresa su agradecimiento a Claire Saundry y Andrew Conn, del NIST, por brindarle la oportunidad de realizar su estancia en el NIST durante mayo de 2024, dónde colaboró con Regina Easley y Antonio Possolo, y trabajó en la elaboración de esta Nota Técnica.

Contribuciones de los autores

Maria Victoria Gelabert: Verificación de contenido, conceptualización, desarrollo de la traducción al español y producción de la versión en LaTeX del documento. **Antonio Possolo, David Brynn Hibbert, Jürgen Stohner, Olha Bodnar, y Juris Meija:** Coautores del artículo original en inglés en el que se basa esta Nota Técnica del NIST.

Convenciones

El término inglés “odds”, cuando se utiliza para comparar las probabilidades de que un evento ocurra o no, y la probabilidad de ocurrencia es p , representa la razón $p/(1 - p)$. Considerando que se utilizan múltiples términos en español para denotar esta cantidad, hemos optado por seguir [1] y usar “posibilidades” como traducción de “odds”. De acuerdo con el anterior, traducimos “odds ratio” como “razón de posibilidades.”

Aunque la coma es el separador decimal más común en español, para conservar la tipografía del original, seguimos el uso del inglés y usamos el punto. Por la misma razón, cuando una unidad de medida implica un cociente y hay una sola unidad en el denominador, usamos “/” en lugar de elevar esta unidad a la -1 , por ejemplo “mg/L”. Sin embargo, cuando el denominador incluye más de una unidad, entonces utilizamos la notación exponencial, por ejemplo en “ $\mu\text{mol L}^{-1} \text{min}^{-1}$ ”.

1. Introducción

La IUPAC ha contribuido al desarrollo de la guía *Guide to the Expression of Uncertainty in Measurement* desde 1997, incluyendo su versión original [2] y documentos relacionados, que en conjunto se conocen como la GUM. Dado que estos documentos hacen un uso extensivo de conceptos y métodos bastante especializados de probabilidad y estadística, muchos usuarios y responsables de evaluaciones de incertidumbre de medición recurren a orientaciones menos técnicas, más accesibles y a ejemplos ilustrativos variados, como los que ofrece la guía EURACHEM *Quantifying Uncertainty in Analytical Measurement* [3] o la guía *Simple Guide for Evaluating and Expressing the Uncertainty of NIST Measurement Results* [4].

La tarea de evaluar la incertidumbre de medición a menudo se reduce a calcular desviaciones estándar, sin un análisis cuidadoso de los modelos de medición que expliquen la interacción entre las contribuciones realizadas por las fuentes relevantes de incertidumbre. Las suposiciones que validan tales modelos y que hacen que las evaluaciones de incertidumbre sean aptas para el propósito a menudo quedan sin examinar. En muchas aplicaciones, las contribuciones relativas de incertidumbre se combinan incorrectamente en una suma cuadrática sin especificar el modelo de medición subyacente, o incluso cuando el mensurando no es un producto de potencias de cantidades de entrada (ver Subsección 4.2).

Esta *Guía Breve* proporciona una reintroducción concisa a los modelos y métodos para la evaluación de la incertidumbre, que enfatiza cómo dicha evaluación depende del propósito para el que se destinan los resultados de medición, e incluye ejemplos prácticos que involucran datos reales que son representativos de la variedad de problemas que seguramente se encuentren en la ciencia y la ingeniería química.

1.1. Notación

Cuando se hace referencia en esta contribución a un conjunto de valores de una o varias cantidades, $\{x_1, \dots, x_n\}$, a menudo representamos por simplicidad a este conjunto como $\{x_i\}$, para distinguir entre el conjunto completo y un solo elemento x_i , de acuerdo con la notación matemática establecida. Los vectores, matrices y conjuntos ordenados más generales se representan con símbolos en negrita. La notación $[x_1 \dots x_m]$ se refiere a un vector fila, que es una matriz con 1 fila y m columnas, mientras que $[x_1 \dots x_m]^T$ denota su transpuesta: un vector columna o una matriz con m filas y 1 columna. De manera similar, si $\mathbf{J}$ denota una matriz con m filas y n columnas, entonces $\mathbf{J}^T$ denota su transpuesta, es decir una matriz con n filas y m columnas, cuyas columnas son las filas de $\mathbf{J}$. Aunque a menudo

se omite por simplicidad, el acento circunflejo (sombrero) significa la estimación de la cantidad. Por ejemplo, la estimación de máxima verosimilitud de una masa, m , se denota como $\hat{m}$. La línea superior, $\overline{m}$, se utiliza para denotar el promedio aritmético (igualmente ponderado) de las réplicas de las determinaciones de m .

El ejemplo 3E introduce las incertidumbres estándar representadas con la notación entre paréntesis, donde la estimación de la masa molar del iridio se muestra como 192.216 62(29) g/mol. Esto indica un valor medido de 192.216 62 g/mol y una incertidumbre estándar asociada de 0.000 29 g/mol. Esta notación entre paréntesis se utiliza extensamente a lo largo de esta *Guía Breve* y de manera más general en muchos campos de la ciencia y la tecnología. El número entre paréntesis denota la incertidumbre estándar, la cual afecta a tantos dígitos finales (menos significativos) del valor medido como dígitos entre los paréntesis, independientemente de la posición del punto decimal, si lo hubiese.

2. Medición

La *medición* es un proceso experimental o computacional que, mediante la comparación con un estándar, produce una estimación del valor verdadero de una propiedad (cuantitativa o cualitativa) de un material, objeto virtual, colección de objetos, proceso o evento o una serie de eventos [5]. Un *resultado de medición* es una estimación de la propiedad que se pretende medir (*mensurando*) junto con una evaluación de la incertidumbre de medición asociada a esta. Las mediciones de algunos tipos de propiedades nominales a menudo se denominan *evaluaciones*, *identificaciones* o *clasificaciones*.

Ejemplo 2A (Mensurandos cuantitativos y cualitativos.) Dado que los mensurandos pueden ser cuantitativos o cualitativos, los resultados de las mediciones correspondientes reflejan esta diferencia. Los mensurandos cuantitativos tienen valores para los cuales las diferencias o los cocientes son significativos. Ninguna de estas operaciones aritméticas es relevante para los mensurandos cualitativos (o categóricos), los cuales pueden ser clasificados según si son ordinales o nominales: los valores de los primeros pueden ser ordenados, de menor a mayor, o de más débil a más fuerte; los valores de los últimos solo pueden ser comparados mediante la equivalencia (si son iguales o diferentes).

CUANTITATIVOS: El certificado de análisis para el material de referencia certificado de agua potable AQUA-1 [6], desarrollado por el National Research Council Canada, indica un valor certificado de 0.222 µg/L para la concentración de masa de arsénico, con una incertidumbre estándar asociada de 0.007 µg/L.

ORDINALES: Mineralab (Eagle, Idaho) produce un kit de prueba de dureza, HD-2, para uso industrial, que incluye ocho puntas que proporcionan valores de 2 a 9 en

la escala de dureza de Mohs [7, §3.5] para minerales y otros productos naturales y manufacturados.

NOMINALES: El National Institute of Standards and Technology (NIST, EE.UU.) tiene plena confianza en la identificación del Material de Referencia 8256, Salmón Coho capturado en estado salvaje, como un compuesto de tejido del *Oncorhynchus kisutch* [8]. La identificación involucró secuenciación de fragmentos de ADN y su comparación con una referencia, identificando secuencias de especímenes de [referencia](#) de esta especie en la biblioteca [Reference Standard Sequence Library](#) de la U.S. Food and Drug Administration (FDA).

La medición tiene como objetivo “*descubrir algo*” [9] y se realiza para respaldar la toma de decisiones [10, 11]. Por ejemplo: la medición de la concentración de troponinas cardíacas específicas en la sangre permite tomar decisiones sobre cómo manejar los infartos de miocardio; la medición del pH del agua en una piscina guía la decisión sobre su mantenimiento; y la medición de la temperatura interna de la carne cocida determina si está lista y es segura para consumir.

La *incertidumbre de medición* es la duda sobre el valor verdadero del mensurando que persiste después de realizar una medición [5, 12]. Puede describirse en su totalidad y cuantitativamente mediante una distribución de probabilidad sobre el conjunto de valores del mensurando, o de manera más concisa para mensurandos escalares, mediante la desviación estándar de dicha distribución (llamada *incertidumbre estándar*), o para mensurandos multivariantes, mediante su [matriz de covarianza](#) (también llamada *matriz de dispersión*).

Si y denota un mensurando escalar, entonces escribimos $u(y)$ para denotar la incertidumbre estándar asociada a y , teniendo en cuenta el hecho que dicha incertidumbre se refiere al valor verdadero del mensurando, no a estimaciones del mismo. En casos donde se están considerando diferentes estimaciones del mismo mensurando, como en el ejemplo 4F, y se necesita hacer referencia a las diferentes evaluaciones de incertidumbre, se pueden denominar, por ejemplo, $u_1(y)$ y $u_2(y)$.

Cabe señalar que el significado de $u(y)$ depende de la distribución de probabilidad que se utilice para modelar la incertidumbre asociada a y . Por ejemplo, la probabilidad de que el intervalo $y \pm u(y)$ cubra el valor real del mensurando es del 68 % cuando y tiene una distribución [gaussiana](#) (o normal), del 58 % cuando y tiene una distribución uniforme (o rectangular), y del 76 % cuando y tiene una distribución de [Laplace](#) (o doble exponencial).

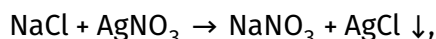
3. Modelos de Medición

Los modelos de medición relacionan el mensurando (salida) con los valores de entrada que son observables durante el experimento, o que han sido medidos previamente [13]. Suponemos en todo momento que tanto las salidas como las entradas son cantidades (es decir, propiedades cuyos valores pueden ser comparados mediante razones o por diferencias). Sin embargo, los modelos de medición generalmente pueden incorporar también entradas cualitativas.

3.1. Modelos de medición determinísticos

Los modelos de medición determinísticos son relaciones funcionales entre los valores verdaderos de salida η y los n valores verdaderos de las cantidades de entrada, $\xi_1, \dots, \xi_n$, expresados como $\eta = f(\xi_1, \dots, \xi_n)$, donde la *función de medición* f está determinada empíricamente, por una teoría física, o por convención. Nótese que la propia definición de f puede estar rodeada de incertidumbre, la cual debe ser identificada y diseminada siempre que haga una contribución no despreciable a la incertidumbre asociada con la estimación de η .

Ejemplo 3A (Modelo teórico) La cantidad de concentración de sustancia de cloruro de sodio, $c(\text{NaCl})$, en una solución acuosa de volumen $V(\text{NaCl})$ se determina mediante titulación con una solución de nitrato de plata en presencia de K_2CrO_4 . La función de la medición se deriva de la teoría química que explica la reacción entre NaCl y AgNO_3 ,



y el punto final de esta titulación se observa por la aparición del cromato de plata marrón-rojizo. Un modelo simple de medición teórico asume que el Ag_2CrO_4 , al ser más soluble que el AgCl , comenzará a formarse solo después de que los iones de cloruro hayan sido consumidos por los iones de plata. Esto conduce a la conocida relación $c(\text{NaCl})V(\text{NaCl}) = c(\text{AgNO}_3)V(\text{AgNO}_3)$, donde $V(\text{AgNO}_3)$ es el volumen de la solución de nitrato de plata que se utilizó para alcanzar el punto final, por lo tanto, el modelo de medición es $c(\text{NaCl}) = c(\text{AgNO}_3)V(\text{AgNO}_3)/V(\text{NaCl})$ [14].

Un modelo de medición físicoquímico más sofisticado considera que el AgCl y el Ag_2CrO_4 son algo solubles, y que el cromato de plata comenzará a precipitar cuando $c^2(\text{Ag}^+) \times c(\text{CrO}_4^{2-}) > K_{\text{sp}}(\text{Ag}_2\text{CrO}_4)$, donde K_{sp} denota el producto de solubilidad. Dependiendo de la concentración de cromato de potasio, esto puede ocurrir ligeramente antes o después de que se alcance el punto de equivalencia [15].

Ejemplo 3B (Modelo empírico) En el análisis espectrométrico de biomoléculas, la señal, A , producida por el espectrómetro de masas generalmente no es directamente proporcional a la concentración de cantidad de sustancia, c , de estas biomoléculas.

Si las razones de estas desviaciones no pueden ser explicadas completamente, a menudo se utilizan funciones matemáticas empíricas adecuadas para modelar la relación aproximada entre la concentración y la señal analítica [16], por ejemplo, la función cuadrática $A = a_0 + a_1c + a_2c^2$.

Ejemplo 3C (Modelo convencional) La dureza total del agua, H , a menudo se define como la concentración de masa de carbonato de calcio equivalente a la suma de las concentraciones de masa de calcio y magnesio disueltos, $H = \gamma(\text{Ca}^{2+}) \times M(\text{CaCO}_3)/M(\text{Ca}) + \gamma(\text{Mg}^{2+}) \times M(\text{CaCO}_3)/M(\text{Mg})$. La función de medición específica que relaciona la cantidad de salida, H , con las cantidades de entrada medidas, $\gamma(\text{Ca}^{2+})$ y $\gamma(\text{Mg}^{2+})$, está definida por convención. Otro ejemplo es la *brecha aniónica*, c_{AG} , definida como la diferencia entre las concentraciones de ciertos cationes y aniones en orina, suero o plasma: $c_{\text{AG}} = c(\text{K}^+) + c(\text{Na}^+) - c(\text{Cl}^-) - c(\text{HCO}_3^-)$ [17].

3.2. Modelos de medición estadística

Los modelos de medición a menudo comprenden componentes determinísticos y estocásticos. Los componentes estocásticos invocan distribuciones de probabilidad para modelar las incertidumbres asociadas con las entradas y salidas. Esto equivale a considerar estas cantidades como *variables aleatorias*, que definimos de manera intuitiva como cantidades que tienen distribuciones de probabilidad como atributos, y donde el adjetivo “aleatorio” no sugiere que el valor verdadero de la cantidad varíe de manera impredecible, simplemente indica que el valor verdadero de la cantidad no se conoce con plena certeza. En consecuencia, la relación entre las cantidades de entrada observadas se reescribe como una regresión con errores en ambas variables [18] $y = f(x_1, \dots, x_n)$, donde f es la función de medición, $y = \eta + \Delta y$ y $x_i = \xi_i + \Delta x_i$, con Δy y Δx_i denotando errores de medición no observables, ξ_i denotando el valor verdadero de x_i , para $i = 1, \dots, n$, y η denotando el valor verdadero de y .

Ejemplo 3D (Modelo de error de medición) Consideremos n determinaciones de la masa, $m_{\text{obs},1}, \dots, m_{\text{obs},n}$, y del volumen, $V_{\text{obs},1}, \dots, V_{\text{obs},n}$, de un líquido, realizadas para estimar su densidad, ρ . El modelo de medición determinista para la densidad es $\rho = m/V$, que relaciona los valores verdaderos (es decir, exentos de errores de medición) de estas cantidades. Los valores observados de la masa y el volumen están relacionados con sus valores verdaderos de la siguiente manera:

$$m_{\text{obs},i} = m + \Delta m_i \quad \text{y} \quad V_{\text{obs},i} = V + \Delta V_i,$$

donde Δm_i y ΔV_i son errores de medición modelados como resultados de variables aleatorias independientes gaussianas, todas con media 0, y desviaciones estándar iguales a sus incertidumbres estándar asociadas reportadas, $u(m_{\text{obs},i})$ y $u(V_{\text{obs},i})$ respectivamente, para $i = 1, \dots, n$.

El componente determinista de un modelo de medición estadístico expresa la magnitud medida, $\eta = g(\theta)$, como una función específica g de un parámetro θ (en la Sección 4.4, por ejemplo, puede visualizarse cómo un vector con varios componentes), que hace referencia a la distribución de probabilidad conjunta de las cantidades de entrada, la cual caracteriza el componente estocástico del modelo. El valor de θ debe ser estimado, típicamente a partir de observaciones replicadas de las cantidades de entrada, para obtener una estimación de la magnitud medida como $\hat{\eta} = g(\hat{\theta})$. Los métodos de momentos, mínimos cuadrados, estimación de máxima verosimilitud (Sección 4.4), y métodos bayesianos (Sección 4.7) son comúnmente utilizados para obtener $\hat{\theta}$ [19]. Los diferentes métodos producen típicamente no solo diferentes estimaciones de θ , sino también diferentes evaluaciones de $u(\theta)$.

Ejemplo 3E (Masa molar del iridio — Modelo de efectos de laboratorio) Prohaska *et al.* [20, §8.4] basaron el cálculo del peso atómico estándar del iridio del año 2021 en tres determinaciones independientes de su relación isotópica, que corresponden a las siguientes estimaciones de masa molar:

$$M_{\text{obs},1} = 192,216\,77(81) \text{ g/mol}$$

$$M_{\text{obs},2} = 192,216\,62(29) \text{ g/mol}$$

$$M_{\text{obs},3} = 192,217\,63(17) \text{ g/mol}$$

La desviación estándar de estas estimaciones de masa molar es 1.6 veces mayor que el promedio geométrico de las incertidumbres estándar asociadas, lo que sugiere que los tres resultados de medición son mutuamente inconsistentes. La prueba Q de Cochran [21], la cual arroja un [valor \$p\$](#) de 0.0082, confirma la relevancia de la inconsistencia. Un modelo de medición estadístico que considera efectos adicionales de laboratorio, que no están incluidos en las incertidumbres de medición individuales, puede ajustar estos tres resultados mutuamente inconsistentes:

$$M_{\text{obs},1} = M + \lambda_1 + \Delta M_1, \quad M_{\text{obs},2} = M + \lambda_2 + \Delta M_2, \quad \text{y} \quad M_{\text{obs},3} = M + \lambda_3 + \Delta M_3,$$

donde M denota el valor verdadero de la masa molar, λ_1 , λ_2 y λ_3 denotan efectos de laboratorio (asumidos como resultados de variables aleatorias con la misma distribución [gaussiana](#) con media 0 g/mol y desviación estándar τ), y ΔM_1 , ΔM_2 y ΔM_3 denotan errores de medición, asumidos como resultados independientes de tres variables aleatorias [gaussianas](#) todas con la misma media 0 g/mol, y desviaciones estándar iguales a las incertidumbres estándar $u(M_{\text{obs},1})$, $u(M_{\text{obs},2})$ y $u(M_{\text{obs},3})$. Estas incertidumbres de medición son confiables, es decir, se asume que las mismas están basadas en un gran número de grados de libertad.

Este modelo de medición estadística tiene dos parámetros de interés primario: M y τ . Una evaluación básica de τ se puede obtener de manera similar a cómo se estiman

a menudo los componentes de varianza en el análisis de varianza [22, capítulo 5]: basada en la diferencia entre la desviación estándar de los valores medidos de M y el promedio geométrico de las incertidumbres estándar asociadas con los valores medidos: $\tau(\text{g/mol}) = \sqrt{0,000\ 54^2 - 0,000\ 34^2} = 0,000\ 42$. Dado que esta evaluación se basa solo en tres observaciones, su confiabilidad es cuestionable y puede aumentarse ya sea teniendo en cuenta mediciones independientes adicionales de M , o mediante la incorporación de conocimientos relevantes y preexistentes sobre τ en forma de una distribución de probabilidad bayesiana informativa para esta desviación estándar.

4. Evaluación de la incertidumbre de medición

4.1. Introducción

Las evaluaciones *Tipo A* implican la aplicación de métodos estadísticos a datos experimentales, consistentes con un modelo de medición. Las evaluaciones *Tipo B* implican la obtención de conocimiento especializado y su emplazamiento en distribuciones de probabilidad que describen estados de conocimiento sobre los valores verdaderos de las entradas. Este conocimiento puede provenir de fuentes legítimas como certificados de calibración o publicaciones técnicas, o de expertos en la materia.

Las tipologías mencionadas previamente se refieren a formas de evaluar la incertidumbre, no a tipos de error de medición. Por ejemplo, la incertidumbre resultante de la presencia de un sesgo desconocido (atribuible a un error persistente o sistemático) puede ser evaluada en algunos casos mediante métodos Tipo A (por ejemplo, a través de una *validación cruzada* [23, Cap. 8]), o mediante métodos Tipo B en otros casos (por ejemplo, generados por un experto en la materia).

Las evaluaciones de incertidumbre *de abajo hacia arriba* implican la enumeración de todas las fuentes relevantes de incertidumbre, cuyas contribuciones luego se resumen en un *presupuesto de incertidumbre*, a menudo complementado por un *diagrama de causa y efecto* [3, Apéndice D.3] que describe cómo influyen en el resultado. Las evaluaciones de incertidumbre *de arriba hacia abajo* implican comparaciones entre valores medidos obtenidos de manera independiente entre sí, y no implican la identificación y caracterización de las fuentes de incertidumbre, como se muestra en los ejemplos 3E y 4F.

4.2. Fórmula de Gauss para la propagación de incertidumbre

Si la función de medición f en el modelo de medición $y = f(x_1, \dots, x_n)$ es derivable, entonces la incertidumbre estándar asociada a y puede ser evaluada de manera

aproximada utilizando la siguiente generalización (Ec. (13) de [2]) de la fórmula de Gauss [24, §18] para la incertidumbre estándar al cuadrado,

$$u^2(y) \approx \sum_{i=1}^n c_i^2 u^2(x_i) + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^n c_i c_j u(x_i) u(x_j) r_{ij}, \quad (1)$$

donde $u(x_1), \dots, u(x_n)$ son las incertidumbres estándar asociadas a las estimaciones de las cantidades de entrada, los $\{r_{ij}\}$ son los **coeficientes de correlación** (producto-momento de Pearson) entre ellas, y los *coeficientes de sensibilidad* $c_1, \dots, c_n$ son los valores de las derivadas parciales de primer orden de f con respecto a las cantidades de entrada, evaluadas en $(x_1, \dots, x_n)$, es decir, $c_i = \partial f(x_1, \dots, x_n) / \partial x_i$, para $i = 1, \dots, n$ [25]. Cuando las cantidades de entrada no están correlacionadas, la Ec. (1) se reduce a

$$u^2(y) \approx \sum_{i=1}^n c_i^2 u^2(x_i). \quad (2)$$

Adicionalmente, cuando la cantidad de salida, y , es una combinación lineal de las cantidades de entrada, es decir, $y = \sum_{i=1}^n a_i x_i$, la Ec. (1) es exacta.

Ejemplo 4A (Dureza del agua — Fórmula de Gauss) En el caso de la dureza total del agua en el ejemplo 3C, los coeficientes de sensibilidad son $c_1 = \partial H / \partial \gamma(\text{Ca}^{2+}) = M(\text{CaCO}_3) / M(\text{Ca}) \approx 2,5$ y $c_2 = \partial H / \partial \gamma(\text{Mg}^{2+}) = M(\text{CaCO}_3) / M(\text{Mg}) \approx 4,1$, y

$$u^2(H) = c_1^2 u^2(\gamma(\text{Ca}^{2+})) + c_2^2 u^2(\gamma(\text{Mg}^{2+}))$$

cuando ignoramos cualquier correlación que pueda existir entre las estimaciones de $\gamma(\text{Ca}^{2+})$ y $\gamma(\text{Mg}^{2+})$. Esta situación se presenta cuando el calcio y el magnesio se miden por separado mediante espectroscopía de absorción atómica. Por lo tanto, para un agua con $\gamma(\text{Ca}^{2+}) = 23,2(8)$ mg/L y $\gamma(\text{Mg}^{2+}) = 9,4(6)$ mg/L, determinando ambos parámetros independientemente uno del otro, tenemos que

$$u^2(H) = (2,5)^2 \times (0,8 \text{ mg/L})^2 + (4,1)^2 \times (0,6 \text{ mg/L})^2 \approx (3 \text{ mg/L})^2.$$

Cuando la cantidad de salida es un producto o un cociente de potencias de cantidades de entrada no correlacionadas, $y = \kappa x_1^{\alpha_1} \dots x_n^{\alpha_n}$, donde κ y $\alpha_1, \dots, \alpha_n$ son constantes (positivas o negativas), la fórmula de Gauss adquiere una forma sencilla de recordar que solo implica las incertidumbres estándar relativas: $[u(y)/y]^2 \approx [\alpha_1 u(x_1)/x_1]^2 + \dots + [\alpha_n u(x_n)/x_n]^2$. Como en el caso del ejemplo 3D, $[u(\rho)/\rho]^2 \approx [u(m_{\text{obs}})/m_{\text{obs}}]^2 + [u(V_{\text{obs}})/V_{\text{obs}}]^2$.

En algunas aplicaciones, la cantidad de salida de la función de medición es un vector con $m > 1$ componentes, que puede escribirse como un vector columna

cuyos elementos son los valores de m funciones de valores verdaderos de los n parámetros de entrada: $\mathbf{f}(\mathbf{x}) = [f_1(\mathbf{x}) \dots f_m(\mathbf{x})]^T$. Por lo tanto, propagar la incertidumbre asociada con las entradas requiere la siguiente versión de la fórmula de Gauss que produce una aproximación de la matriz de covarianza de tamaño $m \times m$ del vector de salida:

$$\Sigma_{\mathbf{y}} \approx \mathbf{J} \Sigma_{\mathbf{x}} \mathbf{J}^T \quad (3)$$

donde $\mathbf{x} = [x_1 \dots x_n]^T$ denota el vector columna de n entradas, $\mathbf{y}$ denota el vector columna de m salidas, y la matriz jacobiana $\mathbf{J}$ de tamaño $m \times n$ tiene los valores de las derivadas parciales $\partial f_i(\mathbf{x}) / \partial x_j$, siendo las filas $i = 1, \dots, m$ y las columnas $j = 1, \dots, n$. Los elementos diagonales de $\Sigma_{\mathbf{y}}$ son las incertidumbres estándar de las salidas al cuadrado, y los elementos fuera de la diagonal son las covarianzas entre estas [26].

Ejemplo 4B (Isótopos de silicio — Propagación de incertidumbre multivariable) La medición de dos de las relaciones isotópicas de silicio, $R_{29/28}$ y $R_{30/28}$, permite la estimación simultánea de tres abundancias isotópicas que suman 1 mol/mol:

$$x_{28} = \frac{1}{1 + R_{29/28} + R_{30/28}}, \quad x_{29} = \frac{R_{29/28}}{1 + R_{29/28} + R_{30/28}}, \quad \text{y} \quad x_{30} = \frac{R_{30/28}}{1 + R_{29/28} + R_{30/28}}.$$

Por lo tanto, tenemos una función de medición de dos parámetros de entrada que corresponden a tres parámetros de salida, con la siguiente matriz jacobiana de tamaño 3×2 [27]:

$$\mathbf{J} = \begin{bmatrix} \frac{\partial x_{28}}{\partial R_{29/28}} & \frac{\partial x_{28}}{\partial R_{30/28}} \\ \frac{\partial x_{29}}{\partial R_{29/28}} & \frac{\partial x_{29}}{\partial R_{30/28}} \\ \frac{\partial x_{30}}{\partial R_{29/28}} & \frac{\partial x_{30}}{\partial R_{30/28}} \end{bmatrix} = x_{28}^2 \begin{bmatrix} -1 & -1 \\ 1 + R_{30/28} & -R_{29/28} \\ -R_{30/28} & 1 + R_{29/28} \end{bmatrix}$$

La [matriz de covarianza](#) de los parámetros de entrada es

$$\Sigma_{\mathbf{R}} = \begin{bmatrix} u^2(R_{29/28}) & ru(R_{29/28})u(R_{30/28}) \\ ru(R_{29/28})u(R_{30/28}) & u^2(R_{30/28}) \end{bmatrix},$$

la cual incluye el [coeficiente de correlación](#), r , entre las estimaciones de los dos parámetros de entrada.

Valkiers *et al.* [28, Tabla 4] enumeran los siguientes resultados de medición de la relación isotópica de silicio para su Mezcla III, $R_{29/28} = 0,050\,0657(25)$ mol/mol y $R_{30/28} = 0,032\,9035(34)$ mol/mol, pero no proporcionan una estimación de la correlación entre ellos. Suponiendo que la misma es cero, y aplicando las fórmulas anteriores, obtenemos $x_{28} = 0,923\,3873(36)$ mol/mol, $x_{29} = 0,046\,2300(22)$ mol/mol, y $x_{30} = 0,030\,3827(30)$ mol/mol. Cuando hay más de dos relaciones isotópicas, la incertidumbre derivada de la falta de conocimiento de las correlaciones entre ellas puede ser modelada utilizando la [distribución de LKJ](#) [29].

4.3. Distribuciones de probabilidad

La incertidumbre de medición puede expresarse de manera más exhaustiva mediante una distribución de probabilidad. Esta expresión se aplica de igual manera a la determinación de propiedades cuantitativas como cualitativas (los ejemplos 4G, 4L y 4L ilustran esto último). Existen muchas familias de distribuciones de probabilidad, como por ejemplo la [distribución normal](#) (también llamada *gaussiana*), la [distribución de Laplace](#), la [distribución triangular](#), la [distribución uniforme](#) (también llamada *rectangular*), la [distribución gamma](#), la [distribución beta](#), etc. Los miembros de la misma familia tienen [funciones de densidad de probabilidad](#) con la misma forma funcional, diferenciándose solo en los valores de sus parámetros. Por ejemplo, una distribución [gaussiana](#) se identifica por su media (también conocida como [valor esperado](#)) y su [desviación estándar](#), mientras que las distribuciones [gamma](#) se identifican por los valores de los parámetros que determinan su forma y su escala.

Aunque la distribución [gaussiana](#) recibe la mayor parte de la atención, el modelado probabilístico de resultados de medición a menudo requiere elegir otras distribuciones. Esta elección debe estar guiada por cuán adecuadamente diferentes distribuciones aglomeran y transmiten, ya sea un nivel de conocimiento sobre un parámetro de entrada de un modelo de medición o la variabilidad de muestreo de una entrada, como se ilustra a continuación y en muchos ejemplos proporcionados por Possolo and Meija [12].

Varios tipos de recuento se modelan utilizando distribuciones discretas como la [distribución de Poisson](#), la [distribución binomial negativa](#) o la [distribución multinomial](#). Esto puede incluir el número de pacientes enfermos en un ensayo clínico, colonias de *E. coli* en una placa de agar o centelleos detectados por un contador de centelleo.

Para los parámetros de entrada que toman valores en un continuo (por ejemplo, una concentración de cantidad medida, cuya magnitud puede ser cualquier número real no negativo), hay una amplia variedad de las llamadas distribuciones de probabilidad continua disponibles para su modelado. Algunos de los criterios de selección en este caso incluyen: (i) el rango de valores que la cantidad de entrada posiblemente puede tomar (por ejemplo, si los valores solo pueden ser positivos o deben estar dentro de un intervalo acotado); (ii) si los valores mayores al centro de la distribución son tan probables como los valores menores a él (simetría); y (iii) qué tan probable es que se esperen valores alejados del centro de la distribución (cola pesada).

No siempre es trivial decidir *a priori* si una cantidad de entrada particular debe ser modelada utilizando una distribución simétrica o asimétrica, o qué tan pesadas deben ser las colas de la distribución. Algunas distribuciones tienen paráme-

tros cuyos valores permiten una elección que responde a características específicas de los datos de medición: por ejemplo, si hay asimetría hacia la izquierda o hacia la derecha.

Ejemplo 4C (Modelando asimetría) Louis Ahrens [30] propone la distribución [log-normal](#), cuya cola derecha es más larga o más pesada que la cola izquierda (es decir, exhibe asimetría positiva), como un modelo para describir la variabilidad natural de la fracción de masa de elementos menores en rocas, por ejemplo, el escandio en muestras de granitos canadienses. Una convolución de las distribuciones [normal](#) y [exponencial](#) (también llamada distribución [gaussiana modificada exponencialmente](#)) se utiliza para modelar la asimetría positiva que comúnmente muestran los picos cromatográficos. La [distribución beta](#) puede modelar la asimetría negativa en la pureza observada en materiales de alta pureza. Tanto la distribución [normal sesgada](#) como la distribución [t-sesgada](#) [31] pueden modelar un sesgo positivo así como negativo.

Ejemplo 4D (Modelando el peso de las colas) Tanto la distribución de [Laplace](#) como la distribución [t de Student](#) pueden modelar diferentes grados del peso de la cola (cuanto menor sea el número de grados de libertad de una distribución [t de Student](#), más pesadas serán sus colas). Por ejemplo, la probabilidad de que una variable aleatoria de [Laplace](#) tome valores de más de 3 desviaciones estándar alejados de la media es 0.014, mientras que la probabilidad correspondiente para las distribuciones [gaussianas](#) es solo de 0.003. Para la distribución [t de Student](#) con 3 grados de libertad, la probabilidad del mismo evento es 0.033. Entre las distribuciones discretas, la distribución [binomial negativa](#) se puede utilizar para modelar estadísticamente cantidades que están más dispersas de lo que permite la [distribución de Poisson](#), como se ilustra por Possolo y Meija en la Ref. [12, páginas 41-45].

4.4. Estimación de máxima verosimilitud

Con el fin de obtener una evaluación de incertidumbre, modelamos los resultados de las mediciones como valores observados de variables aleatorias con distribuciones de probabilidad especificadas. Por lo tanto, tenemos a nuestra disposición todos los métodos aplicables de inferencia estadística que pueden producir no solo estimaciones de las cantidades de interés, sino también evaluaciones de la incertidumbre asociadas a ellas. La función de verosimilitud (definida a continuación) juega un papel destacado en muchos de estos métodos, tanto clásicos como bayesianos.

Supongamos que las entradas $x_1, \dots, x_n$ (introducidas en 3) pueden considerarse como resultados de variables aleatorias independientes cuyas distribuciones de probabilidad involucran parámetros $\theta_1, \dots, \theta_n$, y cuyas funciones de densidad de probabilidad (o funciones de masa de probabilidad) son $p_1, \dots, p_n$. Si denotamos como $\mathbf{x}$ al conjunto de cantidades de entrada, la función de verosimilitud L es

$$L_{\mathbf{x}}(\theta_1, \dots, \theta_n) = p_1(x_1 | \theta_1) \dots p_n(x_n | \theta_n).$$

En la práctica, a menudo se omite el subíndice de L indicativo de los datos, dando por entendido que L es una función de los parámetros de las distribuciones de las cantidades de entrada, mientras que las cantidades de entrada permanecen fijas en sus valores medidos. Cuando $p_1, \dots, p_n$ son funciones de masa de probabilidad, $L_{\mathbf{x}}(\theta_1, \dots, \theta_n)$ es la probabilidad de los datos $\mathbf{x}$ correspondientes a los valores de los parámetros $\theta_1, \dots, \theta_n$. En general, la función L alcanza sus valores más altos para combinaciones de valores de sus argumentos, $\{\theta_j\}$, que hacen que los datos sean “más probables”.

Ejemplo 4E (SiO₂ en piedra caliza — Función de verosimilitud) El certificado del Material de Referencia Certificado 1 del NIST, Caliza Arcillosa, que fue publicado el 1 de julio de 1910 [32], enumera dos determinaciones de la fracción de masa de SiO₂, 0.1811 g/g y 0.1818 g/g, realizadas por W.F. Hillebrand (U.S. Geological Survey) y por C.E. Waters (U.S. National Bureau of Standards) respectivamente, quienes emplearon básicamente el mismo método de análisis en medio líquido [33]. Si modelamos estos dos valores como resultados de variables aleatorias gaussianas con la misma media μ y desviación estándar σ , entonces la función de verosimilitud es el producto de los valores de dos funciones de densidad de probabilidad gaussianas:

$$L(\mu, \sigma) = \frac{\exp\left\{-\frac{(0.1811 \text{ g/g} - \mu)^2}{2\sigma^2}\right\}}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(0.1818 \text{ g/g} - \mu)^2}{2\sigma^2}\right\}}{\sqrt{2\pi\sigma^2}}.$$

Esta función alcanza un máximo cuando μ es igual a $\hat{\mu} = (0.1811 \text{ g/g} + 0.1818 \text{ g/g})/2 = 0.18145 \text{ g/g}$ y σ^2 es igual a $\hat{\sigma}^2 = ((0.1811 \text{ g/g} - \hat{\mu})^2 + (0.1818 \text{ g/g} - \hat{\mu})^2)/2 = (0.00035 \text{ g/g})^2$.

La *estimación de máxima verosimilitud* (EMV) del vector de parámetros $\boldsymbol{\theta} = [\theta_1, \dots, \theta_n]$, simbolizado como $\hat{\boldsymbol{\theta}}$, es un valor de $\boldsymbol{\theta}$ que maximiza la función de verosimilitud L . (Cabe destacar que, en algunos casos, la función de verosimilitud tiene múltiples máximos locales y $\hat{\boldsymbol{\theta}}$ puede no ser único). Por razones de eficiencia computacional y estabilidad, es ventajoso maximizar $\ell = \ln(L)$, lo cual en la mayoría de las aplicaciones requiere optimización numérica. En términos generales, $\hat{\boldsymbol{\theta}}$ tiene propiedades notables: la distribución de probabilidad de $\hat{\boldsymbol{\theta}}$ es aproximadamente *gausiana multivariante* con una media $\boldsymbol{\theta}$ y *matriz de covarianza* $\Sigma_{\hat{\boldsymbol{\theta}}} \approx -\mathbf{H}^{-1}(\hat{\boldsymbol{\theta}})$, donde $\mathbf{H}$ (conocida como *matriz hessiana*) es la matriz de las segundas derivadas parciales de ℓ evaluadas en $\hat{\boldsymbol{\theta}}$, y la calidad de la aproximación aumenta con la cantidad de datos en los que se basa $\hat{\boldsymbol{\theta}}$ (teorema 5.1 en [34]).

Ejemplo 4F (Comparación interlaboratorio — Estimación de máxima verosimilitud) Consideremos la comparación interlaboratorio CCQM-K145 con mediciones de frac-

ción de masa de níquel en hígado bovino proporcionadas por 17 laboratorios expertos [35]. Estos resultados, junto con sus incertidumbres estándar de medición asociadas, se resumen en la siguiente tabla.

Laboratorio	w(Ni)/ (mg/kg)	Laboratorio	w(Ni)/ (mg/kg)	Laboratorio	w(Ni)/ (mg/kg)
JSI	1.930(110)	INACAL	2.010(60)	PTB	2.077(35)
INMC	1.940(64)	NMIA	2.020(50)	LATU	2.080(59)
GLHK	1.942(92)	NIM	2.022(23)	LGC	2.131(42)
LNE	1.958(75)	NMIJ	2.050(20)	UME	2.150(30)
NIST	1.984(20)	RISE	2.055(52)	HSA	2.180(80)
KRISS	1.993(33)	NRC	2.070(50)		

Tabla 1. Resultados de medición de la fracción de masa de níquel obtenidos en la comparación clave CCQM-K145.

Se puede utilizar un modelo de efectos de laboratorio para estos datos, que tenga la misma forma general que el ejemplo 3E, con efectos de laboratorio gaussianos de una media de 0 mg/kg y una desviación estándar desconocida τ . Este τ cuantifica una fuente de incertidumbre que a menudo se llama *incertidumbre oscura* [36], porque su presencia solo se descubre una vez que los resultados de las mediciones obtenidos independientemente uno de otro se comparan entre sí y se encuentran más dispersos de lo que sugieren las incertidumbres reportadas por los laboratorios participantes. Los errores de medición se modelan como variables aleatorias gaussianas con media 0 mg/kg y desviaciones estándar iguales a las incertidumbres informadas, tal cómo fueron reportadas (es decir, se supone que estas se basan en un gran número de grados de libertad). La elección de este modelo estadístico también está respaldada por el *NIST Decision Tree* [37].

Bajo la hipótesis simplificada de que las incertidumbres estándar asociadas con los valores medidos son conocidas, la función de verosimilitud tiene dos argumentos escalares: el valor medio general de la fracción molar de níquel en el material, μ , y la incertidumbre oscura, τ . En este caso, la función de verosimilitud es un producto de 17 funciones de densidad de probabilidad gaussianas, cada una relacionada con el resultado de un participante, similar al ejemplo 4E. Por ejemplo, el término correspondiente al laboratorio JSI es

$$\frac{\exp \left\{ -\frac{(1,930 \text{ mg/kg} - \mu)^2}{2(\tau^2 + (0,110 \text{ mg/kg})^2)} \right\}}{\sqrt{2\pi(\tau^2 + (0,110 \text{ mg/kg})^2)}}.$$

Esto indica que, basándose únicamente en el resultado de la medición de JSI, el valor más probable para μ es el valor 1.930 mg/kg, medido por JSI, y que la incertidumbre asociada con dicho valor debido a la presencia de la incertidumbre oscura τ es $\sqrt{\tau^2 + (0,110 \text{ mg/kg})^2}$.

Como se señaló anteriormente, es preferible maximizar el logaritmo de la función de verosimilitud, por lo que en este caso, se requiere del uso de un método de optimización numérica. Las estimaciones de máxima verosimilitud son $\hat{\mu} = 2,043(15)$ mg/kg y $\hat{\tau} = 0,044(14)$ mg/kg. Las incertidumbres estándar asociadas a estas estimaciones se basan en el hessiano de la función de log-verosimilitud, descritas anteriormente, a pesar de que solo tenemos 17 resultados de medición. El método estadístico del “[bootstrap](#)” paramétrico [38], presentado en el ejemplo 4J, puede proporcionar una evaluación alternativa, que generalmente es una evaluación de incertidumbre confiable para las estimaciones de máxima verosimilitud.

Este ejemplo ilustra cómo el método de máxima verosimilitud puede ser utilizado para combinar medidas independientes de la misma cantidad en un valor de consenso, y cómo la incertidumbre estándar asociada con este resultado puede ser evaluada, en particular cuando las mediciones que se están combinando son mutuamente inconsistentes.

Sin embargo, el método de máxima verosimilitud tiende a producir estimaciones sesgadas de los parámetros de interés [39, §3], y las evaluaciones aproximadas de la incertidumbre utilizando la matriz hessiana de la función log-verosimilitud pueden subestimar seriamente la verdadera varianza de las estimaciones [40, §7], en particular cuando el número de resultados de medición disponibles es pequeño (menos de 15). En general, la estimación de la máxima verosimilitud restringida [41], por ejemplo como se implementa en el paquete R *metafor* [42], o en los procedimientos bayesianos implementados en el [NIST Decision Tree](#) [37] o en el [NIST Consensus Builder](#) [43] son preferibles.

Ejemplo 4G (Molnupiravir — Estimación de máxima verosimilitud de la eficacia del tratamiento) Bernal *et al.* (2022) [44] informaron los siguientes resultados finales del ensayo clínico MOVE-OUT sobre la eficacia de molnupiravir contra COVID-19, para todos los pacientes que habían sido asignados aleatoriamente al grupo de tratamiento (800 mg de molnupiravir dos veces al día durante 5 días) y al grupo de control (placebo): 48 de los $n_1 = 709$ pacientes en el grupo de tratamiento, y 68 de los $n_0 = 699$ pacientes en el grupo de control, habían sido hospitalizados o fallecieron dentro de los 29 días posteriores a la participación en el estudio.

A pesar de que el COVID-19 es una enfermedad infecciosa, modelaremos estas observaciones como resultados de distribuciones binomiales cuyo parámetro p (la probabilidad de hospitalización o muerte) depende de si los pacientes estaban tomando molnupiravir (p_1) o placebo (p_0). Las correspondientes posibilidades de hospitalización o muerte en los dos grupos son $O_1 = p_1/(1 - p_1)$ y $O_0 = p_0/(1 - p_0)$. Dado que las estimaciones de máxima verosimilitud de p_1 y p_0 son $\hat{p}_1 = 48/709 = 0,0677$ y $\hat{p}_0 = 68/699 = 0,0973$, la [razón de posibilidades](#) es $O_1/O_0 = 0,674$. Una razón de posibilidades de 1 significa que el fármaco es tan efectivo como el placebo, y cuanto más pequeño sea esta razón, más eficaz parecerá ser el tratamiento.

Para determinar si la razón de posibilidades es significativamente menor que 1, evaluaremos la incertidumbre del logaritmo de la razón de posibilidades, $\ln(O_1/O_0) = -0,395$, lo cual tiene la ventaja de tener una distribución de probabilidad que es aproximadamente [gaussiana](#) [45]. Dado que el diseño del estudio garantiza que $\hat{p}_1$ y $\hat{p}_0$ no están correlacionados, tenemos:

$$u^2(\ln(O_1/O_0)) = u^2(\ln O_1) + u^2(\ln O_0).$$

Aplicando la fórmula de Gauss a cada término en el lado derecho de esta ecuación, obtenemos:

$$u^2(\ln(O_1/O_0)) \approx \frac{1}{n_1 p_1 (1 - p_1)} + \frac{1}{n_0 p_0 (1 - p_0)} = (0,197)^2.$$

La probabilidad de observar un logaritmo de la razón de posibilidades de $-0,395$ o menor (lo que indica que molnupiravir es al menos tan efectivo como lo sugieren los resultados del ensayo), por simple casualidad y en ausencia de un efecto beneficioso significativo, es aproximadamente igual a la probabilidad de que una variable aleatoria [gaussiana](#) con una media 0 y desviación estándar 1 sea menor que $-0,395/0,197 = -2,01$. Esta probabilidad es solo un 2.2 %. Si el umbral de significancia estadística se establece en un 5 %, entonces los resultados de este ensayo sugieren que molnupiravir es eficaz. Sin embargo, un estudio clínico posterior y mucho más grande [46] estableció que “molnupiravir no redujo la frecuencia de hospitalizaciones asociadas con el COVID-19 o muertes entre adultos vacunados de alto riesgo en la comunidad”.

4.5. Métodos de Monte Carlo

Los métodos de simulación de Monte Carlo se utilizan a menudo para la propagación de la incertidumbre. Ofrecen varias ventajas importantes sobre el enfoque descrito en la Sección 4.2: (i) no requieren el cálculo de derivadas, ya sea analítica o numéricamente; (ii) son aplicables incluso cuando los modelos de medición involucran relaciones notablemente no lineales entre las cantidades de entrada y salida; y (iii) proporcionan los elementos necesarios para caracterizar completamente la distribución de probabilidad de la cantidad de salida.

Los métodos de Monte Carlo implican muestrear de las distribuciones de probabilidad adecuadas un número especificado K de veces. El valor de K debería ser suficientemente grande para asegurar que las evaluaciones de incertidumbre resultantes (incertidumbres estándar, intervalos de cobertura, etc.) alcancen el número necesario de cifras significativas de manera reproducible.

El método de Monte Carlo utilizado para evaluar la incertidumbre asociada a un mensurando en un modelo de medición convencional, $y = f(x_1, \dots, x_n)$, implica los siguientes pasos:

- (a) Especificar una distribución de probabilidad conjunta para las cantidades de entrada.
- (b) Elegir un entero positivo, K , lo suficientemente grande (según lo mencionado anteriormente), y extraer una muestra de tamaño K de la distribución conjunta de las cantidades de entrada para obtener $(x_{1,1}, \dots, x_{n,1}), \dots, (x_{1,K}, \dots, x_{n,K})$.
- (c) Calcular $y_1 = f(x_{1,1}, \dots, x_{n,1}), \dots, y_K = f(x_{1,K}, \dots, x_{n,K})$, que son una muestra de la distribución de probabilidad de la cantidad de salida.

El método de Monte Carlo produce una muestra grande, $y_1, \dots, y_K$, extraída de la distribución de probabilidad de la cantidad de salida, lo que hace posible caracterizar completamente la incertidumbre asociada a y . Esta muestra se puede resumir de varias maneras diferentes, dependiendo del propósito. Por ejemplo, para un mensurando escalar, un *intervalo de cobertura* cuyos extremos son los percentiles 2.5 y 97.5 de la muestra, $y_1, \dots, y_K$, captura un valor “típico” de la cantidad de salida con una probabilidad del 95 %. En general, dicho intervalo no es ni un intervalo de confianza clásico del 95 % ni un intervalo de credibilidad bayesiano del 95 % para la media de la distribución de y . En cambio, es un intervalo de predicción para el valor de la cantidad de salida [19, 47].

Ejemplo 4H (Estándar de calibración de cadmio — Propagación de incertidumbre mediante Monte Carlo) El ejemplo A1 de la Guía CG-4 de EURACHEM/CITAC [3] describe la preparación de un estándar de calibración para espectroscopía de absorción atómica, donde una masa m_{Cd} de cadmio con pureza w_{Cd} se disuelve en ácido para obtener una solución de volumen V . El modelo de medición expresa la concentración másica de cadmio como $\gamma_{\text{Cd}} = w_{\text{Cd}} m_{\text{Cd}} / V$. Las cantidades de entrada tienen los siguientes valores medidos e incertidumbres estándar: $m_{\text{Cd}} = 100,28 \text{ mg}$, $u(m_{\text{Cd}}) = 0,05 \text{ mg}$; $w_{\text{Cd}} = 0,999\,900 \text{ g/g}$, $u(w_{\text{Cd}}) = 0,000\,058 \text{ g/g}$; $V = 100,00 \text{ mL}$, $u(V) = 0,07 \text{ mL}$. La aplicación de la fórmula de Gauss arroja una incertidumbre estándar de $u(\gamma_{\text{Cd}}) = 0,000\,9 \text{ mg/mL}$.

Ahora, con el propósito de aplicar el método de Monte Carlo para la propagación de distribuciones, supongamos que m_{Cd} , w_{Cd} y V se modelan como variables aleatorias mutuamente independientes con medias iguales a sus valores medidos y desviaciones estándar iguales a sus incertidumbres estándar, y con una distribución *gaussiana* truncada en 0 mg para la masa (m_{Cd}), una distribución *beta* para la pureza (w_{Cd}), y una distribución triangular simétrica para el volumen (V).

Dado un valor extraído de cada una de estas tres distribuciones, podemos calcular un valor para γ_{Cd} utilizando el modelo de medición indicado anteriormente. Hacer esto 10^7 veces produce una muestra de Monte Carlo de tamaño $K = 10^7$ extraída de la distribución de probabilidad de la concentración de cadmio. La media y la desviación estándar de esta muestra son $\gamma_{\text{Cd}} = 1,0027 \text{ mg/mL}$ y $u(\gamma_{\text{Cd}}) = 0,000\,9 \text{ mg/mL}$. Esta evaluación de la incertidumbre estándar coincide con la evaluación aproxima-

da utilizando la fórmula de Gauss por una cifra significativa. Sin embargo, el examen de la muestra de Monte Carlo también muestra que la distribución de probabilidad de γ_{Cd} es unimodal y aproximadamente simétrica, pero no [gaussiana](#). Por lo tanto, evaluamos la incertidumbre expandida para una cobertura del 95 % asociada a γ_{Cd} como la mitad de la diferencia entre los percentiles de 97.5 y 2.5 de la muestra de Monte Carlo, obteniendo 0.0017 mg/mL.

El método estadístico del “[bootstrap](#)” paramétrico [38] es un método de Monte Carlo adecuado para evaluar la incertidumbre de medida para mensurandos que son funciones de parámetros de modelos estadísticos de medición. La idea es estimar los parámetros de una distribución de probabilidad (por ejemplo, utilizando mínimos cuadrados, el método de momentos o máxima verosimilitud) y luego proceder a muestrear de esta distribución y usar la dispersión de los resultados para caracterizar la incertidumbre de las funciones relevantes de los parámetros.

Por ejemplo, a menudo se utiliza una función lineal de calibración para expresar la relación entre los valores de la concentración de masa de un analito en una solución y las lecturas instrumentales, véase la página 98 de la Ref. [12]. La pendiente y la ordenada en el origen, que pueden ser estimadas mediante el método de mínimos cuadrados, son parámetros de la distribución de probabilidad de las lecturas instrumentales. Sin embargo, también determinan el valor del mensurando que corresponde a cualquier valor particular de la lectura instrumental. En este caso, el “bootstrap” paramétrico (llamado “propagación de distribuciones” o “método de Monte Carlo” de acuerdo a Ref. JCGM 101:2008) extrae muestras de la distribución de probabilidad conjunta de la pendiente y la ordenada en el origen para evaluar la incertidumbre asociada con las estimaciones del mensurando. Los siguientes ejemplos [4I](#) y [4J](#) ilustran el uso del “bootstrap” paramétrico para la evaluación de la incertidumbre.

Ejemplo 4I (Titulación argentométrica — Propagación de la incertidumbre) Consideremos la medición de la concentración molar de cloruro de sodio en solución acuosa mediante titulación con nitrato de plata y utilizando cromato de potasio como indicador del punto final (método de Mohr), que fue introducido en el ejemplo [3A](#), según [49]. El volumen nominal de la solución de cloruro de sodio es $V(\text{NaCl}) = 100 \text{ mL}$, la concentración del titulante es $c(\text{AgNO}_3) = 100 \text{ mmol/L}$, de los cuales se consumieron $V(\text{AgNO}_3) = 10 \text{ mL}$ hasta alcanzar el punto final. La concentración del indicador fue $c(\text{K}_2\text{CrO}_4) = 1,75 \text{ mmol/L}$. Se asume que las concentraciones del titulante y del indicador son conocidas exactamente.

En este caso consideraremos únicamente el impacto de la incertidumbre en las determinaciones volumétricas y en la determinación visual del punto final. Suponiendo que $V(\text{NaCl})$ se midió con una pipeta volumétrica de Clase A, puede ser modelado utilizando una distribución de probabilidad triangular simétrica cuya función de densidad de probabilidad tiene un gráfico triangular simétrico el cual tiene una base de

ancho 0.16 mL (es decir, la desviación estándar es de $0,08/\sqrt{6}$ mL). Suponiendo que $V(\text{AgNO}_3)$ se midió utilizando una bureta electrónica con una resolución de 0.01 mL, modelamos $\Delta_V V(\text{AgNO}_3)$ utilizando una distribución triangular simétrica cuyo ancho es $\pm 0,01$ mL, y modelamos la detección del punto final utilizando una distribución gaussiana cuyo intervalo de cobertura del 95 % corresponde al volumen de una sola gota (o una desviación estándar de 0.025 mL). El modelo de medición es

$$c(\text{NaCl}) = c(\text{AgNO}_3) \frac{V(\text{AgNO}_3) + \Delta_V V(\text{AgNO}_3) + \Delta_E V(\text{AgNO}_3)}{V(\text{NaCl}) + \Delta_V V(\text{NaCl})},$$

donde $\Delta_V V(\text{AgNO}_3)$ y $\Delta_V V(\text{NaCl})$ denotan los errores al medir los volúmenes de las soluciones del analito y el titulante, y $\Delta_E V(\text{AgNO}_3)$ denota el error en la determinación del punto final. La aplicación de la fórmula de Gauss, o de la [NIST Uncertainty Machine](#) [50], que implementa tanto la fórmula de Gauss como el “bootstrap” paramétrico para la propagación de distribuciones (método de Monte Carlo), da como resultado $c(\text{NaCl}) = 10,000$ mmol/L y $u(c(\text{NaCl})) = 0,026$ mmol/L, por lo tanto, una incertidumbre estándar relativa del 0.26 %.

Un análisis más elaborado de las reacciones químicas que rigen esta titulación, mencionado en el ejemplo 3A, revela que la diferencia entre el punto de equivalencia y el punto final es una fuente de incertidumbre cuya contribución puede ser comparable a la incertidumbre evaluada anteriormente, dependiendo de la concentración del indicador K_2CrO_4 y la estimación preliminar de la concentración desconocida $c(\text{NaCl})$. En lugar de modelar esta fuente de incertidumbre probabilísticamente, como se hizo anteriormente, Juris Meija *et al.* [15, ec. (19)] muestran que la concentración de NaCl se puede obtener resolviendo la siguiente ecuación para el volumen de titulante, $V_E(\text{AgNO}_3)$, necesario para alcanzar el punto final:

$$\frac{K_{sp1}}{\sqrt{K_{sp2} \frac{V(\text{NaCl}) + V_E(\text{AgNO}_3)}{c(\text{K}_2\text{CrO}_4)V(\text{NaCl})}}} - \sqrt{K_{sp2} \frac{V(\text{NaCl}) + V_E(\text{AgNO}_3)}{c(\text{K}_2\text{CrO}_4)V(\text{NaCl})}}} = \frac{c(\text{NaCl})V(\text{NaCl}) - c(\text{AgNO}_3)V_E(\text{AgNO}_3)}{V(\text{NaCl}) + V_E(\text{AgNO}_3)},$$

donde K_{sp1} y K_{sp2} son los productos de solubilidad de AgCl y Ag_2CrO_4 respectivamente, cuyos valores numéricos son $\log_{10}[K_{sp1}/(\text{mol/L})^2] = -9,75$ y $\log_{10}[K_{sp2}/(\text{mol/L})^3] = -11,9$ [49, tabla 1]. La diferencia relativa entre la estimación resultante de $c(\text{NaCl})$ y la estimación convencional dada anteriormente es del 0.24 %.

Ejemplo 4J (Michaelis–Menten — Propagación de incertidumbre mediante Monte Carlo) La enzima β -galactosidasa cataliza la hidrólisis del sustrato o-nitrofenil- β -D-galactopiranosido (ONPG, por sus siglas en inglés) para producir galactosa y o-nitrofenol a una velocidad inicial v_0 , que depende de la concentración inicial de ONPG, c_0 , según la ecuación de Michaelis–Menten [51, §30A-4] [52] [53]:

$$v_0 = v_{\text{MAX}} \frac{c_0}{K_M + c_0}.$$

La velocidad máxima inicial v_{MAX} , y la constante de Michaelis K_M , pueden estimarse a partir de los conjuntos de valores de c_0 y v_0 , que se enumeran en la siguiente tabla [54, 55].

$c_0 / (\text{mmol L}^{-1})$	0.05	0.1	0.25	0.5	1	2.5	5	8	20	30
$v_0 / (\mu\text{mol L}^{-1} \text{min}^{-1})$	3	5.2	14.4	30.3	49	86.2	112.6	136.2	170	177.7

Tabla 2. Resultados de medición utilizados para estimar los parámetros del modelo de Michaelis-Menten para la hidrólisis de ONPG.

El modelo de medición estadística $v_{0,j} = v_{\text{MAX}} c_{0,j} / (K_M + c_{0,j}) + \Delta v_{0,j}$, para $j = 1, \dots, 10$, suponiendo errores $\Delta v_{0,1}, \dots, \Delta v_{0,10}$ que son muestra de una distribución [gaussiana](#) con media $0 \mu\text{mol L}^{-1} \text{min}^{-1}$ y desviación estándar σ , puede ajustarse a los datos minimizando $\sum_{j=1}^n (v_{0,j} - v_{\text{MAX}} c_{0,j} / (K_M + c_{0,j}))^2$, donde $n = 10$ es el número de observaciones. Este problema de mínimos cuadrados no lineales puede resolverse utilizando un optimizador numérico. Las estimaciones resultantes de los parámetros en el modelo de medición estadística mencionado anteriormente son $\hat{v}_{\text{MAX}} = 195 \mu\text{mol L}^{-1} \text{min}^{-1}$ y $\hat{K}_M = 3,27 \text{ mmol L}^{-1}$. La desviación estándar de los $\{\Delta v_{0,j}\}$ se estima en $\hat{\sigma} = 3,09 \mu\text{mol L}^{-1} \text{min}^{-1}$, basada en 8 grados de libertad.

Para evaluar las incertidumbres asociadas a $\hat{v}_{\text{MAX}}$ y $\hat{K}_M$, empleamos el método de “bootstrap” paramétrico, repitiendo las siguientes etapas un gran número K de veces, para $k = 1, \dots, K$:

- (1) Seleccionar aleatoriamente (extraer) una muestra $e_{1,k}, \dots, e_{n,k}$ de una distribución t de Student con 8 grados de libertad, redimensionada para tener una desviación estándar $\hat{\sigma}$, y generar $v_{0,j,k}^* = \hat{v}_{0,j} + e_{j,k}$, donde $\hat{v}_{0,j} = \hat{v}_{\text{MAX}} c_{0,j} / (\hat{K}_M + c_{0,j})$, asegurándose que la distribución de probabilidad de los $v_{0,j,k}^*$ resultantes, para $j = 1, \dots, 10$, esté truncada en $0 \mu\text{mol L}^{-1} \text{min}^{-1}$. Muestreamos de esta distribución t de Student, en lugar de una distribución gaussiana, para identificar el número de grados de libertad en los que se basa $\hat{\sigma}$.
- (2) Ajustar el modelo de Michaelis-Menten a los $n = 10$ pares de valores de c_0 y v_0^* utilizando el mismo procedimiento mencionado anteriormente, para obtener estimaciones de “bootstrap” de sus parámetros, $v_{\text{MAX},k}^*$ y $K_{M,k}^*$.

Las desviaciones estándar de las $\{v_{\text{MAX},k}^*\}$ y de las $\{K_{M,k}^*\}$ son las evaluaciones de $u(v_{\text{MAX}}) = 3 \mu\text{mol L}^{-1} \text{min}^{-1}$ y $u(K_M) = 0,18 \text{ mmol L}^{-1}$. Mientras que tales evaluaciones de incertidumbre pueden, en circunstancias favorables, obtenerse utilizando aproximaciones o incluso fórmulas exactas, la gran ventaja del “bootstrap” paramétrico es su flexibilidad, lo que lo hace aplicable incluso en situaciones desafiantes, independientemente de si se dispone de las fórmulas apropiadas.

El “bootstrap” paramétrico también puede aplicarse para evaluar las incertidumbres asociadas con las estimaciones de v_{MAX} y de K_M producidas por el método ampliamente utilizado propuesto por Lineweaver y Burk [56], basado en la expresión linealizada $1/v_0 = 1/v_{\text{MAX}} + (K_M/v_{\text{MAX}})/c_0$. Incluso cuando esta relación se ajusta a los datos utilizando pesos proporcionales a v_0^4 [54], las estimaciones resultantes aún presentan un sesgo notable y sus incertidumbres asociadas son mayores que sus contrapartes para el modelo no lineal. La elección de los pesos tiene una influencia significativa en la calidad de las estimaciones de los parámetros de velocidad de reacción, lo que es otra razón para abandonar la versión linealizada del modelo. En este caso, el “bootstrap” paramétrico también propaga las incertidumbres asociadas con los pesos utilizados en la regresión lineal que produce las estimaciones WLB (Weighted Lineweaver-Burk, por sus siglas en inglés), cuyas densidades de probabilidad se representan en la figura 1, junto con las densidades de probabilidad de las estimaciones obtenidas mediante optimización no lineal (NL).

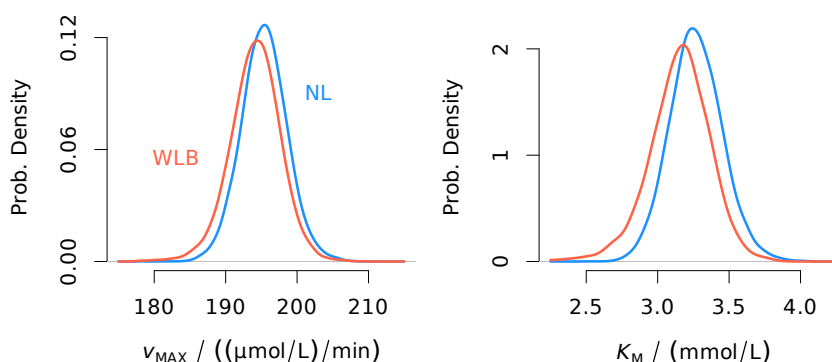


Fig. 1. Densidades de probabilidad de v_{MAX} y K_M estimadas por regresión no lineal (NL) y por regresión lineal ponderada (WLB)

4.6. Métodos no paramétricos

El método estadístico del “*bootstrap* no paramétrico” es aún más audaz que el “*bootstrap* paramétrico” [38]. Considera el conjunto de observaciones replicadas disponibles de la cantidad de interés, $\{x_1, \dots, x_n\}$, como si fueran una muestra infinitamente grande de una distribución subyacente no especificada, y la vuelve a muestrear repetidamente como se describe a continuación. (Algunas versiones del “bootstrap” son aplicables cuando las observaciones deben ser modeladas como resultados de variables aleatorias mutuamente dependientes [57]). Para evaluar la incertidumbre asociada con alguna función, $y = f(x_1, \dots, x_n)$, de las observaciones, repita los siguientes pasos para $k = 1, \dots, K$, un número adecuadamente grande K de veces:

- Extraiga un solo valor del conjunto $\{x_1, \dots, x_n\}$ de tal manera que todos tengan la misma probabilidad de ser extraídos, registre su valor y devuelva el

valor extraído al conjunto. Repita esto n veces para obtener una muestra de “bootstrap” $\{x_{1,k}^*, \dots, x_{n,k}^*\}$: es decir, cada observación x_i tiene la misma probabilidad de ser extraída en la k -ésima muestra de “bootstrap”, y puede ser extraída varias veces en esta muestra porque el muestreo se realiza con reemplazo.

(b) Calcule la réplica de “bootstrap” $y_k^* = f(x_{1,k}^*, \dots, x_{n,k}^*)$.

La desviación estándar de las réplicas de “bootstrap” $y_1^*, \dots, y_K^*$ es una evaluación de $u(y)$, y los percentiles 2.5 y 97.5 del conjunto de estas réplicas son los extremos de un intervalo de cobertura del 95 % (generalmente asimétrico) para el valor verdadero de y .

Los métodos no paramétricos pueden ser útiles para producir evaluaciones de incertidumbre con mínimas suposiciones sobre los resultados de medición, por ejemplo, para evaluar la incertidumbre asociada con la mediana de determinaciones replicadas de la misma cantidad, lo cual puede hacerse ya sea mediante la aplicación del “bootstrap” no paramétrico, o mediante la inversión de la prueba de signos [58, Capítulo 3]. El Material Suplementario detalla la prueba de signos, y explica cómo su inversión produce intervalos de cobertura válidos para la mediana.

Ejemplo 4K (Basaltos del Monte Submarino Verde — “Bootstrap” no paramétrico para la mediana de la fracción de masa de TiO_2) Las siguientes determinaciones de la fracción de masa de TiO_2 , expresadas en mg/g, de $n = 22$ muestras de basaltos de la dorsal mesoceánica del *Monte Submarino Verde* en el Océano Pacífico cerca de México, descritas por Allan *et al.* [59], se enumeran en la base de datos *Smithsonian Abyssal Volcanic Glass Data File* de la institución Smithsonian[60, 61]: 15.2, 14.2, 15.1, 15.0, 12.4, 12.2, 12.7, 14.9, 14.0, 14.0, 14.4, 20.1, 20.5, 11.2, 13.7, 14.1, 12.7, 12.8, 17.9, 11.7, 13.1, 13.5.

Es muy improbable que estas determinaciones sean una muestra de una distribución [gaussiana](#) (la prueba de Shapiro-Wilk [62] de forma [gaussiana](#) arroja un [valor de p](#) de 0.003), pero parecen ser consistentes con una distribución simétrica (la prueba de simetría propuesta por Miao *et al.* [63] arroja un [valor de p](#) de 0.32). Por lo tanto, puede ser razonable considerar un modelo de error de medición para estas determinaciones de la forma $w_{\text{obs},i} = w + \Delta w_i$ para $i = 1, \dots, n$, de manera similar al ejemplo 3D, pero donde w denota la verdadera fracción de masa de TiO_2 , y los errores de medición $\Delta w_1, \dots, \Delta w_n$ son una muestra de una [distribución \(doble exponencial\) de Laplace](#) (una prueba de si la muestra mencionada podría provenir de una distribución de Laplace [64] arroja un [valor de p](#) de 0.99). En estas circunstancias, el estimador de máxima verosimilitud de w es la mediana, 14.0 mg/g.

La incertidumbre asociada con la mediana puede evaluarse utilizando el “bootstrap” no paramétrico, lo que implica tomar un gran número K , de muestras aleatorias de tamaño $n = 22$, extraídas con reemplazo del conjunto de 22 valores observados an-

teriormente, como se explica en la sección 4.6, y calcular la mediana de cada una de estas muestras.

La fiabilidad de la desviación estándar para estas réplicas de “bootstrap” de la mediana, como evaluación de la incertidumbre estándar asociada a la mediana, depende tanto del número total de réplicas, como del número de réplicas que son numéricamente diferentes entre sí. En general, K no debería ser pequeño (al menos 1000), y el número de réplicas numéricamente diferentes debería ser lo suficientemente grande como para que K sea una fracción pequeña del número, N_B , de posibles muestras de “bootstrap” diferentes. En este caso, en el cual las determinaciones comprenden $m = 20$ valores numéricamente diferentes, $K = 10^5$ es una pequeña fracción de $N_B \approx 7 \times 10^{10}$ [65].

La evaluación resultante de la incertidumbre estándar de la mediana de la fracción de masa de TiO_2 fue de 0.41 mg/g, con un intervalo de cobertura del 95 % para la misma cantidad que osciló entre 12.9 mg/g y 14.7 mg/g. Una evaluación alternativa de la incertidumbre para la mediana se puede obtener mediante la inversión de la prueba de signos [58, capítulo 3], que proporciona un intervalo de cobertura del 95 % que varía entre 12.8 mg/g y 14.9 mg/g.

Ejemplo 4L (Secuenciación de ADN — Evaluación de incertidumbre de medida mediante el método de Monte Carlo para una propiedad nominal) La base nitrogenada en una posición particular de un fragmento de ADN extraído del Material de Referencia 8256 del NIST, Salmón Coho capturado en estado salvaje [8], fue identificado como adenina (A) mediante secuenciación de alto rendimiento (secuenciación de “última generación”). Sin embargo, la probabilidad asociada a la identificación errónea, reportada en el archivo de formato FASTQ resultante, fue inusualmente alta: 0.16. En ausencia de otra información, las probabilidades de las cuatro bases nitrogenadas en ese locus son de 0.84 (que es $1 - 0.16$) para A (adenina), y $0.16/3 = 0.053$ para cada una de las restantes bases: C (citosina), G (guanina) y T (timina). Esta distribución de probabilidad discreta caracteriza completamente la incertidumbre en la identificación de la base nitrogenada en esa posición del ADN. El muestreo de Monte Carlo de distribuciones similares para las diferentes posiciones del fragmento de ADN genera réplicas de la composición del fragmento, que luego pueden ser alineadas y comparadas con secuencias de referencia de [especímenes](#) de especies estrechamente relacionadas, con el fin de identificar el material (ver ejemplo 4L).

4.7. Métodos bayesianos

La estimación bayesiana de los parámetros del modelo y sus incertidumbres asociadas proporciona el medio para combinar información previa sobre el mensurando con nueva información de los resultados de medición recién adquiridos [66]. En modelos estadísticos de medición (ver Sección 3), el mensurando es una función de los parámetros de la distribución de probabilidad de las cantidades de entrada. El enfoque bayesiano para la estimación y evaluación de la incerti-

dumbre en modelos estadísticos de medición implica modelar esos parámetros como valores (no observables) de variables aleatorias, y los datos de medición como resultados observados de variables aleatorias cuyas distribuciones dependen de los valores desconocidos de estos parámetros [12, página 204].

Los métodos bayesianos requieren que el conocimiento previo sobre esos parámetros, antes de llevar a cabo el experimento de medición, esté contenido en una distribución de probabilidad, que es convenientemente llamada la *distribución a priori*. La estimación del mensurando y evaluación de incertidumbre asociada se derivan de la distribución conjunta de los parámetros del modelo, condicionadas por la información proporcionada por los datos. Esta distribución condicional, llamada la *distribución a posteriori*, se obtiene mediante la aplicación del Teorema de Bayes [67], que establece que la función de densidad de probabilidad a posteriori de los parámetros es proporcional al producto de su función de densidad de probabilidad *a priori* conjunta y la función de verosimilitud.

Ejemplo 4M (Masa molar del iridio — Teorema de Bayes) Consideremos las tres determinaciones independientes de la masa molar del iridio del ejemplo 3E. La función de verosimilitud es un producto de tres funciones de densidad de probabilidad gaussianas, de manera similar al ejemplo 4F. Se asume que los dos parámetros del modelo de medición, M y τ , son *a priori* independientes.

La masa molar M tiene una distribución *a priori* gaussiana cuya media y desviación estándar corresponden al peso atómico estándar (masa atómica relativa) del iridio, $192,217 \pm 0,003$, publicado en 1993 [68]. La *distribución a priori* para τ es semi-Cauchy, con una mediana de 0.0004 g/mol, como se menciona en el ejemplo 3E. (La distribución de semi-Cauchy, usada comúnmente para componentes de varianza [69], es la distribución t de Student con 1 grado de libertad, truncada en cero). Esta elección de datos previos equivale a decir que esperamos que la media de los tres valores medidos listados en el ejemplo 3E se alinee con el valor del peso atómico estándar (masa atómica relativa) del iridio publicado en 1993, sin descartar la posibilidad de considerables discrepancias entre los valores medidos actualmente.

La función de densidad de probabilidad q , de la *distribución a posteriori* conjunta de M y τ , es

$$q(M, \tau) \propto \underbrace{p_M(M) \times p_\tau(\tau)}_{a \text{ priori conjunta}} \times \underbrace{L_{\text{datos}}(M, \tau)}_{\text{verosimilitud}},$$

donde los datos comprenden los tres valores medidos de la masa molar del iridio dados en el ejemplo 3E ($M_{\text{obs},1}$, $M_{\text{obs},2}$ y $M_{\text{obs},3}$) junto con las incertidumbres de medición asociadas $u(M_{\text{obs},1})$, $u(M_{\text{obs},2})$ y $u(M_{\text{obs},3})$. Además, p_M es la función de densidad de probabilidad de la *distribución a priori* de M , que es gaussiana con media 192.217 g/mol y desviación estándar 0.0015 g/mol, y p_τ es la función de densidad de probabilidad de la *distribución a priori* de τ , que es la distribución semi-Cauchy con mediana 0.00042 g/mol mencionada anteriormente. (Usamos la mediana para espe-

cificar una distribución semi-Cauchy particular, porque los miembros de esta familia de distribuciones no tienen media finita.)

En la práctica, el Teorema de Bayes rara vez se utiliza directamente dado que involucra integrales (o sumas) que no pueden ser evaluadas analíticamente o pueden ser demasiado complicadas de calcular numéricamente. En su lugar, se utilizan los métodos de Monte Carlo basados en cadenas de Markov (MCMC, por sus siglas en inglés) para obtener muestras de las *distribuciones a posteriori* independientemente de su complejidad [69]. El algoritmo Random Walk Metropolis (RWM, por sus siglas en inglés), el muestreo de Gibbs y el algoritmo hamiltoniano de Monte Carlo son versiones ampliamente utilizadas de MCMC [70, 71].

Ejemplo 4N (Masa molar del iridio — Métodos de Montecarlo basados en cadenas de Markov) Utilizaremos el muestreo MCMC para obtener muestras de la *distribución a posteriori* de M y τ , como se especifica en el ejemplo 4M. Emplearemos una versión particularmente simple de MCMC: el algoritmo RWM con una *distribución propuesta* simétrica (cuyo significado se aclara más adelante). Esto implica repetir los siguientes pasos un gran número K de veces, después de hacer una elección inicial de valores razonables, M_0 y τ_0 , para asignar a M y τ [12], para $k = 1, \dots, K$:

- (1) Calcular $p(M_k, \tau_k) = p_M(M_k)p_\tau(\tau_k)L(M_k, \tau_k)$ (suprimiendo la referencia a los datos que normalmente se haría en el subíndice de L).
- (2) Generar un par de candidatos (M^*, τ^*) en las proximidades de (M_k, τ_k) , y calcular $p(M^*, \tau^*)$ como se mencionó anteriormente. Suponemos que (M^*, τ^*) es una extracción de una distribución circularmente simétrica, *gaussiana* bivariada centrada en (M_k, τ_k) con desviaciones estándar especificadas y pequeñas, la llamada *distribución propuesta*.
- (3) Calcular $\alpha = p(M^*, \tau^*)/p(M_k, \tau_k)$.
- (4) Si $\alpha > 1$, entonces aceptar los valores propuestos de los parámetros y establecer $M_{k+1} = M^*$ y $\tau_{k+1} = \tau^*$. Si $\alpha \leq 1$, entonces obtener un número z de una distribución uniforme entre 0 y 1, y aceptar los valores propuestos solo si $\alpha > z$. De lo contrario, mantener los valores actuales estableciendo $M_{k+1} = M_k$ y $\tau_{k+1} = \tau_k$.

La porción inicial (normalmente la primera mitad) de la secuencia de muestras, $(M_1, \tau_1), \dots, (M_K, \tau_K)$, suele ser descartada para asegurar que el MCMC haya alcanzado el equilibrio y esté efectivamente generando muestras de la *distribución a posteriori* prevista. También se puede hacer una reducción de la secuencia (conservando por ejemplo solo los pares de valores cuyos índices k sean múltiplos de 10 o 25), antes de usarla para producir estimaciones de los parámetros y evaluaciones de incertidumbre.

Aplicar la versión de este algoritmo implementada en la función `MCMCmetrop1R` del software R, definida en el paquete `MCMCpack` [72], con $K = 5 \times 10^5$, utilizando los tres

resultados de medición para M listados en el ejemplo 3E, descartando los primeros $K/2$ pares de valores de los parámetros, y reduciendo el resto para mantener solo cada par de valores múltiplo de 25, produce una muestra MCMC de un tamaño de 10 000. La media de la *distribución a posteriori* de M es $\hat{M} = 192,217\,1(4)\,\text{g/mol}$, y la estimación de la incertidumbre oscura es la mediana de la *distribución a posteriori* de τ , $\hat{\tau} = 0,000\,5\,\text{g/mol}$ (elegimos la mediana para τ , dado que la distribución de τ es notablemente asimétrica, con una cola derecha mucho más pesada que la izquierda).

Los modelos y métodos bayesianos nos permiten incorporar información proporcionada por los resultados de mediciones recientes junto con cualquier información preexistente relevante sobre los mensurandos o incluso sobre otras cantidades que sean relevantes para la medición, como se ilustra en los ejemplos 4O y 4P. Las técnicas bayesianas también ofrecen formas flexibles de respetar restricciones que algunos parámetros o mensurandos deben cumplir, por ejemplo, que las fracciones de masa deben ser positivas o que la pureza no puede exceder 1 g/g, e incluso pueden propagar la incertidumbre que a menudo rodea a los modelos empíricos de medición (por ejemplo, en relación con la elección del grado de un polinomio que se utiliza como curva de calibración o análisis en metrología de gases [73]).

Ejemplo 4O (Nitrato en agua de mar — Caracterización bayesiana de la precisión de medición) Cuatro determinaciones espectrofotométricas de la fracción de masa de nitrato en una muestra de agua de mar se obtuvieron bajo condiciones de repetibilidad de la medición utilizando el método de Griess: $w_{\text{obs}}/(\text{mg/kg}) = 0.1514, 0.1523, 0.1531$ y 0.1545 .

Basándonos en el desempeño pasado de este método de medición, sabemos que la incertidumbre relativa es típicamente del 1 %, pero a veces podría ser tan baja como 0.33 % o tan alta como 3 %. Esta información previa sobre la desviación estándar de la repetibilidad puede ser modelada utilizando una distribución **gamma** centrada en alrededor 1 % con la mayoría (por ejemplo, el 90 %) de su probabilidad distribuida entre 0.33 % y 3 %. Dicha distribución **gamma** tiene una forma $\alpha = 2,619$ y una escala $\beta = 1248\,\text{kg/mg}$.

El modelo estadístico de medición $w_{\text{obs},i} = w + \Delta w_{\text{obs},i}$, para $i = 1, \dots, 4$, relaciona las observaciones de la fracción de masa de nitrato con su valor verdadero w , donde los errores de medición $\{\Delta w_{\text{obs},i}\}$ se asumen como una muestra de una distribución **gaussiana** con media 0 mg/kg y desviación estándar σ . La información previa sobre σ está contenida en la mencionada distribución **gamma**. Una distribución previa **gaussiana** poco informativa para w , con media 0 mg/kg y una desviación estándar de 1 mg/kg, truncada en cero, refleja la expectativa previa sobre el orden de magnitud de w .

MCMC se utiliza para producir una muestra grande de la distribución condicional de los parámetros, w y σ , dadas las cuatro observaciones. El promedio de los valores de

w en esta muestra, 0.1528 mg/kg, es la estimación bayesiana de w , y su desviación estándar, 0.0010 mg/kg, es la estimación bayesiana de $u(w)$.

Es importante tener en cuenta que esta evaluación bayesiana de la incertidumbre estándar es aproximadamente un 50 % más grande que la evaluación convencional de Tipo A de la incertidumbre estándar asociada al promedio de las cuatro réplicas. Esto no es sorprendente, dado que los métodos bayesianos suelen considerar las contribuciones de incertidumbre de manera más exhaustiva que los métodos clásicos. Este ejemplo muestra que tener en cuenta información previa sobre una cantidad auxiliar — la precisión de la medición σ , en este caso — puede afectar significativamente las conclusiones alcanzadas sobre el mensurando de interés primario, que era la fracción de masa de nitritos.

Ejemplo 4P (Molnupiravir — Reconsideración escéptica mediante modelado bayesiano) Pogue and Yusuf [74] sugieren que “la mayoría de las intervenciones clínicamente relevantes probablemente reducirán el riesgo relativo de resultados importantes, como infarto de miocardio, accidente cerebrovascular o muerte, en aproximadamente un 10-20 %”, y Higgins y Spiegelhalter [75] señalan que es común en la investigación médica ver cómo el optimismo de los primeros resultados se apaga con investigaciones posteriores. El ejemplo 4G describe un caso de tal optimismo inicial sobre el molnupiravir, impulsado por los resultados informados por Bernal *et al.* [44], que mostraron una reducción del $(0,0973 - 0,0677)/0,0973 = 30$ % en la frecuencia de hospitalizaciones o muertes. El posterior, y mucho más grande, ensayo PANORAMIC redujo esas altas expectativas [46].

Higgins y Spiegelhalter [75] han demostrado cómo un enfoque bayesiano para el análisis de los resultados de ensayos clínicos puede inyectar una medida de escepticismo que contrarresta el optimismo potencialmente excesivo y temprano sobre una intervención. Si el tratamiento y el placebo tienen aproximadamente el mismo efecto, entonces la razón de posibilidades correspondiente debería ser 1. Por lo tanto, una reducción del 25 % en las posibilidades de hospitalización o muerte situaría la razón de posibilidades en 0.75. Para que una variable aleatoria gaussiana con media $\ln(1) = 0$ tome un valor menor a $\ln(0,75)$ con una probabilidad del 5 %, su desviación estándar necesita ser $\ln(0,75)/(-1,645) = 0,175$, donde -1,645 es el percentil 5 de una distribución gaussiana con media cero y desviación estándar unitaria.

Por lo tanto, podemos utilizar una *distribución a priori* gaussiana con una media de 0 y una desviación estándar de 0.175 para expresar la creencia previa de que es bastante improbable (probabilidad menor al 5 %) que un nuevo tratamiento pueda reducir las posibilidades de hospitalización o muerte en más de un 25 %. Dado que la distribución de probabilidad del logaritmo de la razón de posibilidades medida en el ejemplo 4G es aproximadamente gaussiana [45], un resultado de la teoría de la probabilidad [76, teorema 1] implica que la distribución posterior del valor verdadero del logaritmo de la razón de posibilidades también es gaussiana, con una media de $(-0,395/0,197^2 + 0/0,175^2)/(1/0,197^2 + 1/0,175^2) = -0,174$ y una desviación estándar de $\sqrt{1/(1/0,197^2 + 1/0,175^2)} = 0,131$.

El análisis convencional en el ejemplo 4G arrojó un z-score de $-0,395/0,197 = -2,01$ para la eficacia del molnupiravir, lo que sugiere una significancia cuestionable para el efecto del tratamiento. Sin embargo, una vez que se introduce una mirada escéptica acerca del valor verdadero del logaritmo de la razón de posibilidades, el molnupiravir no parece ser decisivamente más eficaz que el placebo (la probabilidad posterior de que así sea es solo del 9.1 %). Por lo tanto, este reanálisis escéptico y bayesiano de los resultados del ensayo clínico reportado por Bernal *et al.* [44] arroja conclusiones similares a las que se obtienen del ensayo más grande y reciente reportado por Butler *et al.* [46]: que el molnupiravir no es un tratamiento significativamente beneficioso para la COVID-19.

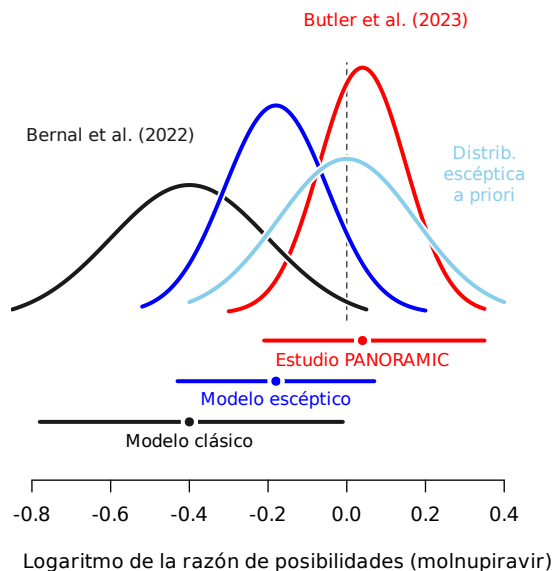


Fig. 2. Evaluar los resultados de los ensayos clínicos con modelos bayesianos escépticos puede conducir a conclusiones más confiables que los métodos clásicos de análisis estadístico.

5. Reporte de incertidumbre de medición

Las incertidumbres de medición deben ser reportadas de tal manera que sean adecuadas para el uso previsto, lo que debería modelar tanto la forma de expresión como el número de cifras utilizadas. La cláusula 7.2.6 de la GUM [2] recomienda que las evaluaciones de incertidumbre se informen “con un máximo de dos cifras significativas, aunque en algunos casos puede ser necesario mantener cifras adicionales para evitar errores de redondeo en cálculos posteriores.”

En la mayoría de los casos, especificar un conjunto de valores del mensurando que se cree que incluyen su valor verdadero con una probabilidad específica, es suficiente como expresión de la incertidumbre de medición. La forma en que

se construye dicho conjunto o intervalo depende del método utilizado para la evaluación de la incertidumbre como se describe en la sección 4, pero necesariamente implica alguna suposición sobre la distribución de probabilidad del mensurando.

Si dicho intervalo, $y \pm U_\alpha(y)$, cubre el valor verdadero del mensurando con una alta probabilidad, típicamente $\alpha > 0,9$, entonces decimos que $U_\alpha(y)$ es una *incertidumbre expandida* de y con una probabilidad de cobertura del 100α %. Por convención, la incertidumbre expandida para una cobertura del 95 % es el doble de la incertidumbre estándar (ver Sección 6.5 de [77]). Esto está motivado por el sorprendente hecho de que, en muchos casos, un intervalo de cobertura de la forma $y \pm 2u(y)$ logra una probabilidad de cobertura de aproximadamente el 95 % incluso cuando la distribución de probabilidad es notablemente asimétrica (ver Páginas 68 y A-9 de [78]).

Ejemplo 5A (Resultado de medición de una masa) $m = 100,02147$ g con una incertidumbre estándar asociada de $u(m) = 0,00035$ g, también expresado como $m = 100,02147(35)$ g. El valor verdadero de m se encuentra en el intervalo $(100,02147 \pm 0,00070)$ g, o entre $100,02077$ g $\leq m \leq 100,02217$ g, con aproximadamente un 95 % de probabilidad.

Ejemplo 5B (Peso atómico estándar del boro) El peso atómico (masa atómica relativa) de varios elementos varía significativamente entre los materiales terrestres. Dado que el boro es uno de estos elementos, su peso atómico estándar se expresa en forma de un intervalo, $A_r^o(\text{B}) = [10,806, 10,821]$, lo que significa que el peso atómico del boro en un material terrestre no especificado, $A_r(\text{B})$, está entre $10,806 \leq A_r(\text{B}) \leq 10,821$, con al menos un 95 % de probabilidad [20].

Cuando el resultado de una evaluación de incertidumbre de medición está destinado a ser utilizado en ejercicios posteriores de propagación de incertidumbre que involucren métodos de Monte Carlo o bayesianos, entonces la expresión de la incertidumbre de medición debe ser una distribución de probabilidad completamente especificada para la magnitud medida, o la muestra extraída de una distribución de probabilidad que caracterice el nivel de conocimiento sobre la magnitud medida debe ser lo suficientemente grande.

Ejemplo 5C (Fracción de masa de ácido benzoico - incertidumbre de medición para una propiedad cuantitativa) El nivel de conocimiento sobre la fracción de masa de ácido benzoico en el NIST PS1, Estándar Primario para RMN Cuantitativo (Ácido Benzoico), está caracterizado por una distribución [beta](#) que se aproxima a la *distribución posterior bayesiana* con parámetros de forma $\alpha = 88873,7$ y $\beta = 7,11$ [79]. La desviación estándar de esta *distribución beta* es una evaluación de la incertidumbre estándar asociada con la fracción de masa del ácido benzoico del NIST PS1, y los intervalos creíbles de su valor verdadero, con una probabilidad específica, que también pueden derivarse de la misma distribución.

La incertidumbre de medición para los mensurandos nominales puede expresarse en términos de la entropía de Shannon [80, 81] de la correspondiente distribución de probabilidad discreta .

Ejemplo 5D (Identificación del salmón — incertidumbre de medición para una propiedad nominal) Un gran número de réplicas de Monte Carlo (véase ejemplo 4L) de la composición de un fragmento de ADN particular extraído del Material de Referencia NIST 8256, Salmón Coho capturado en estado salvaje [8], fueron alineadas y comparadas con secuencias de referencia de especímenes de referencia de varias especies estrechamente relacionadas de salmón. Las frecuencias de las coincidencias más cercanas fueron las siguientes: Coho ($p_1 = 0,921$), Steelhead ($p_2 = 0,048$), Chinook ($p_3 = 0,022$), Sockeye ($p_4 = 0,005$), Pink ($p_5 = 0,003$), Chum ($p_6 = 0,001$) y Atlantic ($p_7 = 0,000$). La entropía de Shannon de esta distribución de probabilidad discreta, $S = -\sum_{i=1}^7 p_i \ln(p_i) = 0,36$, resume la dispersión de la unidad de probabilidad de las siete posibles identidades del material, por lo tanto cuantifica la incertidumbre de medición para un mensurando nominal. Cuanto mayor sea la entropía de Shannon, mayor será la incertidumbre. (El valor más alto que puede tener la entropía de Shannon en este caso es 1.95, que se logra cuando las siete especies se consideran igualmente probables. Cuando se está seguro sobre la especie en cuestión, la entropía de Shannon alcanza su valor mínimo de 0). La entropía de Shannon también es relevante para distribuciones de probabilidad continuas. Por ejemplo, una [distribución gaussiana](#) con desviación estándar σ tiene una entropía de Shannon $S = \ln(\sigma) + \frac{1}{2} \ln(2\pi e)$.

El intervalo más corto con una probabilidad de cobertura especificada no necesariamente debe ser simétrico alrededor del valor medido. El método de Monte Carlo puede utilizarse para propagar incertidumbres expresadas asimétricamente, ya sea como intervalos asimétricos (ejemplo 5E) o como distribuciones de probabilidad asimétricas, sin “simetrizarlas” primero [82].

Ejemplo 5E (Intervalos de cobertura asimétricos) La fracción de masa de ácido benzoico en el Estándar Primario PS1 del NIST para RMN cuantitativo es de $999,92^{+0,04}_{-0,06}$ mg/g con un 95 % de confianza. Esto significa que se cree que el valor verdadero de esta fracción de masa está entre 999.86 mg/g y 999.96 mg/g con el nivel de confianza indicado.

Membresía de organismos patrocinadores

La membresía del Comité de la División de Química Analítica de la IUPAC para el período 2020-2021 fue la siguiente: **Presidente:** Z. Mester (Canadá); **Secretario:** D. Craston (Reino Unido); **Vicepresidente:** D. Shaw (EEUU); **Presidente anterior:** J. Labuda (Eslovaquia); **Miembros Titulares:** V. Baranovskaia (Rusia); H. Kim (Corea); P. Krystek (Países Bajos); C. Magalhães (Portugal); T. Takeuchi (Japón); S. Wiedmer (Finlandia); **Miembros Asociados:** R. Apak (Turquía); J. Barek (República Checa); F. Emmerling (Alemania); E. Flores (Brasil); I. Kuselman (Israel); H. Li (China/Pekín); **Representantes Nacionales:** M. F. Camões (Portugal); O. Chailapakul (Tailandia); Y.-J. Chen (China/Taipei); A. Felinger (Hungría); B. Hibbert (Australia); S. Ndiaye (Senegal); M. Piston (Uruguay); R. Sha'Ato (Nigeria); L. Torsi (Italia); F. Vanhaecke (Bélgica)

Aclaración

Algunas entidades comerciales, equipos, software o materiales pueden ser identificados en esta contribución para describir un procedimiento experimental o computacional, o para ilustrar adecuadamente un concepto. Tal identificación no implica recomendación o respaldo por parte de la IUPAC, el National Research Council (NRC) Canada o el National Institute of Standards and Technology (NIST), ni tampoco implica que las entidades, equipos, software o materiales identificados sean necesariamente los mejores disponibles para su propósito previsto.

Referencias

- [1] Tapia-Granados JA (1997) Posibilidades, oportunidades, momios: un comentario sobre la traducción del término odds. *Salud Pública de México* 39(1):69–71. Available at <https://www.saludpublica.mx/index.php/spm/article/view/5978>.
- [2] Joint Committee for Guides in Metrology (JCGM) (2008) *Evaluation of Measurement Data — Guide to the Expression of Uncertainty in Measurement* (International Bureau of Weights and Measures (BIPM), Sèvres, France). BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML, JCGM 100:2008, GUM 1995 with minor corrections Available at https://www.bipm.org/documents/20126/2071204/JCGM_100_2008_E.pdf/cb0ef43f-baa5-11cf-3f85-4dcd86f77bd6.
- [3] Ellison SLR, Williams A (eds) (2012) *Quantifying Uncertainty in Analytical Measurement*. EURACHEM/CITAC Guide CG-4, QUAM:2012.P1 (EURACHEM), 3rd Ed. Available at <https://www.eurachem.org>.
- [4] Possolo A (2015) *Simple Guide for Evaluating and Expressing the Uncertainty of NIST Measurement Results* (National Institute of Standards and Technology, Gaithersburg, MD). <https://doi.org/10.6028/NIST.TN.1900>. NIST Technical Note 1900
- [5] NIST (2019) *NIST Quality Manual for Measurement Services* (National Institute of Standards and Technology, U.S. Department of Commerce, Gaithersburg, Maryland). NIST QM-I Version 11 Available at <https://www.nist.gov/nist-quality-system/>.
- [6] Grinberg P, Nadeau K, Yang L, Gedara IP, Meija J, Mester Z (2017) *AQUA-1: Drinking water Certified Reference Material* (National Research Council Canada, Ottawa, ON). <https://doi.org/10.4224/crm.2017.aqua-1>
- [7] Klein C, Philpotts A (2017) *Earth Materials: Introduction to Mineralogy and Petrology* (Cambridge University Press, Cambridge, UK), 2nd Ed. <https://doi.org/10.1017/9781316652909>
- [8] Ellisor DL, Place B, Phillips M, Yen J (2021) *Analysis of Seafood Reference Materials: RM 8256, RM 8257, RM 8258, and RM 8259, Wild-Caught Coho Salmon (RM 8256), Aquacultured Coho Salmon (RM 8257), Wild-Caught Shrimp (RM 8258), Aquacultured Shrimp (RM 8259)*. NIST Special Publication 260-214 (National Institute of Standards and Technology, Gaithersburg, MD). <https://doi.org/10.6028/NIST.SP.260-214>
- [9] Michell J (2005) The logic of measurement: A realist overview. *Measurement* 38(4):285–294. <https://doi.org/10.1016/j.measurement.2005.09.004>
- [10] White R (2011) The meaning of measurement in metrology. *Accreditation and Quality Assurance* 16:31–41. <https://doi.org/10.1007/s00769-010-0698-1>
- [11] Possolo A (2018) *Measurement. Advanced Mathematical and Computational Tools in Metrology and Testing: AMCTM XI*, eds Forbes AB, Zhang NF, Chunovkina A, Eichstädt S, Pavese F (World Scientific Publishing Company, Singapore),

- Series on Advances in Mathematics for Applied Sciences*, Vol. 89, pp 273–285. https://doi.org/10.1142/9789813274303_0027
- [12] Possolo A, Meija J (2022) *Measurement Uncertainty: A Reintroduction* (Sistema Interamericano de Metrologia (SIM), Montevideo, Uruguay), 2nd Ed. <https://doi.org/10.4224/1tqz-b038>
- [13] Joint Committee for Guides in Metrology (2020) *Guide to the expression of uncertainty in measurement — Part 6: Developing and using measurement models* (International Bureau of Weights and Measures (BIPM), Sèvres, France). BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML, JCGM GUM-6:2020 Available at https://www.bipm.org/documents/20126/2071204/JCGM_106_2012_E.pdf/fe9537d2-e7d7-e146-5abb-2649c3450b25.
- [14] Camões MF, Christian GD, Hibbert DB (2018) Mass and volume in analytical chemistry (IUPAC Technical Report). *Pure and Applied Chemistry* 90(3):563–603. <https://doi.org/10.1515/pac-2017-0410>
- [15] Meija J, Michałowska-Kaczmarczyk AM, Michałowski T (2016) Solution to Mohr’s method challenge. *Analytical and Bioanalytical Chemistry* 408(17):4469–4471. <https://doi.org/10.1007/s00216-016-9555-3>
- [16] Pagliano E, Mester Z, Meija J (2015) Calibration graphs in isotope dilution mass spectrometry. *Analytica Chimica Acta* 896:63–67. <https://doi.org/10.1016/j.aca.2015.09.020>
- [17] Oh MS, Carroll HJ (1977) The Anion Gap. *New England Journal of Medicine* 297(15):814–817. <https://doi.org/10.1056/NEJM197710132971507>
- [18] Carroll RJ, Ruppert D, Stefanski LA, Crainiceanu CM (2006) *Measurement Error in Nonlinear Models — A Modern Perspective* (Chapman & Hall/CRC, Boca Raton, Florida), 2nd Ed.
- [19] Possolo A, Iyer HK (2017) Concepts and tools for the evaluation of measurement uncertainty. *Review of Scientific Instruments* 88(1):011301. <https://doi.org/10.1063/1.4974274>
- [20] Prohaska T, Irrgeher J, Benefield J, Böhlke JK, Chesson LA, Coplen TB, Ding T, Dunn PJH, Gröning M, Holden NE, Meijer HAJ, Moossen H, Possolo A, Takahashi Y, Vogl J, Walczyk T, Wang J, Wieser ME, Yoneda S, Zhu XK, Meija J (2022) Standard atomic weights of the elements 2021 (IUPAC Technical Report). *Pure and Applied Chemistry* 94(5):573–600. <https://doi.org/10.1515/pac-2019-0603>
- [21] Cochran WG (1954) The combination of estimates from different experiments. *Biometrics* 10(1):101–129. <https://doi.org/10.2307/3001666>
- [22] Searle SR, Casella G, McCulloch CE (2006) *Variance Components* (John Wiley & Sons, Hoboken, NJ).
- [23] Mosteller F, Tukey JW (1977) *Data Analysis and Regression* (Addison-Wesley Publishing Company, Reading, Massachusetts).
- [24] Gauss CF (1821) *Theoria combinationis observationum erroribus minimis obnoxiae: Pars prior. Commentationes Societatis Regiae Scientiarum Gottin-*

- gensis Recentiores, Classis Mathematicae* 5 (1819–1822):33–62. Available at https://resolver.sub.uni-goettingen.de/purl?PPN35283028X_0005_2NS.
- [25] Ku HH (1966) Notes on the use of propagation of error formulas. *Journal of Research of the National Bureau of Standards — C Engineering and Instrumentation* 70C:263–273. <https://doi.org/10.6028/jres.070C.025>
- [26] Joint Committee for Guides in Metrology (2011) *Evaluation of Measurement Data — Supplement 2 to the “Guide to the Expression of Uncertainty in Measurement” — Extension to any Number of Output Quantities* (International Bureau of Weights and Measures (BIPM), Sèvres, France). BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML, JCGM 102:2011 Available at https://www.bipm.org/documents/20126/2071204/JCGM_102_2011_E.pdf/6a3281aa-1397-d703-d7a1-a8d58c9bf2a5.
- [27] Meija J, Mester Z (2008) Uncertainty propagation of atomic weight measurement results. *Metrologia* 45(1):53–62. <https://doi.org/10.1088/0026-1394/45/1/008>
- [28] Valkiers S, Mana G, Fujii K, Becker P (2011) Si primary standards for the calibration of ion-current ratios in the molar-mass measurement of natural Si single crystals. *Metrologia* 48(2):S26–S31. <https://doi.org/10.1088/0026-1394/48/2/S04>
- [29] Lewandowski D, Kurowicka D, Joe H (2009) Generating random correlation matrices based on vines and extended onion method. *Journal of Multivariate Analysis* 100(9):1989–2001. <https://doi.org/10.1016/j.jmva.2009.04.008>
- [30] Ahrens LH (1954) The lognormal distribution of the elements (a fundamental law of geochemistry and its subsidiary). *Geochimica et Cosmochimica Acta* 5:49–73. [https://doi.org/10.1016/0016-7037\(54\)90040-X](https://doi.org/10.1016/0016-7037(54)90040-X)
- [31] Azzalini A, Capitanio A (2014) *The Skew-Normal and Related Families* (Cambridge University Press, Cambridge, UK). <https://doi.org/10.1017/cbo9781139248891>
- [32] Wise SA, Watters RL (2005) *Standard Reference Material 1d, Argillaceous Limestone* (Office of Reference Materials, National Institute of Standards and Technology, Department of Commerce, Gaithersburg, Maryland). Available at www.nist.gov/srm/.
- [33] Hillebrand WF (1900) *Some Principles and Methods of Rock Analysis*. *Bulletin of the United States Geological Survey*, Vol. 176 (Government Printing Office, Washington, D.C.). Available at <https://www.biodiversitylibrary.org/item/310844>.
- [34] Lehmann EL, Casella G (1998) *Theory of Point Estimation*. Springer Texts in Statistics (Springer-Verlag, New York, NY), 2nd Ed.
- [35] Wang J, et al. (2020) Final report of the CCQM-K145: toxic and essential elements in bovine liver. *Metrologia* 57(1A):08013–08013. <https://doi.org/10.1088/0026-1394/57/1a/08013>

- [36] Thompson M, Ellison SLR (2011) Dark uncertainty. *Accreditation and Quality Assurance* 16:483–487. <https://doi.org/10.1007/s00769-011-0803-0>
- [37] Possolo A, Koepke A, Newton D, Winchester MR (2021) Decision Tree for Key Comparisons. *Journal of Research of the National Institute of Standards and Technology* 126:126007. <https://doi.org/10.6028/jres.126.007>
- [38] Efron B, Tibshirani RJ (1993) *An Introduction to the Bootstrap* (Springer-Science+Business Media, Dordrecht, The Netherlands). <https://doi.org/10.1201/978-0-4292-4659-3>
- [39] Rukhin AL (2011) Maximum likelihood and restricted likelihood solutions in multiple-method studies. *Journal of Research of the National Institute of Standards and Technology* 116(1).
- [40] Rukhin AL (2009) Weighted means statistics in interlaboratory studies. *Metrologia* 46(3):323–331. <https://doi.org/10.1088/0026-1394/46/3/021>
- [41] Rukhin A, Biggerstaff B, Vangel M (2000) Restricted maximum likelihood estimation of a common mean and the Mandel-Paule algorithm. *Journal of Statistical Planning and Inference* 83:319–330. [https://doi.org/10.1016/S0378-3758\(99\)00098-1](https://doi.org/10.1016/S0378-3758(99)00098-1)
- [42] Viechtbauer W (2010) Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software* 36(3):1–48. <https://doi.org/10.18637/jss.v036.i03>
- [43] Koepke A, Lafarge T, Possolo A, Toman B (2017) Consensus building for interlaboratory studies, key comparisons, and meta-analysis. *Metrologia* 54(3):S34–S62. <https://doi.org/10.1088/1681-7575/aa6c0e>
- [44] Bernal AJ, et al. (2022) Molnupiravir for oral treatment of Covid-19 in nonhospitalized patients. *New England Journal of Medicine* 386(6):509–520. <https://doi.org/10.1056/NEJMoa2116044>
- [45] Bland JM, Altman DG (2000) Statistics Notes: The odds ratio. *British Medical Journal* 320:1468. <https://doi.org/10.1136/bmj.320.7247.1468>
- [46] Butler CC, et al. (2023) Molnupiravir plus usual care versus usual care alone as early treatment for adults with COVID-19 at increased risk of adverse outcomes (PANORAMIC): an open-label, platform-adaptive randomised controlled trial. *The Lancet* 401:281–293. [https://doi.org/10.1016/S0140-6736\(22\)02597-1](https://doi.org/10.1016/S0140-6736(22)02597-1)
- [47] Stoudt S, Pintar A, Possolo A (2021) Coverage intervals. *Journal of Research of the National Institute of Standards* 126:126004. <https://doi.org/10.6028/jres.126.004>
- [48] Joint Committee for Guides in Metrology (2008) *Evaluation of measurement data — Supplement 1 to the “Guide to the expression of uncertainty in measurement” — Propagation of distributions using a Monte Carlo method* (International Bureau of Weights and Measures (BIPM), Sèvres, France). BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML, JCGM 101:2008 Available at https://www.bipm.org/documents/20126/2071204/JCGM_101_2008_E.pdf/325dc aad-c15a-407c-1105-8b7f322d651c.

- [49] Meija J, Michałowska-Kaczmarczyk AM, Michałowski T (2016) Mohr's method challenge. *Analytical and Bioanalytical Chemistry* 408(7):1721–1722. <https://doi.org/10.1007/s00216-015-9273-2>
- [50] Lafarge T, Possolo A (2015) The NIST Uncertainty Machine. *NCSLI Measure Journal of Measurement Science* 10(3):20–27. <https://doi.org/10.1080/19315775.2015.11721732>
- [51] Skoog DA, West DM, Holler FJ, Crouch SR (2014) *Fundamentals of Analytical Chemistry* (Brooks/Cole, Belmont, CA), 9th Ed.
- [52] Johnson KA, Goody RS (2011) The original Michaelis constant: Translation of the 1913 Michaelis-Menten paper. *Biochemistry* 50(39):8264–8269. <https://doi.org/10.1021/bi201284u>
- [53] Michaelis L, Menten ML (1913) Die Kinetik der Invertinwirkung. *Biochemische Zeitschrift* 49:333–369.
- [54] Aledo JC (2021) Enzyme kinetic parameters estimation: A tricky task? *Biochemistry and Molecular Biology Education* 49:633–638. <https://doi.org/10.1002/bmb.21522>
- [55] Aledo JC (2022) renz: an R package for the analysis of enzyme kinetic data. *BMC Bioinformatics* 23:182. <https://doi.org/10.1186/s12859-022-04729-4>
- [56] Lineweaver H, Burk D (1934) The determination of enzyme dissociation constants. *Journal of the American Chemical Society* 56(3):658–666. <https://doi.org/10.1021/ja01318a036>
- [57] Lahiri SN (2003) *Resampling Methods for Dependent Data* (Springer-Verlag, New York, NY).
- [58] Hollander M, Wolfe DA, Chicken E (2014) *Nonparametric Statistical Methods* (John Wiley & Sons, Hoboken, NJ), 3rd Ed.
- [59] Allan JF, Batiza R, Lonsdale P (1987) Petrology and chemistry of lavas from seamounts flanking the East Pacific Rise axis, 21°N: implications concerning the mantle source composition for both seamount and adjacent EPR lavas. *Seamounts, Islands, and Atolls*, eds Keating BH, Fryer P, Batiza R, Boehlert GW (American Geophysical Union (AGU), Washington, D.C.), *Geophysical Monograph Series*, Vol. 43, pp 255–282. <https://doi.org/10.1029/GM043p0255>
- [60] Melson WG, O'Hearn T, Jarosewich E (2002) A data brief on the Smithsonian Abyssal Volcanic Glass Data File. *Geochemistry, Geophysics, Geosystems* 3(4):1–11. <https://doi.org/10.1029/2001GC000249>
- [61] Melson WG, O'Hearn T, Jarosewich E (2020) Smithsonian abyssal volcanic glass data file, National Museum of Natural History: Dataset. <https://doi.org/10.25573/data.11833128.v1>. Available at https://smithsonian.figshare.com/articles/dataset/Smithsonian_Abyssal_Volcanic_Glass_Data_File/11833128
- [62] Shapiro SS, Wilk MB (1965) An analysis of variance test for normality (complete samples). *Biometrika* 52(3,4):591–611. <https://doi.org/10.2307/2333709>
- [63] Miao W, Gel YR, Gastwirth JL (2006) A new test of symmetry about an unknown median. *Random Walk, Sequential Analysis and Related Topics: A Festschrift*

- in Honor of Yuan-Shih Chow*, eds Hsiung AC, Ying Z, Zhang CH (World Scientific Publishing Company, Singapore), pp 199–214. https://doi.org/10.1142/9789812772558_0013
- [64] González-Estrada E, Villaseñor JA (2016) A ratio goodness-of-fit test for the Laplace distribution. *Statistics & Probability Letters* 119:30–35. <https://doi.org/10.1016/j.spl.2016.07.003>
 - [65] Davison AC, Hinkley D (1997) *Bootstrap Methods and their Applications* (Cambridge University Press, Cambridge, UK). Available at www.epfl.ch/davison/BMA/.
 - [66] O’Hagan A (2008) The Bayesian Approach to Statistics. *Handbook of Probability: Theory and Applications*, ed Rudas T (Sage Publications, Thousand Oaks, CA). <https://doi.org/10.4135/9781452226620.n6>
 - [67] Possolo A, Toman B (2011) *Tutorial for Metrologists on the Probabilistic and Statistical Apparatus Underlying the GUM and Related Documents* (National Institute of Standards and Technology, Gaithersburg, MD). <https://doi.org/10.13140/RG.2.1.2256.8482>. Available at <https://www.itl.nist.gov/div898/possolo/TutorialWEBServer/TutorialMetrologists2011Nov09.xht>
 - [68] De Laeter JR, Böhlke JK, Bièvre PD, Hidaka H, Peiser HS, Rosman KJR, Taylor PDP (2003) Atomic weights of the elements: Review 2000. *Pure and Applied Chemistry* 75(6):683–800. <https://doi.org/10.1351/pac200375060683>. (IUPAC Technical Report)
 - [69] Gelman A, Carlin JB, Stern HS, Rubin DB (2003) *Bayesian Data Analysis* (Chapman & Hall / CRC, Boca Raton, FL), 2nd Ed.
 - [70] Robert CP, Casella G (2010) *Introducing Monte Carlo Methods with R* (Springer, New York, NY). <https://doi.org/10.1007/978-1-4419-1576-4>
 - [71] Neal RM (2011) MCMC using Hamiltonian Dynamics. *Handbook of Markov Chain Monte Carlo*, eds Brooks S, Gelman A, Jones G, Meng XL (Chapman & Hall/CRC, Boca Raton, FL), pp 113–162.
 - [72] Martin AD, Quinn KM, Park JH (2011) MCMCpack: Markov Chain Monte Carlo in R. *Journal of Statistical Software* 42(9):22. <https://doi.org/10.18637/jss.v042.i09>
 - [73] Cecelski CE, Toman B, Liu FH, Meija J, Possolo A (2022) Errors-in-variables calibration with dark uncertainty. *Metrologia* 59(4):045002. <https://doi.org/10.1088/1681-7575/ac711c>
 - [74] Pogue J, Yusuf S (1998) Overcoming the limitations of current meta-analysis of randomised controlled trials. *Lancet* 351:47–52. [https://doi.org/10.1016/S0140-6736\(97\)08461-4](https://doi.org/10.1016/S0140-6736(97)08461-4)
 - [75] Higgins JPT, Spiegelhalter DJ (2002) Being sceptical about meta-analyses: a Bayesian perspective on magnesium trials in myocardial infarction. *International Journal of Epidemiology* 31(1):96–104. <https://doi.org/10.1093/ije/31.1.96>

- [76] Lindley DV (1965) *Introduction to Probability and Statistics from a Bayesian Viewpoint — Part 2, Inference* (Cambridge University Press, Cambridge, UK). <https://doi.org/10.1017/CBO9780511662973>
- [77] Taylor BN, Kuyatt CE (1994) *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*. NIST Technical Note 1297 (National Institute of Standards and Technology, Gaithersburg, MD). Available at <https://physics.nist.gov/Pubs/guidelines/TN1297/tn1297s.pdf>.
- [78] Freedman DA (2009) *Statistical Models: Theory and Practice* (Cambridge University Press, New York, NY).
- [79] Nelson MA, Waters JF, Toman B, Lang BE, Rück A, Breitruick K, Obkircher M, Windust A, Lippa KA (2018) A new realization of SI for organic chemical measurement: NIST PS1 primary standard for quantitative NMR (Benzoic Acid). *Analytical Chemistry* 90:10510–10517. <https://doi.org/10.1021/acs.analchem.8b02575>
- [80] Shannon CE (1948) A mathematical theory of communication. *Bell System Technical Journal* 27:379–423. Available at <https://people.math.harvard.edu/~ctm/home/text/others/shannon/entropy/entropy.pdf>.
- [81] Shannon CE (1948) A mathematical theory of communication. *Bell System Technical Journal* 27:623–656. Available at <https://people.math.harvard.edu/~ctm/home/text/others/shannon/entropy/entropy.pdf>.
- [82] Possolo A, Merktas C, Bodnar O (2019) Asymmetrical uncertainties. *Metrologia* 56(4):045009. <https://doi.org/10.1088/1681-7575/ab2a8d>

Apéndice: Material complementario

Códigos de programación escritos en los lenguajes R y Stan. El entorno R para el modelado estadístico y el análisis de datos está disponible gratuitamente para su descarga en [The R Project for Statistical Computing](#), con material de apoyo en [Manuals](#).

Los abundantes comentarios intercalados dentro del código no tienen como objetivo enseñar los lenguajes R o Stan, con los que se supone que el lector ya está familiarizado. En cambio, sirven para aclarar el propósito del código que les continúa. Las porciones de líneas de código después de “##” (R) o de “//” (Stan) son comentarios ignorados en los cálculos.

Una vez que se hayan instalado los paquetes complementarios necesarios, los siguientes códigos R deberían ejecutarse si se procesan a través de la consola R, en [Emacs Speaks Statistics \(ESS\)](#) en [GNU Emacs](#), o en [RStudio Integrated Development Environment](#).

Las referencias bibliográficas aparecen con números entre corchetes, y su detalle aparece en *Referencias* a partir de la página 31. Cualquier referencia que no figure en dicho apartado se especifica en detalle a continuación del código.

Ejemplo 3E (Masa molar del iridio — Modelo de efectos de laboratorio)

```
## Chang et al. (1992) : Ir191/Ir193 = 0.59399, u(Ir191/Ir193) = 0.00103
## Walczyk et al. (1993) : Ir191/Ir193 = 0.59418, u(Ir191/Ir193) = 0.00037
## Zhu et al. (2017) : Ir191/Ir193 = 0.59290, u(Ir191/Ir193) = 0.00021

r.C = 0.59399; ur.C = 0.00103 ## mol/mol
r.W = 0.59418; ur.W = 0.00037
r.Z = 0.59290; ur.Z = 0.00021

## AME 2020 Atomic Mass Evaluation (Wang et al. 2021)
Ir191.x = 190960591.5 / 1e6 ## dalton
Ir191.u = 1.4 / 1e6
Ir193.x = 192962923.8 / 1e6
Ir193.u = 1.4 / 1e6

## Fracciones de cantidades isotópicas correspondientes a la relación
## isotópica C (Chang et al., 1992), estimación correspondiente de la
## masa molar del iridio y evaluación de la incertidumbre asociada
## mediante la aplicación del método de Monte Carlo

Ir191.C.x = r.C/(1+r.C)
Ir193.C.x = 1-Ir191.C.x
```

```
K = 1e7
rB = rnorm(K, mean=r.C, sd=ur.C)
Ir191.C.xB = rB/(1+rB)
Ir193.C.xB = 1 - Ir191.C.xB
Ir.C.x = (Ir191.C.x*Ir191.x + Ir193.C.x*Ir193.x)
Ir191.xB = rnorm(K, mean=Ir191.x, sd=Ir191.u)
Ir193.xB = rnorm(K, mean=Ir193.x, sd=Ir193.u)
Ir.C.u = sd(Ir191.C.xB*Ir191.xB + Ir193.C.xB*Ir193.xB)

## Fracciones de cantidades isotópicas correspondientes a la relación
## isotópica W (Walczyk et al., 1993) estimación correspondiente de la
## masa molar del iridio y evaluación de la incertidumbre asociada
## mediante la aplicación del método de Monte Carlo

Ir191.W.x = r.W/(1+r.W)
Ir193.W.x = 1-Ir191.W.x

K = 1e7
rB = rnorm(K, mean=r.W, sd=ur.W)
Ir191.W.xB = rB/(1+rB)
Ir193.W.xB = 1 - Ir191.W.xB
Ir.W.x = (Ir191.W.x*Ir191.x + Ir193.W.x*Ir193.x)
Ir191.xB = rnorm(K, mean=Ir191.x, sd=Ir191.u)
Ir193.xB = rnorm(K, mean=Ir193.x, sd=Ir193.u)
Ir.W.u = sd(Ir191.W.xB*Ir191.xB + Ir193.W.xB*Ir193.xB)

## Fracciones de cantidades isotópicas correspondientes a la relación
## isotópica Z (Zhu et al., 2017), estimación correspondiente de la
## masa molar del iridio y evaluación de la incertidumbre asociada
## mediante la aplicación del método de Monte Carlo

Ir191.Z.x = r.Z/(1+r.Z)
Ir193.Z.x = 1-Ir191.Z.x

K = 1e7
rB = rnorm(K, mean=r.Z, sd=ur.Z)
Ir191.Z.xB = rB/(1+rB)
Ir193.Z.xB = 1 - Ir191.Z.xB
Ir.Z.x = (Ir191.Z.x*Ir191.x + Ir193.Z.x*Ir193.x)
Ir191.xB = rnorm(K, mean=Ir191.x, sd=Ir191.u)
Ir193.xB = rnorm(K, mean=Ir193.x, sd=Ir193.u)
Ir.Z.u = sd(Ir191.Z.xB*Ir191.xB + Ir193.Z.xB*Ir193.xB)

## Estimaciones de la masa molar del iridio correspondientes a las
## determinaciones de las relaciones isotópicas de C, W y Z, y las
## incertidumbres estándar asociadas

options(digits=8)
c(Ir.C.x, Ir.W.x, Ir.Z.x)
## 192.21677 192.21662 192.21763

round(c(Ir.C.u, Ir.W.u, Ir.Z.u), 5)
```

```
## 0.00081 0.00029 0.00017

## Evaluación simplista de la incertidumbre oscura: exceso de
## variabilidad de los valores medidos, por encima de lo
## que las incertidumbres estándar asociadas a ellos sugieren que
## debería ser la variabilidad.

## Desviación estándar de las estimaciones de la masa molar de C, W y Z
s = sd(c(Ir.C.x, Ir.W.x, Ir.Z.x)) ## 0.00054
## Promedio geométrico de las incertidumbres estándar asociadas
g = exp(mean(log(c(Ir.C.u, Ir.W.u, Ir.Z.u)))) ## 0.00034
s/g ## 1.6

## Evaluación simplista de tau (incertidumbre oscura)
sqrt(s^2 - g^2) ## 0.00043 g/mol
sqrt(0.00054^2 - 0.00034^2) ## 0.00042

## Prueba Q de Cochran [20] de consistencia mutua de los resultados de
## medición y evaluación de DerSimonian-Laird de la incertidumbre oscura
require(metafor)
rma(yi=c(Ir.C.x, Ir.W.x, Ir.Z.x), sei=c(Ir.C.u, Ir.W.u, Ir.Z.u), method="DL")

## Prueba de heterogeneidad:  $Q(df = 2) = 9.61$ ,  $p\text{-val} = 0.0082$ 
## Como el valor de  $p$  es muy pequeño, la prueba Q de Cochran rechaza la
## hipótesis de consistencia mutua de los dos resultados de
## medición. Cuanto mayor sea el criterio Q (9.61 en este caso), más
## sólida será la evidencia contra la hipótesis de consistencia mutua.

## El valor  $p$  de la prueba estadística de la hipótesis de consistencia
## mutua es la probabilidad de observar un valor del criterio de
## prueba (Q en este caso) más alejado que el valor observado, cuando
## los resultados de las mediciones son en realidad mutuamente consistentes

## La estimación de la incertidumbre oscura depende del método de estimación
rma(yi=c(Ir.C.x, Ir.W.x, Ir.Z.x), sei=c(Ir.C.u, Ir.W.u, Ir.Z.u), method="DL")
## tau (square root of estimated  $\tau^2$  value): 0.0006
rma(yi=c(Ir.C.x, Ir.W.x, Ir.Z.x), sei=c(Ir.C.u, Ir.W.u, Ir.Z.u), method="REML")
## tau (square root of estimated  $\tau^2$  value): 0.0007
rma(yi=c(Ir.C.x, Ir.W.x, Ir.Z.x), sei=c(Ir.C.u, Ir.W.u, Ir.Z.u), method="ML")
## tau (square root of estimated  $\tau^2$  value): 0.0004

## Compárese con la estimación simplista anterior: 0.00041
```

Ejemplo 4B (Isótopos de silicio — Propagación de incertidumbre multivariable)

```
## Valkiers et al. [27, Tabla 4] -- Mezcla III

R2928.x = 0.0500657 ## mol/mol
R2928.u = 0.0000025

R3028.x = 0.0329035
R3028.u = 0.0000034

## Relaciones entre las relaciones isotópicas y la composición
## isotópica: x28, x29 y x30 indican fracciones de cantidad de los
## tres isótopos estables del silicio.
x28 = 1 / (1 + R2928.x + R3028.x)
x29 = R2928.x / (1 + R2928.x + R3028.x)
x30 = R3028.x / (1 + R2928.x + R3028.x)

## Las fracciones de cantidad son no negativas y suman 1
any(c(x28, x29, x30) < 0) ## FALSE
sum(c(x28, x29, x30)) ## 1

J = x28^2 * rbind(c(-1, 1), c(1+R3028.x, -R2928.x), c(-R3028.x, 1+R2928.x))
print(J)
## -0.85264410 0.85264410
## 0.88069907 -0.04268822
## -0.02805498 0.89533232

## RHO es la correlación entre las dos razones, R2928.x y R3028.x,
## que asumimos que es 0
rho = 0

sigmaR = rbind(c(R2928.u^2, rho*R2928.u*R3028.u),
               c(rho*R2928.u*R3028.u, R3028.u^2))
## Matriz de covarianza de las fracciones de cantidad (x28, x29, x30)
sigmaX = J %*% sigmaR %*% t(J)

cbind(round(c(x28=x28, x29=x29, x30=x30), 7),
      round(sqrt(diag(sigmaX)), 7))
## x28 0.9233873 3.6e-06 ## mol/mol
## x29 0.0462300 2.2e-06
## x30 0.0303827 3.0e-06
```

Ejemplo 4D (Modelando el peso de las colas)

```
## Laplace
require(extraDistr)
## La desviación estándar de la distribución de Laplace con escala
## sigma es sqrt(2)*sigma
```

```
sd(rlaplace(1e6, mu=0, sigma=1/sqrt(2))) ## 1
2 * plaplace(-3, mu=0, sigma=1/sqrt(2)) ## 0.0143696

## Gaussiana
2 * pnorm(-3, mean=0, sd=1) ## 0.002699796

## t de Student con 3 grados de libertad
## La desviación estándar de la t de Student con nu grados de libertad
## es sqrt(nu/(nu-2)) for nu > 2
sd(rt(1e6, df=3)/sqrt(3/(3-2))) ## 1
2 * pt(-3 * sqrt(3/(3-2)), df=3) ## 0.01384683
```

Ejemplo 4E (SiO₂ en piedra caliza — función de verosimilitud)

```
## Estimación de máxima verosimilitud de la media y la varianza de
## una distribución gaussiana, dada una muestra de tamaño 2 de dicha
## distribución. Las estimaciones de máxima verosimilitud de mu
## y sigma2 se pueden calcular analíticamente, pero a continuación
## utilizamos el enfoque general, que implica la optimización numérica
## de la función de verosimilitud

w.SiO2 = c(0.1811, 0.1818) ## g/g

negLogLik = function (theta, w, tol=sqrt(.Machine$double.eps))
{
  mu = theta[1]; sigma2 = theta[2]
  if (sigma2 < tol) { return(Inf)
  } else { return(-1*sum(dnorm(w, mean=mu, sd=sqrt(sigma2), log=TRUE))) }
}

SiO2.optim = optim(par=c(mean(w.SiO2), var(w.SiO2)), fn=negLogLik,
  method="Nelder-Mead", w=w.SiO2)
signif(SiO2.optim$par[1], 5) ## 0.18145 ## g/g
signif(SiO2.optim$par[2], 2) ## 1.2e-7 ## (g/g)^2

## La estimación de máxima verosimilitud (EMV) de la media es el
## promedio de los valores medidos, pero la EMV de la desviación
## estándar es diferente de la desviación estándar de la muestra.
signif(mean(w.SiO2), 5) ## 0.18145 ## g/g
signif(var(w.SiO2), 2) ## 2.4e-7 ## (g/g)^2
signif(mean((w.SiO2 - mean(w.SiO2))^2), 2) ## 1.2e-7 ## (g/g)^2

## NOTA: La estimación de máxima verosimilitud de sigma es la raíz
## cuadrada de la estimación de máxima verosimilitud de sigma2
```

Ejemplo 4F (Comparación interlaboratorio — Estimación de máxima verosimilitud)

```
## Los siguientes resultados de medición de la fracción de masa de
## níquel en hígado bovino difieren de los del informe final de
## CCQM-K145 [34] solo en que la incertidumbre informada por INMC fue
## redondeada a 2 cifras significativas

Ni = read.table(header=TRUE, text="
  lab      w      uw
  JSI 1.930 0.110
  INMC 1.940 0.064
  GLHK 1.942 0.092
  LNE 1.958 0.075
  NIST 1.984 0.020
  KRISS 1.993 0.033
  INACAL 2.010 0.060
  NMIA 2.020 0.050
  NIM 2.022 0.023
  NMIJ 2.050 0.020
  RISE 2.055 0.052
  NRC 2.070 0.050
  PTB 2.077 0.035
  LATU 2.080 0.059
  LGC 2.131 0.042
  UME 2.150 0.030
  HSA 2.180 0.080") ## mg/kg

## Estimación de máxima verosimilitud de la media mu y la
## incertidumbre oscura tau

require(bbmle)

negLogLik2 = function (mu, tau, w, uw)
{
  return(-1*sum(dnorm(w, mean=mu, sd=sqrt(uw^2 + tau^2), log=TRUE)))
}

## tauNaive es una estimación simplista de la dispersión "adicional" de
## los valores medidos, en comparación con las incertidumbres
## informadas, las cuales se resumen a través de su promedio geométrico

tauNaive = sqrt(var(Ni$w) - exp(mean(2*log(Ni$uw))))
Ni.mle = mle2(negLogLik2, start=list(mu=mean(Ni$w), tau=tauNaive),
              method="L-BFGS-B", lower=c(mu=min(Ni$w), tau=0),
              skip.hessian=FALSE, data=list(w=Ni$w, uw=Ni$uw))

summary(Ni.mle)
##      Estimate Std. Error
## mu  2.043      0.015      ## mg/kg
```

```
## tau 0.044    0.014

## La matriz de covarianza de las estimaciones de mu y tau se obtiene
## utilizando la aproximación de una gran muestra basada en la matriz hessiana
## de la función log-verosimilitud (invocandola especificando
## "skip.hessian=FALSE" arriba), incluso si solo se tienen 17 resultados
## de medición

Ni.mle@vcov
##              mu              tau
## mu  2.384880e-04 -7.760112e-06
## tau -7.760112e-06  2.027808e-04

## Suponiendo que el modelo gaussiano es apropiado para estos
## valores medidos, el estimador de máxima verosimilitud restringida
## (REML, por sus siglas en inglés) es generalmente preferible al EMV.

require(metafor)
rma(yi=w, sei=uw, data=Ni, method="REML", level=95, test="knha")

## tau = 0.0463 ## mg/kg
## Test for Heterogeneity: Q(df = 16) = 40.1205, p-val = 0.0007
## estimate      se    ci.lb    ci.ub
##  2.0428  0.0159    2.0091    2.0765 ## mg/kg

## El intervalo cuyos puntos finales son ci.lb y ci.ub es un intervalo
## de cobertura del 95 % para el valor real de la fracción de masa de
## níquel en el material, obtenido utilizando el ajuste de
## Knapp-Hartung, invocado especificando 'test="knha"' arriba.

## Knapp, G., & Hartung, J. (2003). Improved tests for a random
## effects meta-regression with a single covariate. Statistics in
## Medicine, 22(17), 2693-2710. DOI 10.1002/sim.1482

## IntHout, J., Ioannidis, J. P. A., & Borm, G. F. (2014). The
## Hartung-Knapp-Sidik-Jonkman method for random effects meta-analysis
## is straightforward and considerably outperforms the standard
## DerSimonian-Laird method. BMC Medical Research Methodology, 14,
## 25, DOI 10.1186/1471-2288-14-25
```

Ejemplo 4G (Molnupiravir — Estimación de máxima verosimilitud de la eficacia del tratamiento)

```
## Al final del ensayo informado por Butler et al. (2023), 105 de los
## 12529 pacientes que habían estado tomando molnupiravir sumado a la
## atención habitual, y 98 de 12525 pacientes del grupo de atención
## habitual, murieron o fueron hospitalizados
```

```
M.events    = 48
```

```
M.patients = 709

P.events = 68
P.patients = 699

M.p = M.events/M.patients ## 0.06770099
P.p = P.events/P.patients ## 0.09728183

OR = (M.p/(1-M.p)) / (P.p/(1-P.p)) ## 0.6738453
logOR = log(OR) ## -0.3947547

logOR.u = sqrt(1/(M.patients*(M.p*(1-M.p))) + 1/(P.patients*(P.p*(1-P.p))))
## 0.1965626

## Nótese que OR.u arroja el mismo valor que la aproximación clásica
## para el error estándar del logaritmo de la razón de posibilidades
sqrt(1/M.events + 1/(M.patients-M.events) +
      1/P.events + 1/(P.patients-P.events)) ## 0.1965626

## Z-score
logOR / logOR.u ## -2.00829
pnorm(logOR / logOR.u, mean=0, sd=1) ## 0.02230626

## Un análisis alternativo de los resultados del ensayo clínico se
## basa en la regresión logística, que es un tipo particular de modelo
## lineal generalizado.

## El "estado" de un paciente describe si el paciente está tomando
## molnupiravir (M) o un placebo (P), y el "resultado" correspondiente
## es 1 si el paciente ha sido hospitalizado o ha fallecido, o 0 en
## caso contrario

status = relevel(factor(c(rep("M", M.patients),
                           rep("P", P.patients))), ref="P")
outcome = c(rep(1, M.events), rep(0, M.patients-M.events),
             rep(1, P.events), rep(0, P.patients-P.events))
z = data.frame(status, outcome)

z.glm = glm(outcome ~ status, family=binomial, data=z)
summary(z.glm)
##               Estimate Std. Error.
## (Intercept)  -2.2278      0.1276
## statusM      -0.3948      0.1966

## Log odds ratio      = -0.3948
## Standard uncertainty = 0.1966

## Razón de posibilidades de hospitalización o muerte entre pacientes que
## recibieron placebo

oddsP = exp(-2.2278)
```

```
## Probabilidad de hospitalización o muerte entre pacientes que
## recibieron placebo
probP = oddsP/(1+oddsP) ## 0.0973

## Razón de posibilidades
oddsRatioMP = exp(-0.3948)

## Razón de posibilidades de hospitalización o muerte entre pacientes que
## recibieron molnupiravir

oddsM = oddsRatioMP * oddsP

## Probabilidad de hospitalización o muerte entre pacientes que
## recibieron molnupiravir
probM = oddsM/(1+oddsM) ## 0.0677
```

Ejemplo 4H (Estándar de calibración de cadmio — Propagación de incertidumbre mediante Monte Carlo)

```
## Ejemplo A1 de la Guía CG-4 de EURACHEM/CITAC: preparación de un
## patrón de calibración para la determinación de cadmio mediante
## espectroscopia de absorción atómica

m.x = 100.28      ## mg
P.x = 0.9999      ## g/g
V.x = 100.0       ## mL

m.u   = 0.05      ## mg
P.u   = 0.000058  ## g/g
V.u   = 0.07      ## mL

K = 1e7

## Distribución gaussiana para m truncada en 0 mg
require(truncnorm)
m = rtruncnorm(K, a=0, b=+Inf, mean=m.x, sd=m.u)

## Distribución beta para P
alpha = P.x * ( P.x * (1 - P.x)/P.u^2 - 1)
beta = (1 - P.x) * alpha / P.x
curve(dbeta(x, alpha, beta), n=2^12, from=0.9995, to=1)
P = rbeta(K, shape1=alpha, shape2=beta)

## Distribución triangular para V
require(triangle)
a = V.x - V.u*sqrt(6) ## 99.82854
b = V.x + V.u*sqrt(6) ## 100.1715
c = (a + b)/2
curve(dtriangle(x, a=a, b=b, c=c), from=a, to=b)
```

```
V = rtriangle(n=K, a=a, b=b, c=c)

## Calcular la concentración de cadmio
cCd = m*P/V

## Siempre examine los resultados gráficamente antes de resumirlos,
## para garantizar que los resúmenes elegidos sean significativos y
## adecuados al propósito

plot(density(cCd), bty="n", lwd=3, col="DodgerBlue", main="",
      xlab=expression(italic(c)(plain(Cd))~{}/{}~(plain(mg)/plain(mL))),
      ylab="Prob. Density")
curve(dnorm(x, mean=mean(cCd), sd=sd(cCd)), add=TRUE, n=2^12, lwd=1, col="Red")
## NOTA: La distribución de probabilidad de cCd se desvía
## significativamente de la forma gaussiana

## Resumir resultados (media, incertidumbre estándar, incertidumbre
## expandida del 95 %)

round(c(mean(cCd), sd(cCd)), 4) ## 1.0027 0.0009 ## mg/mL
round(diff(quantile(cCd, probs=c(0.025, 0.975)))/2, 4) ## 0.0017 ## mg/mL
```

Ejemplo 4I (Titulación argentométrica — Propagación de la incertidumbre)

```
V0.x = 100/1000 ## L      (100 mL de solución de NaCl)
Ca.x = 100/1000 ## mol/L (100 mmol/L de AgNO3 -- Titulante)
Va.x = 10/1000  ## L      (10 mL de AgNO3 solución consumida)
Ck.x = 1.75/1000 ## mol/L (1.75 mmol/L de K2CrO4 -- Indicador)

## V0 (volumen de solución de NaCl) medido con una pipeta volumétrica
## de clase A, cuyo error de medición se modela utilizando una
## distribución triangular simétrica con un semiancho de 0.08 mL.

K = 1e7
require(triangle)
V0 = V0.x + rtriangle(K, a=-0.08/1000, b=+0.08/1000, c=0)

## Va (volumen de AgNO3 consumido para alcanzar el punto final) medido
## con una bureta electrónica con una resolución de 0.01 mL, cuyo
## error de medición se modela utilizando una distribución triangular
## simétrica con un semiancho de 0.01 mL. Va también se ve afectado por
## el error en la determinación en alcanzar el punto final, que
## se modela utilizando una distribución gaussiana centrada en 0 mL y
## con una desviación estándar de 0.025 mL.

Va = Va.x + rtriangle(K, a=-0.01/1000, b=+0.01/1000, c=0) +
      rnorm(K, mean=0, sd=0.025/1000)

## Se asume que la incertidumbre asociada con la concentración de
```

```
## AgNO3 en el titulante es insignificante.

Ca = Ca.x

C0 = Ca * Va/V0
C0.x = mean(C0)
C0.u = sd(C0)
c(C0.x, C0.u) ## 0.010000 0.000026 ## mol/L

## Productos de solubilidad de AgCl (Ksp1) y Ag2CrO4 (Ksp2)

Ksp1 = 10^(-9.75)
Ksp2 = 10^(-11.9)

## VE es el valor real de Va (volumen añadido de titulante) que tiene
## en cuenta estos productos de solubilidad

f = function (VE, Ca, Ksp1, Ksp2, V0, C0, Ck)
{
  (Ksp1 / sqrt(Ksp2*(V0+VE)/(Ck*V0))) - sqrt(Ksp2*(V0+VE)/(Ck*V0)) -
    (C0*V0 - Ca*VE) / (V0 + VE)
}

VE = uniroot(f, interval=c(0, 0.05), tol=sqrt(.Machine$double.eps),
  C=Ca.x, Ksp1=Ksp1, Ksp2=Ksp2, V0=V0.x, C0=C0.x, Ck=Ck.x)$root
## 0.010024 ## L

## Estimación simplista de la concentración de NaCl
C0.x = Va.x * Ca.x / V0.x

## Estimación de la concentración de NaCl que tiene en cuenta
## los productos de solubilidad de AgCl y Ag2CrO4
C0.STAR = VE * Ca.x / V0.x

## Error relativo de la estimación simplista de la concentración de NaCl
(C0.STAR - C0.x) / C0.x ## 0.0024
```

Ejemplo 4J (Michaelis-Menten — Propagación de incertidumbre mediante Monte Carlo)

```
## La enzima beta-galactosidasa cataliza la hidrólisis de
## o-nitrofenil-beta-galactopiranosido (ONPG) para producir galactosa y
## o-nitrofenol a una velocidad, v, que depende de la concentración
## de ONPG, c, como se describe en la ecuación de Michaelis-Menten
## [51,52,53,54,55]

c = c(0.05, 0.1, 0.25, 0.5, 1, 2.5, 5, 8, 20, 30) ## mmol/L
v = c(3, 5.2, 14.4, 30.3, 49, 86.2, 112.6, 136.2, 170, 177.7) ## umol/L/min
```

```
v.nls = nls(v ~ vMax*c/(KM + c), start=list(vMax=195, KM=3),
           nls.control(maxiter=5000, tol=1e-07, minFactor= 1/2^12),
           data=data.frame(c=c, v=v))

summary(v.nls)
##      Estimate Std. Error
## vMax 195.3556      3.2342 ## umol/L/min
## KM    3.2662      0.1837 ## mmol/L
## Residual standard error: 3.085 on 8 degrees of freedom ## umol/L/min

vMaxHAT = coef(v.nls)[1]
KMHAT = coef(v.nls)[2]
sigmaHAT = summary(v.nls)$sigma ## 3.085365

## Evaluación de Monte Carlo de las incertidumbres asociadas con las
## estimaciones de vMax y KM, las cuales básicamente reproducen, de manera
## independiente, los valores enumerados en "Estimate Std. Error" arriba

n = length(c)
K = 10000
theta = array(rep(NA, K*2), dim=c(K,2))
for (k in 1:K)
{
  if (k %% 1000 == 0) {cat(k, "of", K, "\n")}
  vB = vMaxHAT*c/(KMHAT+c) + sigmaHAT*rt(n, df=8)/sqrt(8/(8-2))

  vB.nls = try(nls(v ~ vMax * c/(KM+c),
                  start=list(vMax=195, KM=3),
                  nls.control(maxiter=5000, tol=1e-07, minFactor= 1/2^12),
                  data=data.frame(c=c, v=vB)))
  if (inherits(vB.nls, "try-error")) {next}
  } else {theta[k,] = coef(vB.nls)}
}

iNA = apply(theta, 1, function(x){is.na(x)})
if (sum(iNA) > 0) {theta = theta[!iNA,]}

## Siempre examine los resultados gráficamente antes de resumirlos,
## para garantizar que los resúmenes elegidos sean significativos y
## adecuados al propósito

par(mfrow=c(1,2))
plot(density(theta[,1], n=2^12, adjust=1.5), lwd=2, col="DodgerBlue",
     main="", bty="n",
     xlab=expression(italic(v) [plain(MAX)] ~{}/{}~##
                     ((mu*plain(mol)/plain(L))/plain(min))),
     ylab="Prob. Density")
plot(density(theta[,2], n=2^12, adjust=1.5), lwd=2, col="Tomato",
     main="", bty="n",
     xlab=expression(italic(K) [plain(M)] ~{}/{}~(plain(mmol)/plain(L))),
     ylab="Prob. Density")
par(mfrow=c(1,1))
```

```
apply(theta, 2, sd) ## 3.2 0.18 ## umol/L/min, mmol/L
## Compárese con los errores estándar enumerados anteriormente después de
## "summary(v.nls)"
```

Ejemplo 4K (Basaltos del Monte Submarino Verde – “Bootstrap” no paramétrico para la mediana de la fracción de masa de TiO_2)

```
## Evaluación de la incertidumbre de la mediana de determinaciones
## replicadas de la fracción de masa de dióxido de titanio en muestras
## de vidrio volcánico del Monte Submarino Verde, en el Océano
## Pacífico, a unos 435 km al oeste de Puerto Vallarta, México [58,59,60]

w = c(15.2, 14.2, 15.1, 15, 12.4, 12.2, 12.7, 14.9, 14, 14, 14.4,
      20.1, 20.5, 11.2, 13.7, 14.1, 12.7, 12.8, 17.9, 11.7, 13.1,
      13.5) ## / (mg/g)

## No es probable que la muestra sea de una distribución gaussiana
shapiro.test(w) ## p-value = 0.003052
## La muestra es consistente con una distribución simétrica
require(symmetry)
symmetry_test(w, stat="MGG", B=25000)$p.value ## 0.32

## Adopte el modelo de error de medición del Ejemplo 3D, pero con los
## errores de medición de Laplace. La mediana es el estimador de
## máxima verosimilitud de la media de la distribución de Laplace

median(w) ## 14.0 mg/g

## La evaluación de la incertidumbre asociada con la mediana a través
## de un remuestreo de bootstrap no paramétrico requiere un conjunto
## "grande" de determinaciones de réplicas; aquí tenemos únicamente 30 pero,
## como veremos, aún es suficiente para obtener una evaluación
## aceptablemente precisa

m = length(w)
muB = replicate(1e5, { median(sample(w, size=m, replace=TRUE)) })
round(sd(muB), 2) ## 0.41 mg/g
round(quantile(muB, c(0.025,0.975)), 1) ## 12.9, 14.7 mg/g

## Para validar los resultados del bootstrap, también obtenemos otra
## evaluación independiente de la incertidumbre asociada con la
## mediana, calculando un intervalo de cobertura del 95 % para la
## mediana verdadera a través de la inversión de la prueba de signos

## La prueba de signos de la hipótesis de que la mediana es un valor
## específico, digamos 15 mg/g, se basa en la cantidad de signos
## positivos de las diferencias {w[i] - 15 mg/g}.
```

```
## Si 15 mg/g fuera de hecho la mediana verdadera, entonces cada
## determinación tendría la misma probabilidad de estar por encima o
## por debajo de 15 mg/g, y la cantidad de diferencias positivas
## tendría una distribución binomial basada en 22 ensayos con un
## 50 % de probabilidad de "éxito" en cada ensayo.

## Dado que 5 determinaciones son mayores que 15 mg/g, la probabilidad
## de observar estas o menos diferencias positivas es pbinom(5,
## size=22, prob=0.5) = 0.0085. Sin embargo, estamos probando si hay
## muy pocas diferencias positivas o muy pocas diferencias
## negativas. Por lo tanto, el valor p de la prueba es  $2 * 0.0085 = 0.017$ .

## Ahora, supongamos que elegimos rechazar la hipótesis de que la
## mediana verdadera es M cuando el valor p de la prueba es 0.05 o
## menor. El intervalo de valores de M que la prueba no rechaza es un
## intervalo de cobertura del 95 % para la mediana verdadera:
## la determinación de este intervalo se llama "invertir la prueba"
## y se puede hacer usando la función de R "SIGN.test" definida en el
## paquete de R BSDA

require(BSDA)
round(SIGN.test(w, conf.level=0.95)$conf.int, 1)
## 12.8, 14.9 mg/g

## En muchas aplicaciones, la incertidumbre estándar de la mediana se
## evalúa como  $\text{mad}(w) * \sqrt{\pi/2} / \sqrt{\text{length}(w)} = 0.46 \text{ mg/g}$ , que en este
## caso resulta ser  $0.46/0.41 = 1.12$  veces más grande (pero
## también podría haber sido demasiado pequeña).

## Esta fórmula, que se utiliza de forma incorrecta, es válida solo
## cuando n es grande y los datos son una muestra de una distribución
## gaussiana. Pero si los datos son una muestra de una distribución
## gaussiana, entonces la mejor estimación es el promedio de las
## determinaciones, no la mediana, cuya incertidumbre estándar
## correspondiente es  $s/\sqrt{n}$ .
```

Ejemplo 4N (Masa molar del iridio — Métodos de Monte Carlo basados en cadenas de Markov (MCMC))

```
## Recomendamos que, en lugar de implementar los procedimientos
## estadísticos mencionados en esta breve guía, los usuarios confíen
## en implementaciones ya probadas y confiables disponibles en un software
## confiable y preexistente, que se puede conseguir en muchas fuentes
## para R y para la mayoría de los demás entornos de modelado
## estadístico y análisis de datos. Esta recomendación es
## especialmente importante para todos y cada uno de los
## procedimientos que involucran muestreo de Monte Carlo con cadena de
## Markov
```

```
Ir.M = c(C=192.2168, W=192.2166, Z=192.2176) ## g/mol
Ir.uM = c(C= 0.0008, W= 0.0003, Z= 0.0002) ## g/mol

require(MCMCpack)

require(extraDistr)
median(rhcauchy(1e6, sigma=0.0004)) ## 0.0003997044

## La función definida por el usuario "lognum" calcula el logaritmo
## del numerador del teorema de Bayes, que comprende el producto de la
## función de densidad de probabilidad a priori de los parámetros y de
## la función de verosimilitud

lognum = function (theta, Ir.M, Ir.uM)
{
  M = theta[1]; tau = theta[2]
  ## Distribución a priori para la masa molar del iterbio
  s = dnorm(M, mean=192.217, sd=0.0015, log=TRUE)
  ## La distribución a priori para la incertidumbre oscura, tau,
  ## tiene una mediana de 0.0004 g/mol
  s = s + dhcauchy(tau, sigma=0.0004, log=TRUE)
  ## Verosimilitud gaussiana
  s = s + dnorm(Ir.M["C"], mean=M,
                sd=sqrt(tau^2 + Ir.uM["C"]^2), log=TRUE)
  s = s + dnorm(Ir.M["W"], mean=M,
                sd=sqrt(tau^2 + Ir.uM["W"]^2), log=TRUE)
  s = s + dnorm(Ir.M["Z"], mean=M,
                sd=sqrt(tau^2 + Ir.uM["Z"]^2), log=TRUE)
  return(s)
}

Ir.mcmc =
  MCMCmetrop1R(fun=lognum,

  ## Valores iniciales donde comenzará la cadena MCMC, extraídos
  ## aleatoriamente de las distribuciones a priori correspondientes
  theta.init=c(rnorm(1, mean=192.217, sd=0.0015),
               rhcauchy(1, sigma=0.0004)),

  ## La cantidad de pasos de arranque ('`burn-in'', por sus siglas
  ## en inglés) debe ser aproximadamente la mitad de la cantidad de
  ## pasos de muestro MCMC reales. La entrada "thin" especifica que
  ## solo se conservará 1 de cada 25 muestras generadas por MCMC,
  ## para reducir las correlaciones entre valores consecutivos de
  ## los parámetros
  burnin=250000, mcmc=250000, thin=25,

  ## El parámetro "tune" (que puede ser un escalar o un
  ## vector con tantos componentes como parámetros haya) debe
  ## elegirse de manera que se alcance una tasa de aceptación de
  ## Metropolis de aproximadamente 0.3; la función reporta esta tasa
  ## de aceptación al finalizar, y "tune" debe ajustarse según sea
```

```

## necesario, y la ejecución debe repetirse hasta que se alcance
## una tasa de aceptación satisfactoria
tune=c(4, 1.5),

## Si hay restricciones en el intervalo en los parámetros, como las hay
## aquí para tau, el cual debe ser no negativo, entonces el método
## de optimización debe ser "L-BFGS-B"
optim.method="L-BFGS-B",

## Las siguientes entradas definen los límites inferior y superior
## para los valores "válidos" de los parámetros
optim.lower=c(-Inf, 0), optim.upper=c(Inf, Inf),
optim.control=list(fnscale=-1, trace=0, REPORT=10,
                  maxit=5000, ndeps=c(0.5e-4,0.5e-5)),
## Valores necesarios para calcular el numerador del teorema de Bayes
Ir.M=Ir.M, Ir.uM=Ir.uM)

dimnames(Ir.mcmc)[[2]] = c("M", "tau")
dim(Ir.mcmc) ## 10000      2

summary(Ir.mcmc)

## Iterations = 250001:499976
## Thinning interval = 25
## Number of chains = 1
## Sample size per chain = 10000
##
## 1. Empirical mean and standard deviation for each variable,
##    plus standard error of the mean:
##
##           Mean          SD Naive SE Time-series SE
## M    1.922e+02 0.0004173 4.173e-06      4.173e-06
## tau  5.598e-04 0.0004138 4.138e-06      6.654e-06
##
## 2. Quantiles for each variable:
##
##           2.5%       25%       50%       75%       97.5%
## M    1.922e+02 1.922e+02 1.922e+02 1.922e+02 1.922e+02
## tau  6.064e-05 2.979e-04 4.638e-04 7.060e-04 1.652e-03

M.TILDE = Ir.mcmc[, "M"]
tau.TILDE = Ir.mcmc[, "tau"]

par(mfrow=c(1,2))
plot(density(M.TILDE, adj=1.5, from=192.2145, to=192.2195), main="",
     xlab=expression(italic(M) [plain(Ir)] ~{} / {} ~ (plain(g) / plain(mol))),
     ylab="Posterior Prob. Density", lwd=2, col="DodgerBlue")
plot(density(tau.TILDE, adj=1.5, from=0, to=0.002), main="",
     xlab=expression(tau ~{} / {} ~ (plain(g) / plain(mol))),
     ylab="Posterior Prob. Density", lwd=2, col="Tomato")
par(mfrow=c(1,1))

```

```
round(c(M=mean(M.TILDE), "u(M)"=sd(M.TILDE)), 4)
##      M      u(M)
## 192.2171 0.0004 ## g/mol

round(median(tau.TILDE), 4) ## 0.0005 ## g/mol
```

Ejemplo 40 (Nitrato en agua de mar — Caracterización bayesiana de la precisión de medición)

```
## Cuatro determinaciones espectrofotométricas de la fracción másica
## de nitritos en una muestra de agua de mar obtenida en condiciones
## de repetibilidad de medición utilizando el método de Griess [11]

w = c(0.1514, 0.1523, 0.1531, 0.1545) ## mg/kg

## La incertidumbre relativa de medición podría ser tan baja como
## 0.33 % o tan alta como 3 %, ubicándose en algún punto entre estos
## dos valores con una probabilidad del 90 %

## Calibración de una distribución a priori gamma para sigma, que refleja
## este vago conocimiento previo sobre su valor verdadero

q = c(0.0033, 0.03)*median(w)
p = c(0.05, 0.95)
g = function (theta, q, p)
{
  if (any(theta < 0)) { return(Inf)
  } else {
    alpha = theta[1]; beta = theta[2]
    return(sum((qgamma(p, shape=alpha, rate=beta) - q)^2))
  }
}

optim(par=c(2, 500), fn=g, gr=NULL, q=q, p=p, method="Nelder-Mead")$par
## Shape = 2.619, Rate = 1248 kg/mg

## Modelo bayesiano implementado utilizando el lenguaje de
## programación probabilística Stan
## [B. Carpenter et al., 2017, "Stan: A Probabilistic
## Programming Language." Journal of Statistical Software 76(1):
## 1-32. DOI 10.18637/jss.v076.i01]

require(rstan)
options(mc.cores = parallel::detectCores())
rstan_options(auto_write=TRUE, javascript=FALSE)

model =
  "data
  {
    // Las determinaciones se almacenan en un vector de longitud 4
```

```
    real w[4];
  }
  parameters
  {
    // Fracción de masa media real de nitritos, que debe ser no negativa
    real<lower=0> omega;

    // Desviación estándar verdadera de los errores de medición,
    // la cual no debe ser negativa, según el método de Griess
    real<lower=0> sigma;
  }
  model
  {
    // La distribución previa elegida para la fracción de masa media
    // real de nitritos en la muestra es gaussiana truncada, con una
    // media de 0 mg/kg y una desviación estándar de 1 mg/kg (ambas
    // antes del truncamiento), truncada en 0 mg/kg dado que omega
    // debe ser no negativo
    omega ~ normal(0, 1);

    // Distribución previa elegida para la desviación estándar de
    // los errores de medición que tiene en cuenta el conocimiento
    // de la repetibilidad del método
    sigma ~ gamma(2.619, 1248);

    // Función de verosimilitud correspondiente al supuesto de que
    // los valores medidos son como una muestra de una distribución
    // gaussiana, cuya media es el valor real de la fracción de masa
    // y cuya desviación estándar es el valor real de la
    // repetibilidad
    w ~ normal(omega, sigma);
  }
}"

fit = stan(model_code = model, data = list(w=w),

  ## Se recomienda utilizar la mitad inicial de los pasos
  ## (el número total de pasos solicitados en este caso es 500000)
  ## para el ajuste inicial y ajustar automáticamente los samplers
  warmup=250000, iter=500000,

  ## Ejecute cuatro muestreadores de cadena de Markov
  ## independientemente uno del otro, en paralelo, utilizando
  ## cuatro núcleos de CPU (si la CPU que se utiliza no tiene
  ## 4 núcleos, entonces "cores=4" debe eliminarse de la
  ## siguiente línea)
  chains=4, cores=4,

  ## Mantenga solo los conjuntos de valores de parámetros
  ## múltiples de 25 muestreados, para reducir las
  ## correlaciones (serial) entre ellos
  thin=25
```

```
print(fit, digits=4)

## Los valores de Rhat deben ser muy cercanos a 1: de lo contrario, el
## modelo debería ajustarse utilizando un mayor número de
## iteraciones. Los valores de n_eff son tamaños de muestra efectivos
## en los que se basa el siguiente resumen

##          mean      sd    2.5%    50%   97.5%  n_eff    Rhat
## omega  0.1528  0.0010  0.1509  0.1528  0.1548  40478  0.99997
## sigma  0.0018  0.0007  0.0008  0.0016  0.0036  39757  0.99999

fit.post = extract(fit)

## La incertidumbre estándar (a posteriori) asociada con la estimación
## bayesiana de la fracción de masa real es 1,45 veces mayor que la
## evaluación convencional de tipo A calculada según la GUM

sd(fit.post$omega) / (sd(w)/sqrt(length(w))) ## 1.45
```

Ejemplo 4P (Molnupiravir — Reconsideración escéptica mediante modelado bayesiano)

```
## Media y desviación estándar de la distribución a priori escéptica
## para la razón de posibilidades logarítmica
muP = 0; sigmaP = 0.175
## Razón de posibilidades logarítmica observada en el ensayo MOVE-OUT
## informado por Bernal et al. (2022) e incertidumbre estándar
## asociada (del EJEMPLO 4G)
logOR = -0.395; logOR.u = 0.197

## Media y desviación estándar a posteriori
muQ = (logOR/logOR.u^2 + muP/sigmaP^2) / (1/logOR.u^2 + 1/sigmaP^2)
sigmaQ = sqrt(1/(1/logOR.u^2 + 1/sigmaP^2))

c(muQ, sigmaQ) ## -0.174  0.131
## Probabilidad a posteriori de que el molnupiravir no sea más eficaz que
## el placebo
1-pnorm(0, mean=muQ, sd=sigmaQ) ## 0.091
```