

LA INCERTIDUMBRE EXPANDIDA Y LOS GRADOS DE LIBERTAD: UN ANÁLISIS COMPARATIVO ENTRE EL MÉTODO RECOMENDADO POR LA GUM Y MÉTODOS SIMPLIFICADOS

Wolfgang A. Schmid

Centro Nacional de Metrología, División de Óptica y Radiometría
km 4,5 carretera a Los Cués, Municipio del Marqués, Querétaro
Tel.: (442) 211 0554, Fax: 211 0553, e-mail wschmid@cenam.mx

Resumen: Una determinación rigurosa de la incertidumbre expandida de la medición, correspondiente a un nivel definido de confianza, requiere la consideración de los grados de libertad. Este concepto implica aumentar los valores de incertidumbre, cuando fueron obtenidos con base en informaciones limitadas, p. ej. un número limitado de mediciones.

Se presenta una comparación numérica entre el método recomendado por la "Guía BIPM/ISO para la expresión de la incertidumbre en las mediciones" - GUM, basándose en el número efectivo de los grados de libertad, y varios métodos simplificados que frecuentemente se aplican en laboratorios de metrología y se dan recomendaciones sobre el uso de los métodos.

INTRODUCCIÓN

El resultado de la estimación de la incertidumbre de una medición, generalmente se presenta en terminos de un valor U_p , llamado incertidumbre expandida, que corresponde a un nivel de confianza p . El nivel de confianza es la probabilidad de que el intervalo de $\pm U_p$ alrededor del mejor estimado para el mensurando contenga al valor verdadero (del mensurando).

En el campo de la metrología, a menudo se usa un nivel de confianza $p \cong 95\%$. En el caso frecuente de que la dispersión de los valores del mensurando corresponda a una distribución normal (distribución gaussiana), U_p se obtiene expandiendo la incertidumbre estándar u por un factor $k = 2$:

$$U_{95\%} = k \cdot u = 2 \cdot u \quad (1)$$

Para ser preciso, $2u$ corresponde a un nivel de confianza de 95,45%.

Generalmente la misma incertidumbre estándar no puede ser obtenida con una certeza absoluta, es decir u misma tiene "incertidumbre".

Por ejemplo, con los métodos de evaluación tipo A, que estiman la incertidumbre a partir de la dispersión de mediciones repetidas, la certeza en la estimación no sólo del resultado mismo de la medición, sino también de su incertidumbre, respectivamente desviación estándar, depende del número n de mediciones realizadas. Mientras la

incertidumbre u se estima con la desviación estándar experimental $s(\bar{q})$ de la media \bar{q} de las n mediciones repetidas, la "incertidumbre de la incertidumbre u " se puede estimar mediante la desviación estándar de los valores obtenidos para $s(\bar{q})$, repitiendo el conjunto de las n mediciones varias veces. Si los datos siguen una distribución normal, teóricamente se espera la siguiente relación (ver [1], E.4):

$$\frac{s(s(\bar{q}))}{s(\bar{q})} \cong \frac{1}{\sqrt{2 \cdot (n-1)}} \quad (2)$$

Mientras $s(\bar{q})$ es el valor que se utiliza para estimar la incertidumbre estándar, $s(\bar{q}) = u$ es la incertidumbre estándar "verdadera" o teórica, que se obtendría si se determinara la dispersión verdadera de los \bar{q} por un número infinito de mediciones. Aplicando (2) se observa que por ejemplo en el caso de $n = 3$ repeticiones, la estimación de u mediante $s(\bar{q})$ tiene una incertidumbre de alrededor de 50% y todavía de 24% en el caso de $n = 10$.

En el caso de un número pequeño ($n < 10$) de mediciones, algo muy común en mediciones industriales, se debe considerar este efecto, sustituyendo (1) por:

$$U_p = t_p(n) \cdot u \quad (3)$$

El factor $t_p(n)$ se obtiene a partir de la distribución t de Student para el número n de grados de libertad.

Los valores de $t_p(n)$ se encuentran en tablas [1, 2]. En el caso de la incertidumbre por repetibilidad, n es el número de mediciones menos uno:

$$n = n - 1 \quad (4)$$

La ec. (3) refleja el hecho que la distribución de los \bar{q} , dividida entre $s(\bar{q})$, sigue una distribución t de Student, que, en el caso de pocas mediciones difiere notablemente de una distribución normal.

Por ejemplo, para obtener un nivel de confianza $p = 95,45\%$ en el caso de sólo 3 mediciones, hay que expandir la incertidumbre estándar por un factor $t_{95,45}(2) = 4,53$ en lugar de $k = 2$.

INCERTIDUMBRE COMBINADA Y EXPANDIDA DE ACUERDO A LA GUM

A la incertidumbre de un mensurando generalmente contribuye una serie de fuentes de incertidumbre, que se combinan según la ley de propagación de incertidumbres [1], obteniendo de esta manera la incertidumbre combinada u_c .

En equivalencia a (3), la incertidumbre expandida U_p para un nivel de confianza p se estima por:

$$U_p = t_p(n_{ef}) \cdot u_c \quad (5)$$

n_{ef} es el número efectivo de los grados de libertad, que se calcula según la ecuación de Welch-Satterthwaite (ver p. ej. [1]):

$$\frac{1}{n_{ef}} = \sum_{i=1}^N \left(\frac{u_i}{u_c} \right)^4 \cdot \frac{1}{n_i} \quad (6)$$

donde u_i es la contribución de la fuente de incertidumbre i al mensurando y n_i son los grados de libertad de la fuente i . Para incertidumbres por repetibilidad o reproducibilidad, n_i se obtiene de acuerdo a (4). Para incertidumbres obtenidas por un método tipo B la GUM menciona una regla para asignarles un número de grados de libertad, ampliando el concepto de (2) (ver [1], G.4).

Con el número efectivo de grados de libertad n_{ef} se determina el factor $t_p(n_{ef})$.

PROBLEMÁTICA

A diferencia de este método recomendado por la GUM, en muchos laboratorios se aplican métodos que pretenden ser simplificados, evitando el cálculo de n_{ef} [3, 4, 5, 6]. El concepto común consiste en incrementar la incertidumbre obtenida por un método tipo A, por ejemplo la incertidumbre por repetibilidad, antes de la combinación:

$$u_A^+ = t_p(n-1) \cdot u_A \quad (7)$$

Después se calcula la incertidumbre combinada u_c utilizando u_A^+ en lugar de u_A y se expande u_c con $k = 2$ en lugar de $t_{95}(n_{ef})$ para aspirar un nivel de confianza de 95%.

Rigurosamente, de acuerdo a la GUM ([1], E.4.1), este procedimiento no es válido, debido a que la ley de propagación de incertidumbres sólo aplica para combinar incertidumbres estándar, mientras u_A^+ representa una incertidumbre expandida.

A continuación se presenta una comparación numérica entre el método recomendado por la GUM y varios métodos simplificados.

MÉTODOS ANALIZADOS Y PROCEDIMIENTO DE LA COMPARACIÓN NUMÉRICA

En este análisis todos los métodos tienen como meta un nivel de confianza $p = 95,45\%$.

Para la comparación se consideran dos contribuciones a la incertidumbre: una contribución u_A obtenida por un método de evaluación tipo A, por ejemplo repetibilidad, y una componente u_B , que es la combinación, en forma de una incertidumbre estándar, de todas las demás contribuciones obtenidas por métodos tipo B. No se consideran correlaciones, así que se obtiene la incertidumbre combinada por:

$$u_c = \sqrt{u_A^2 + u_B^2} \quad (8)$$

A continuación se muestran los métodos analizados. El número en el subíndice indica su identificación en las gráficas presentadas. n_A son los grados de libertad asociados con u_A .

$$M2 \quad U_2 = 2 \cdot \sqrt{\left[\frac{t_{95,45}(n_A)}{2} \cdot u_A \right]^2 + u_B^2} \quad (9)$$

$$M3 \quad U_3 = 2 \cdot \sqrt{\left[t_{68,27}(n_A) \cdot u_A \right]^2 + u_B^2} \quad (10)$$

$$M4 \quad U_4 = \sqrt{\left[t_{95}(n_A) \cdot u_A \right]^2 + 3 \cdot u_B^2} \quad (11)$$

M4 es un método adicional discutido en la GUM (ver [1], G.5, pero no es el método recomendado por ésta).

Adicionalmente se analiza el caso más simple de no hacer corrección alguna por el pequeño número de mediciones:

$$M1 \quad U_1 = 2 \cdot \sqrt{u_A^2 + u_B^2} = 2 \cdot u_c \quad (12)$$

Para poder comparar los métodos, se supone que la incertidumbre combinada siempre es 1:

$$u_c = \sqrt{u_A^2 + u_B^2} = 1 \quad (13)$$

En el análisis se varía la contribución de u_A a la incertidumbre combinada entre 0 y 1, para diferentes grados de libertad n_A de u_A :

$$0 \leq u_A \leq 1 \quad (14)$$

Debido a que $u_B = \sqrt{1 - u_A^2}$ resulta que (14) corresponde a $1 \geq u_B \geq 0$.

Nótese que $u_A = 0,71$ significa que las contribuciones de u_A y u_B a la incertidumbre combinada son iguales.

El número de grados de libertad n_B asociado con u_B se considera muy grande (infinito), lo que significa que los valores de u_B son "seguros".

El análisis se realiza en hojas de cálculo de Excel™, calculando, en función de u_A y para un n_A dado, n_{ef} y la incertidumbre expandida U_x para cada método, incluyendo el método recomendado por la GUM $U_{95\%}$. Posteriormente se realizan las comparaciones respectivas, graficando la razón $U_x / U_{95\%}$. Cabe destacar que en esta representación el valor de 1 corresponde a la

incertidumbre obtenida por el método recomendado por la GUM.

RESULTADOS

En la Figura 1 se presenta la comparación de la incertidumbre expandida, obtenida de acuerdo a (12), por no considerar los grados de libertad y expandir con $k = 2$ (M1), con el método recomendado por la GUM, dada por la razón:

$$\frac{2 \cdot u_c}{U_{95\%}} = \frac{2 \cdot u_c}{t_{95,45}(n_{ef}) \cdot u_c} = \frac{2}{t_{95,45}(n_{ef})} \quad (15)$$

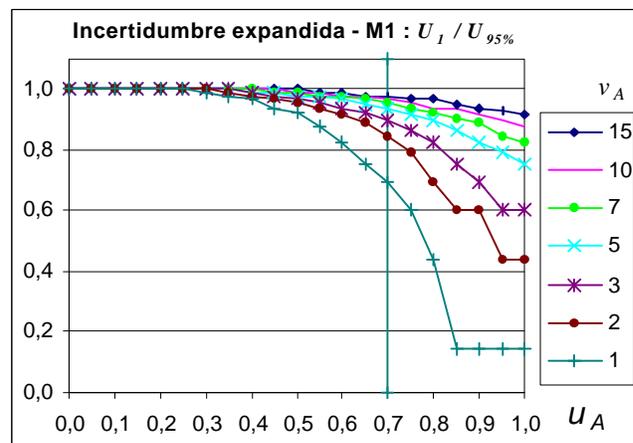


Fig. 1 Incertidumbre expandida como función de u_A y para diferentes valores n_A , obtenida con M1 y normalizada con respecto al método recomendado por la GUM.

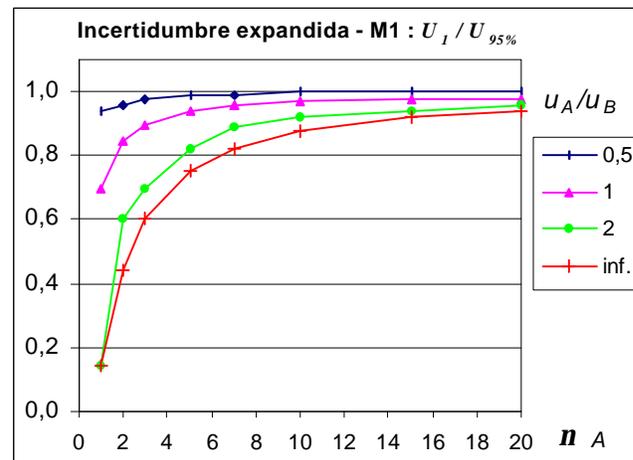


Fig. 2 Incertidumbre expandida como función de n_A y para diferentes razones u_A / u_B , obtenida con M1 y

normalizada con respecto al método recomendado por la GUM. La última curva ("inf.") significa que $u_B=0$.

Como se esperaba, se observa una subestimación considerable en el caso de una contribución dominante de u_A con pocos grados de libertad.

De forma similar, la Figura 2 muestra esta subestimación por M1 en función del número efectivo de grados de libertad, para diferentes razones u_A/u_B .

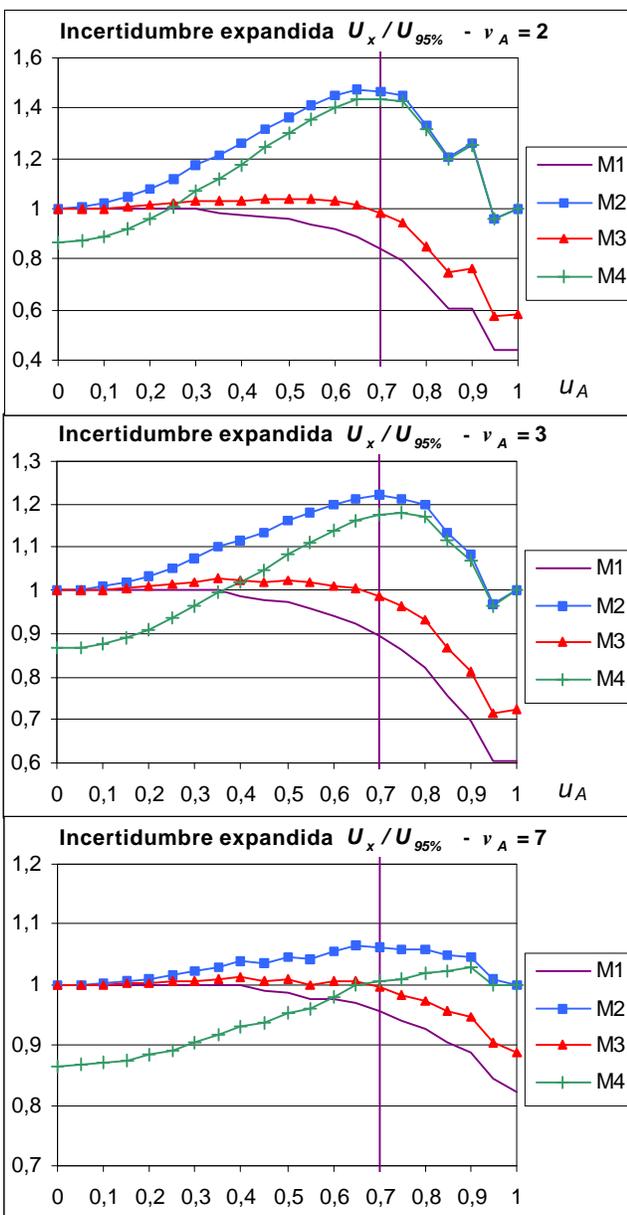


Fig. 3 Incertidumbre expandida como función de u_A para $\nu_A = 2, 3$ y 7 obtenida con los cuatro métodos y normalizada con respecto al método recomendado por la GUM. La línea vertical a $u_A=0,71$ indica la igualdad $u_A=u_B$

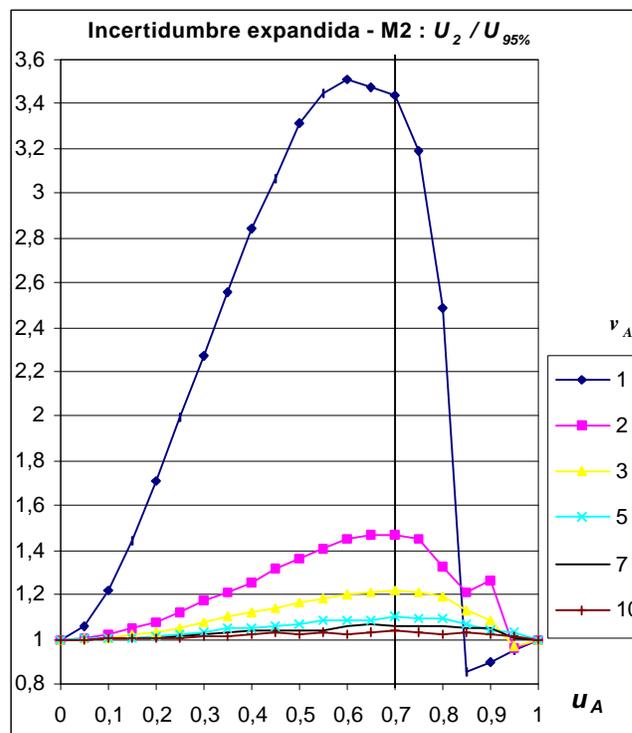
para este valor. Las discontinuidades (escalones) en las gráficas resultan de un truncamiento de los valores exactos obtenidos para ν_{ef} .

La Figura 3 muestra una comparación de diferentes métodos U_x con el método recomendado por la GUM $U_{95\%}$ para $\nu_A = 2, 3$ y 7 , presentando la razón $U_x / U_{95\%}$. Valores mayores a 1 significan una sobreestimación de la incertidumbre, mientras valores menores a 1 significan una subestimación.

Se observa que M2 siempre resulta en una sobreestimación, mientras M3 resulta en una subestimación si u_A es mayor de u_B . Esta subestimación puede ser bastante grande, para ν_A pequeño y $u_A \gg u_B$. Sin embargo, M3 da mejores resultados que M1 (o sea sin corrección por los grados de libertad).

El M4 muestra un comportamiento parecido a M2, salvo en el caso de una dominancia de u_B . En este caso, puede subestimar la incertidumbre hasta 13%, si $u_B \gg u_A$ (ver también [1], G.5.2).

Fig. 4 Incertidumbre expandida como función de u_A y para diferentes valores ν_A , obtenida con M2 y normalizada con respecto al método recomendado por la



GUM.

Para el caso extremo de $n_A = 1$ la sobreestimación por M2 y M4 llega hasta un factor 3,5 (para $u_A \approx u_B$, ver Figura 4) y la subestimación por M3 y M1 llega hasta 86%.

Para n_A grandes, todos los métodos se acercan cada vez más al método recomendado por la GUM. La Figura 4 muestra este comportamiento para M2 variando n_A

CONCLUSIONES

Los resultados presentados se pueden resumir en las siguientes conclusiones:

- 1) No considerar los grados de libertad en el cálculo de la incertidumbre expandida en M1 resulta en una subestimación notable, solamente cuando u_A es dominante y tiene pocos grados de libertad.
- 2) M3 da una subestimación de la incertidumbre expandida cuando u_A supera u_B , que puede ser grande si u_A es dominante y tiene pocos grados de libertad.
- 3) M2 siempre sobreestima la incertidumbre expandida y puede ser considerada como un "método conservador". La sobreestimación generalmente es moderada (menor de 22%), salvo en los casos de
 - a) $n_A = 1$
 - b) $n_A = 2$ y las contribuciones de u_A y u_B son parecidas
- 4) M4 da resultados parecidos a M2, con desviaciones de éste (subestimación de la incertidumbre expandida) en el caso de que domine u_B .

Se resume que ninguno de los métodos analizados representa una aproximación muy cercana al método recomendado por la GUM. Todos los métodos simplificados, salvo M2, pueden presentar subestimaciones de la incertidumbre, lo que por sí puede ser crítico. En algunos casos y dentro de ciertos límites, p. ej. u_A obtenido a partir de por lo menos $n = 3$ mediciones, pueden ser aceptados como aproximaciones, cuando los resultados de la estimación de la incertidumbre no son sumamente críticos. Esto vale, por ejemplo, cuando el daño provocado por una medición equivocada no es muy alto, o cuando los resultados de las mediciones no

se encuentran cerca de las tolerancias del producto medido (verificado).

Sin embargo, hay que cuestionar las ventajas de la simplificación de M2, M3 y M4, que, de todas maneras, requieren conocimientos del metrologo sobre los grados de libertad. Utilizando hojas de cálculo para el análisis de los resultados, como p. ej. Excel™, el cálculo del número efectivo de los grados de libertad puede ser automatizado, así que el método recomendado por la GUM realmente no representa una complicación significativa en comparación con M2, M3 y M4, pero tiene la ventaja de ser el método recomendado por un documento reconocido como la referencia internacional en la estimación de la incertidumbre de medición.

AGRADECIMIENTOS

El autor quiere agradecer a Héctor González, CENAM, por valiosas discusiones.

REFERENCIAS

- [1] a) Guide to the Expression of Uncertainty in Measurement, ISO 1995
b) Versión traducida al español: J. M. Figueroa, Guía BIPM/ISO para la expresión de la incertidumbre en las mediciones, CNM-MED-PT-002, CENAM 1997
- [2] W. Schmid, R. Lazos, Guía para estimar la incertidumbre de la medición, CNM-INC-PT-001, CENAM 2000
- [3] ISO/TR 14353-2:1999(E), Geometrical product specification (GPS) – Inspection by measurement of workpieces and measuring equipment, Part 2
- [4] F. Motolonia, Cálculo de incertidumbre en calibración de indicadores de carátula, Boletín CIDESI, 1/99, 1999, p.3
- [5] C. Hentschel, A. Gerster, Setting up a Calibration System for Optical Power Meters and Attenuators – A Recommendation, Hewlett Packard, Solution Note 153-2, 1996
- [6] Comunicación personal con varios metrologos del CENAM