EL NÚMERO EFECTIVO DE GRADOS DE LIBERTAD

Enrique Villa Diharce
Centro de Investigación en Matemáticas, A. C.
Jalisco s/n, Mineral de Valenciana, Guanajuato, Gto. CP. 36240
Tel.: (473) 732 71 55 ext. 49562, Fax: (473) 732 57 49, Email: villadi@cimat.mx

Resumen: En el análisis de datos de medición, presentamos el resultado de un proceso de medición, mostrando el resultado de la medición y su incertidumbre. En la mayoría de los casos el mensurando no es directamente medible, sino que depende de otras cantidades medibles, a través de una función que generalmente es no lineal. En este caso, la determinación de la incertidumbre expandida de la estimación del mensurando, se obtiene haciendo uso de una estadística, cuya distribución se desconoce, pero se aproxima por una distribución t de Student, en la que el número de grados de libertad se obtiene, de acuerdo con la recomendación de la GUM, utilizando la aproximación de Welch-Satterthwaite. A este número de grados de libertad se le conoce como el número efectivo de grados de libertad. En este artículo comentamos el concepto de grados de libertad y presentamos una derivación de la expresión conocida para obtener el número efectivo de grados de libertad.

1. INTRODUCCIÓN

En el análisis de datos de medición, presentamos el resultado de un proceso de medición, mostrando el resultado de la medición y su incertidumbre. Esta última puede darse en forma estándar o expandida. La conveniencia de dar la versión expandida de la incertidumbre, es que le asociamos una cobertura explicita, lo cual facilita su interpretación, ya sea en el terreno del enfoque frecuentista, o en el subjetivo (grado de creencia). La incertidumbre expandida corresponde al rango de un intervalo de confianza del mensurando. Para construir tal intervalo de confianza, utilizamos una estadística T, dada por el cociente de una estadística normal estándar entre una estadística que incluye a la incertidumbre combinada de las variables de influencia del modelo de medición. Cuando la estadística del denominador tiene asociada una distribución ji-cuadrada, entonces el cociente tiene una distribución t de Student, con un número de grados de libertad igual a los grados de libertad de la estadística ji-cuadrada del denominador.

Frecuentemente, la estadística que involucra a la incertidumbre combinada no tiene una distribución ji-cuadrada exacta, por lo tanto la estadística T no tiene una distribución t de Student exacta. Esta complicación nos impide tener un intervalo de confianza para el mensurando, con la cobertura nominal deseada exacta.

Un procedimiento recomendado por la GUM [1], para estimar la incertidumbre expandida, consiste en aproximar la distribución de la estadística T con una distribución t de Student con un número de grados de libertad dado por la aproximación de Welch – Satterthwaite. A este número de grados de libertad se le denomina el número efectivo de grados de libertad.

En la sección 2, se muestra el origen de los grados de libertad asociados a una incertidumbre expandida. En la sección 3, presentamos el desarrollo de la formula de Welch - Satterthwaite, para calcular el número efectivo de grados de libertad.
En la sección 4 se comentan las restricciones que tiene la aproximación de Welch-Sattertwaite y se comenta un procedimiento alternativo para la determinación de la incertidumbre expandida. En la sección 5, se dan las conclusiones del trabajo.

2. LOS GRADOS DE LIBERTAD

En la mayoría de los casos el mensurando no es directamente medible, sino que depende de otras cantidades medibles $X_1, X_2, ..., X_k$, a través de una función $f(X_1, X_2, ..., X_k)$. Así la incertidumbre de la medición Y del mensurando, es el resultado de combinar las incertidumbres de medición de las diferentes cantidades medidas $X_1, X_2, ..., X_k$, con diferentes pesos, que dependerán de la importancia que cada componente tiene en el modelo de medición $Y = f(X_1, X_2, ..., X_k)$.

Una magnitud X que varia aleatoriamente tiene asociados dos parámetros muy importantes: su esperanza $\mu_X = E(X)$ que nos indica el valor alrededor de cual se tienen las observaciones de X, y su varianza σ_X^2 que expresa la magnitud de la variabilidad de las observaciones alrededor de su valor esperado μ_X.

Cuando se tienen n observaciones independientes $x_1, x_2, ..., x_n$ de X, estimamos su esperanza como el promedio $\overline{x}$ de las n observaciones,

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \ ,$$

mientras que la varianza se estima por

$$s_X^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2 \ .$$

Cuando tenemos un mensurando, del que tomamos mediciones con un instrumento, ocurre que observamos diversos valores que varian aleatoriamente, de acuerdo a alguna distribución. Generalmente asociamos el valor del mensurando con el valor esperado de la distribución, de aquí que estimar el valor del mensurando equivale a estimar el valor esperado μ_X, La incertidumbre de la estimación $\overline{x}$ proviene de la variabilidad de $\overline{x}$ alrededor de μ_X, que cuantificamos por su varianza $\sigma_{\overline{x}}^2 = \sigma_X^2 / n$, y que estimamos como $s_{\overline{x}}^2 = s_X^2 / n$. Tambien podemos tomar como incertidumbre de medición la desviación estándar de la media, $\sigma_{\overline{x}} = \sqrt{\sigma_{\overline{x}}^2}$.

Cuando la información que tenemos del mensurando es un conjunto de n observaciones independientes $x_1, x_2, ..., x_n$, entonces tenemos que para un valor dado de la estimación $\overline{x} = (x_1 + x_2 + ... + x_n)/n$, tenemos solo $n-1$ observaciones x_i que pueden tomar valores libremente, ya que el término restante tiene un valor determinado por estos $n-1$ términos elegidos y el promedio $\overline{x}$. A este número de términos independientes se le conoce como el número de grados de libertad. De aquí que podemos interpretar el número de grados de libertad como la cantidad de información útil en la determinación de la estimación. En cambio si las observaciones no son independientes, tenemos información redundante en las observaciones, por lo que la información útil o efectiva será menor

Otra interpretación del número de grados de libertad, es el parámetro de la distribución ji-cuadrada asociada al estimador de la varianza, calculada a partir de n observaciones independientes. En este caso sabemos que la estadística

$$W = \frac{(n-1)S_{\overline{X}}^2}{\sigma^2} \qquad (1)$$

tiene una distribución ji-cuadrada con $n-1$ grados de libertad. Con este enfoque entonces, el número efectivo de grados de libertad viene a ser el parámetro asociado a la distribución ji-cuadrada asociada al estimador de la incertidumbre de la medición considerada.

La estadística anterior interviene en la estadística T utilizada para determinar la incertidumbre expandida en forma de intervalo de confianza,

$$T = \frac{Z}{\sqrt{\dfrac{W}{(n-1)}}} = \frac{\dfrac{\overline{X} - \mu_X}{\sigma_{\overline{X}}}}{\sqrt{\dfrac{S_{\overline{X}}^2}{\sigma_{\overline{X}}^2}}} = \frac{\overline{X} - \mu_X}{S_{\overline{X}}} \ . \qquad (2)$$

Así, además de expresar en forma numérica la incertidumbre, ya sea como $s_{\overline{x}}^2$, o $s_{\overline{x}}$, podemos expresarla en forma expandida, como un intervalo de confianza del mensurando,

$$(\overline{x} - t_{(n-1,1-\alpha/2)} s_{\overline{x}}, \overline{x} + t_{(n-1,1-\alpha/2)} s_{\overline{x}}) \ , \qquad (3)$$

donde $\overline{x}$ es la estimación del mensurando, $s_{\overline{x}}$ es la incertidumbre de la medición, y $t_{(n-1,1-\alpha/2)}$ el coeficiente de confianza para la distribución t de

Student con $\nu = n-1$ grados de libertad. Este intervalo tiene una cobertura nominal de $(1-\alpha)100\%$.

Cuando el mensurando de interés Y es resultado de varias magnitudes de influencia $X_1, X_2, ..., X_k$, de acuerdo a un modelo de medición dado por la función $Y = f(X_1, X_2, ..., X_k)$, entonces hay que convertir la información que tenemos sobre los resultados de medición e incertidumbres de $X_1, X_2, ..., X_k$, en el resultado de medición y la incertidumbre de Y. Hacer esto en forma exacta puede ser un problema complicado, cuando la función f es nolineal, y mas difícil todavía, si además, las variables de influencia son correlacionadas.

La manera de atacar este problema consiste en tomar una aproximación de Taylor, para Y, alrededor de su valor esperado μ_Y. De esta forma, cuando las variables de influencia son no correlacionadas, obtenemos las aproximaciones para μ_Y y σ_Y^2:

$$\mu_Y = f(\mu_{X_1}, \mu_{X_2}, ..., \mu_{X_k}),$$

$$\sigma_Y^2 = \sum_{i=1}^{k} \left(\frac{\partial f}{\partial x_i}\right)^2 \sigma_{X_i}^2 .$$

Estos valores son estimados por:

$$y = f(x_1, x_2, ..., x_k),$$

$$s_Y^2 = \sum_{i=1}^{k} \left(\frac{\partial f}{\partial x_i}\right)^2_{(x_1, x_2, ..., x_k)} s_{X_i}^2 .$$

La incertidumbre numérica está dada por s_Y^2, mientras que la incertidumbre expandida está dada por el intervalo de confianza

$$(y - t_{(\nu_{ef.}, 1-\alpha/2)} s_y, y + t_{(\nu_{ef.}, 1-\alpha/2)} s_y), \qquad (4)$$

donde y es el resultado de medición, s_y su incertidumbre combinada, $t_{(\nu_{ef.}, 1-\alpha/2)}$ es el coeficiente de confianza y $\nu_{ef.}$ el número de grados de libertad efectivos, dado por

$$\nu_{ef.} = \frac{s_y^4}{\sum_{i=1}^{k} \frac{s_i^4}{\nu_i}}, \qquad (5)$$

siendo $s_1, s_2, ..., s_k$ las incertidumbres estándar de las variables de influencia y $\nu_1, \nu_2, ..., \nu_k$ sus grados de libertad correspondientes.

Para entender el origen del número efectivo de grados de libertad, hay que considerar la estadística T, utilizada para obtener el intervalo de confianza (4)

$$T = \frac{y-\mu_Y}{S_y} = \frac{\frac{y-\mu_Y}{\sigma_Y}}{\sqrt{\frac{S_Y^2}{\sigma_Y^2}}} = \frac{Z}{\sqrt{\frac{W}{\nu_{ef}}}} . \qquad (6)$$

Aquí, la estadística Z tiene una distribución Normal estándar y la estadística $W = \nu_{ef} S_y^2 / \sigma_y^2$ tiene aproximadamente una distribución ji-cuadrada con ν_{ef} grados de libertad.

En este caso, el número efectivo de grados de libertad ν_{ef} viene a ser el parámetro de la distribución ji-cuadrada asociada a la incertidumbre expandida s_Y^2, que se obtiene como una combinación lineal de las incertidumbres $s_1^2, s_2^2, ..., s_k^2$, de las mediciones de las variables de influencia.

La expresión del número efectivo de grados de libertad, se conoce como la aproximación de Welch-Satterthwaite, que describimos en la siguiente sección.

3. APROXIMACIÓN DE WELCH-SATTERTHWAITE

Sean $W_1, W_2, ..., W_k$, k variables aleatorias independientes e idénticamente distribuidas como ji-cuadrada con $r_1, r_2, ..., r_k$ grados de libertad respectivamente. Es un hecho conocido [2] que la suma $W_1 + W_2 + ... + W_k$ se distribuye también como ji-cuadrada con $r = r_1 + r_2 + ... + r_k$ grados de libertad. Esta propiedad de herencia de distribución no se tiene cuando, en lugar de una suma, consideramos una combinación lineal $a_1W_1 + a_2W_2 + ... + a_kW_k$ de las variables aleatorias.

La complicación en la determinación del número efectivo de grados de libertad, se debe a que la incertidumbre combinada s_y^2 es una combinación lineal de las incertidumbres de las variables de influencia. Si en lugar de una combinación lineal, fuese una suma de términos no correlacionados, entonces el número de grados de libertad asociado a la incertidumbre combinada seria la suma de los grados de libertad de las variables de influencia.

Cuando la incertidumbre combinada es una combinación lineal de incertidumbres, entonces la información de cada una de ellas dentro del total, tiene un efecto modulado por la constante que multiplica al término correspondiente. Esta combinación no aditiva de la información proporcionada por cada componente es lo que tenemos en la aproximación de Welch-Satterthwaite.

La lógica en el desarrollo de la aproximación, consiste en suponer que la combinación lineal $a_1W_1 + a_2W_2 + ... + a_kW$, tiene aproximadamente una distribución ji-cuadrada con un número de grados de libertad que debemos determinar. Para esto, procedemos ajustando una distribución ji-cuadrada a la combinación lineal, y estimamos su parámetro (grados de libertad) por el método de momentos.

Suponemos entonces, que la combinación lineal es equivalente al cociente dado por una variable aleatoria U distribuida ji-cuadrada, dividida entre sus grados de libertad ν, esto es, que $\sum_{i=1}^{k} a_iW_i$ y U/ν tienen la misma distribución. Esto se expresa como la relación de equivalencia,

$$\sum_{i=1}^{k} a_iW_i \sim \frac{U}{\nu} \quad .$$

Según el metodo de momentos, para estimar el parámetro ν, igualamos los dos primeros momentos de las variables que tenemos en ambos lados de la relación de equivalencia anterior. De esta forma obtenemos las ecuaciones

$$E(\sum_{i=1}^{k} a_iW_i) = E(\frac{U}{\nu}) \, , \tag{7}$$

$$E(\sum_{i=1}^{k} a_iW_i)^2 = E(\frac{U}{\nu})^2 \, . \tag{8}$$

Como la esperanza de una variable aleatoria distribuida como ji-cuadrada es igual a sus grados de libertad, y su varianza es dos veces sus grados de libertad, entonces

$$E(U) = \nu$$
$$V(U) = E(U)^2 - \left[E(U)\right]^2 = 2\nu \, .$$

De lo anterior obtenemos que

$$E(\frac{U}{\nu}) = 1 \, ,$$

$$E(\frac{U}{\nu})^2 = \frac{2}{\nu} - 1.$$

Sustituyendo estas esperanzas en (7) y (8), obtenemos,

$$E(\sum_{i=1}^{k} a_iW_i) = 1 \text{ , y} \tag{9}$$

$$E(\sum_{i=1}^{k} a_iW_i)^2 = \frac{2}{\nu} - 1. \tag{10}$$

De la ecuación (9) no podemos deducir el valor de ν, solo obtenemos una restricción para los coeficientes de la combinación lineal. Despejamos ν de la ecuación (10) y obtenemos

$$\nu = \frac{2}{E(\sum_{i=1}^{k} a_iW_i)^2 - 1}. \tag{11}$$

De esta ecuación obtenemos un estimador para el número de grados de libertad, sustituyendo en el denominador la esperanza del cuadrado de la combinación lineal, por la combinación lineal cuadrática únicamente. Así tenemos,

$$\hat{\nu} = \frac{2}{(\sum_{i=1}^{k} a_i W_i)^2 - 1}.$$

Este estimador tiene el inconveniente de que nos puede dar valores negativos, lo cual es inadmisible. Esta expresión para $\hat{\nu}$ no considera la restricción de los coeficientes de la combinación lineal. Enseguida consideramos una versión modificada de este estimador, incluyendo la restricción.

$$E(\sum_{i=1}^{k} a_i W_i)^2 = V(\sum_{i=1}^{k} a_i W_i) + \left[E(\sum_{i=1}^{k} a_i W_i)\right]^2$$

$$= \left[E(\sum_{i=1}^{k} a_i W_i)\right]^2 \left\{ \frac{V(\sum_{i=1}^{k} a_i W_i)}{\left[E(\sum_{i=1}^{k} a_i W_i)\right]^2} + 1 \right\},$$

y considerando de (9) que $E(\sum_{i=1}^{k} a_i W_i) = 1$, entonces

$$E(\sum_{i=1}^{k} a_i W_i)^2 = \frac{V(\sum_{i=1}^{k} a_i W_i)}{\left[E(\sum_{i=1}^{k} a_i W_i)\right]^2} + 1.$$

Sustituyendo esto en la expresión (11) para ν, resulta

$$\nu = \frac{2\left[E(\sum_{i=1}^{k} a_i W_i)\right]^2}{V(\sum_{i=1}^{k} a_i W_i)}. \qquad (12)$$

Finalmente, considerando que $W_1, W_2, ..., W_k$ son variables aleatorias independientes distribuidas como ji-cuadradas, tenemos que

$$V(\sum_{i=1}^{k} a_i W_i) = \sum_{i=1}^{k} a_i^2 V(W_i) = 2\sum_{i=1}^{k} \frac{a_i^2 \left[E(W_i)\right]^2}{\nu_i} \qquad (13)$$

La última igualdad se debe a que

$$V(W_i) = 2\left[E(W_i)\right]^2 / \nu_i.$$

Sustituyendo ahora, la expresión (13) de la varianza de la combinación lineal, en la expresión (12) de ν tenemos

$$\nu = \frac{\left[E(\sum_{i=1}^{k} a_i W_i)\right]^2}{\sum_{i=1}^{k} \frac{a_i^2}{\nu_i} E(W_i)}.$$

De esta expresión para los grados de libertad como función de esperanzas de las variables aleatorias, obtenemos su estimador $\hat{\nu}$, sustituyendo las esperanzas por sus observaciones, esto es,

$$\hat{\nu} = \frac{(\sum_{i=1}^{k} a_i W_i)^2}{\sum_{i=1}^{k} \frac{a_i^2}{\nu_i} W_i^2}. \qquad (14)$$

Esta aproximación obtenida en la década de los 40, sigue aún hoy siendo de gran utilidad, su uso se sustituye algunas veces por procedimientos de simulación para la determinación de incertidumbres expandidas. Cabe hacer notar que el procedimiento de simulación de observaciones de las variables de influencia, cuando éstas están correlacionadas no es sencillo.

La expresión (14) obtenida para estimar los grados de libertad coincide con la expresión conocida para el número efectivo de grados de libertad, después de hacer las siguientes sustituciones: $a_i = (\partial f / \partial x_i)$, y $W_i = s_i^2$.

Los trabajos originales sobre esta aproximación fueron desarrollados en forma independiente por Welch [3,4,7] y Satterthwaite [5,6]. Este problema es similar al problema de inferencia para una diferencia de medias, cuando las varianzas son desconocidas y diferentes. Este último problema conocido en la literatura estadística como el problema de Behrens-Fisher [8], ha generado una polémica interesante, y ha atraído la atención de un gran número de estadísticos, y por lo tanto se han propuesto varias soluciones desde diferentes enfoques.

4. RESTRICCIONES O LIMITACIONES

El desarrollo de la aproximación de Welch-Satterthwaite supone que las variables de influencia tienen distribución normal y además son nocorrelacionadas. Estos supuestos con frecuencia no se cumplen, ya que además de la distribución normal, las distribuciones uniforme y triangular intervienen como modelos de medición de las variables de influencia. Y además, la correlación entre dichas variables en algunos casos es significativa.

Frecuentemente, el supuesto de no-correlación no se cumple en la practica metrológica, ya que las mediciones que se hacen de las diferentes variables de influencia que intervienen en un proceso de medición, están sujetas a un ambiente común, lo cual genera una cierta estructura de dependencia entre dichas variables.

M Ballico [9] compara la formula de Welch-Satterthwaite con otros procedimientos basados en series de potencias utilizados para determinar la incertidumbre expandida, y encuentra que la formula sobreestima el número efectivo de grados de libertad cuando una de las principales componentes tiene pocos grados de libertad. Esto genera una subestimación del correspondiente intervalo de confianza, lo cual es aun mas notorio para altos niveles de confianza.

Una alternativa que se ha elegido algunas veces para la determinación de la incertidumbre expandida o intervalo de confianza, ha sido la técnica de simulación [10], también conocida como Monte Carlo. Cuando las variables de influencia que se simulan son nocorrelacionadas, este procedimiento es sencillo, aun cuando se tengan distribuciones diferentes a la normal. En cambio, si las variables son correlacionadas, el problema de simulación en forma exacta es muy complicado, ya que se requiere conocer la distribución conjunta de las variables, lo cual no es fácil. Existe software como por ejemplo, Cristal Ball [11] y Risk [12], que simulan observaciones de variables aleatorias correlacionadas, con cualquier conjunto de distribuciones marginales, y con una relación de correlación definida por el coeficiente de correlación por rangos, siguiendo el procedimiento de Iman y Conover [13].

5. CONCLUSIONES

En este artículo hemos comentado el concepto de grados de libertad asociado a la determinación de la incertidumbre de medición.

Se ha mostrado el desarrollo de la aproximación de Welch-Satterthwite para obtener el número efectivo de grados de libertad.

La estadística de la incertidumbre combinada no tiene una distribución ji-cuadrada exacta, debido a que en el numerador no tiene una suma de cuadrados sino una combinación lineal de cuadrados.

Se han discutido limitaciones de esta aproximación, como son los supuestos de distribución normal de las variables de influencia, así como la no correlación entre ellas.

Además se ha comentado que una alternativa que se puede seguir al evaluar la incertidumbre de un proceso de medición, cuando las variables de influencia son no-correlacionadas, es la técnica de Monte Carlo. Cuando se usa algún software para simular observaciones de variables correlacionadas, se recomienda conocer el procedimiento que dicho software sigue para introducir la correlación entre las variables, para estar seguros de que la correlación que el software considera es la que nosotros hemos determinado.

REFERENCIAS

[1] Guide to the Expression of Uncertainty in Measurement, Geneva, International Organization for standardization, 1993, ISBN 92-67-10188-9.

[2] Casella, G. and Berger, R. L., Statistical Inference, Wadsworth & Brooks/Cole, Pacific Grove, CA. 1990.

[3] Welch, B. L., The Specification of Rules for Rejection too Variable a Product, with Particular Reference to an Electric Lamp Problem, Supplement to the Journal of the Royal Statistical Society, 3(2), 1936, 29-48.

[4] Welch, B. L., The Significance of the Difference Between two Means when the Population Variances are Unequal, 29, 1938, 350-362. Satterthwaite,F. E., Psychometrika, 1941, 6(5), 309-316.

[5] Satterthwaite,F. E., Psychometrika, 1941, 6(5), 309-316.

[6] Satterthwaite,F. E., An approximate distributión of estimates of variance components, Biometrics, Bull., 2(6), 1946, 110-114.

[7] Welch, B. L., The generalization of ´Student´s´ problem when several different population variances are involved, Biometrika 34, (1947), 28-35.

[8] Kendall, M. and Stuart, A., The Advanced Theory of Statistics, , Vol. II: Inference and

Relationship, 4th. Ed., Macmillan, New York, NY., 1979.

[9] Ballico, M., Limitations of the Welch-Satterthwaite aproximation for measurement uncertainty calculations, Metrologia, 37, 2000, 61-64.

[10] Cox, M. G., Dainton, A. B., Forbes, A. B., Harris, P. M., Schwenke, P. M., Siebert, B. R. L., and Woger, W., Use of Monte Carlo Simulation for uncertainty evaluation in metrology, in Ciarlini, P., Cox, M. G., Filipe, E, Pavese, F, and Richter, D., editors, Advanced Mathematical Tools in Metrology V. Series on Advances in Mathematics for Applied Sciences, Vol. 57, World Scientific, 2001, 93-104.

[11] Decisionengineering, Inc., Crystal Ball Version 4.0 User Manual.

[12] Winston, W. L., Simulation usiong @RISK, Duxbury, Pacific Grove CA. 2001.

[13] Iman R. L. and Conover W. J., A Distribution-Free Aproach to Inducing Rank Correlation Among Input Variables, Communications in Statistics - Simulation and Computation, 11(3), 1982, 331-334.