Some remarks on a bivariate analysis in the propagation of measurement uncertainty as an alternative approach to the Monte Carlo method

M. Catelani, L. Ciani, A. Zanobini

Department of Electronics and Telecommunications, University of Florence,
Via S. Marta 3, 50139, Florence (Italy)
marcantonio.catelani@unifi.it , andrea.zanobini@unifi.it, lorenzo.ciani@unifi.it

Abstract **–** Recent literature shown that in cases where there is strong interdependence between direct measures involved in an indirect measurement, or when the uncertainties of direct measures are not as small as one would wish, or again, when a strong nonlinearity ties the measures involved to each other, the Monte Carlo method, proposed by Supplement 1 to the "Guide to the expression of uncertainty in measurement" (GUM), manages in most cases to obtain a more accurate measurement uncertainty evaluation than the one actually proposed by the GUM and bound to Taylor series developed.
This paper aims to offer a new approach related to bivariate analysis of standard models such as sum and product of measures when it is either possible, or unsuitable, to apply the two methods quoted in Supplement 1.

I. Theoretical Introduction

Considering *X* and *Y* the measures of two scalar measurand and time-invariant. The respective uncertainty intervals are defined as:

$$\mu_X \pm u_X = \mu_X(1 \pm u_X'); \;\; \mu_Y \pm u_Y = \mu_Y(1 \pm u_Y') \tag{1}$$

where

$$\mu_x = E\{X\}, \quad u_x = k_x\sqrt{Var\{X\}} = k_x\sigma_x; \quad \mu_y = E\{Y\}, \quad u_y = k_y\sqrt{Var\{Y\}} = k_y\sigma_y \tag{2}$$

We have also:

$$\begin{cases} E\{X^2\} = \sigma_X^2 + \mu_X^2 = \dfrac{u_X^2}{k_X^2} + \mu_X^2 = \mu_X^2\left[1 + \left(\dfrac{u_X'}{k_X}\right)^2\right] \\ E\{Y^2\} = \sigma_Y^2 + \mu_Y^2 = \dfrac{u_Y^2}{k_Y^2} + \mu_Y^2 = \mu_Y^2\left[1 + \left(\dfrac{u_Y'}{k_Y}\right)^2\right] \end{cases} \tag{3}$$

Remembering the correlation coefficient between *X* and *Y*:

$$\rho_{xy} = \rho = \frac{Cov\{X,Y\}}{\sigma_x\sigma_y} \tag{4}$$

with

$$Cov\{X,Y\} = E\{(X-\mu_X)(Y-\mu_Y)\} = E\{XY\} - \mu_X\mu_Y \tag{5}$$

From Eq. (4) we obtain:

$$Cov\{X,Y\} = \rho\,\sigma_X\sigma_Y \tag{6}$$

then

$$E\{X,Y\} = Cov\{X,Y\} + \mu_X\mu_Y = \rho\,\sigma_X\sigma_Y + \mu_X\mu_Y = \mu_X\mu_Y\left[1 + \rho\frac{u_X'}{k_X}\frac{u_Y'}{k_Y}\right] \tag{7}$$

For the covariance is easily checked the following property:

$$Cov\{a + bX, c + dY\} = bd\ Cov\{XY\} \tag{8}$$

Moreover, if $Y = c + dX$ we have:

$$\sigma_Y^2 = d^2\sigma_X^2 \quad and \quad \sigma_Y = |d|\sigma_X;\ Cov\{X, c + dX\} = d\ \ Cov\{X, X\} = d\sigma_X^2 \tag{9}$$

Therefore from Eq. (4) it can be deduced that the correlation coefficient is:

$$\rho = \frac{d}{|d|} \tag{10}$$

II. Bivariate analysis of indirect measurements

In the case of indirect measurements, X and Y known the corresponding uncertainty intervals and the correlation coefficient according to Eq. (4), it is a matter of evaluating the uncertainty interval of $Z = g(X, Y)$. In general terms, one can suppose the following uncertainty interval for Z:

$$E\{g(X, Y)\} \pm k\sqrt{Var\{g(X, Y)\}} \tag{11}$$

whereas k is an imposed coverage factor (for examples the maximum between k_x and k_y). As practical examples, let's consider the sum and the product of measures as follows.

II A. Sum of measures

A generalization of the basic case $Z = X + Y$ consists of the bilinear transformation model:

$$Z = a + bX + dY \tag{12}$$

In this case, remembering that the variance is a squared operator so, the linear transformation variance of a random variable is the product of the square of the scale factor and the variance of such variable. Recalling Eq. (8) we deduce:

$$\mu_Z = a + b\mu_X + d\mu_Y \tag{13}$$

$$\sigma_Z^2 = b^2 Var\{X\} + d^2 Var\{Y\} + 2\,bd\,Cov\,\{X, Y\} \tag{14}$$

In the case of independent measures, we have an important property for the probability density function of the sum which turns out to be the convolution of the two marginal probability densities functions $f_1(m_1)$ and $f_2(m_2)$, it can be demonstrated that the density $f(m)$ of the sum *(M_1+M_2)* is obtained by:

$$f(m) = \int_{-\infty}^{+\infty} f_1(x)\ f_2(m - x)\,dx = \int_{-\infty}^{+\infty} f_1(m - x)\ f_2(x)\,dx \tag{15}$$

Another important condition occurs when the two measures X and Y are jointly normal distributed. In this instance, X and Y are also separately normal so it can be stated that $X \cong N(\mu_X, \sigma_X^2)$, $Y \cong N(\mu_Y, \sigma_Y^2)$. Besides that, the theory demonstrates that Z expresses by the Eq. (12) also appears normal so we can write $Z \cong N(\mu_Z, \sigma_Z^2)$, where the mean and the variance are obtained respectively from Eq. (13) and (14).

II A.1 First example

Let's suppose that two measures M_1 and M_2 associated with two different uncertainty components, can be characterized with uniform distribution $[-|x_1|, +|x_1|]$ and $[-|x_2|, +|x_2|]$ as $|x_1|$ and $|x_2|$ are the maximum variability values. If the two variables are independent, or in the case of they have a negligible correlation index, it can be demonstrated by Eq. (14) that the standard uncertainty of the sum $M = M_1 + M_2$ can be written as:

$$u_M = \sqrt{\frac{|x_1|^2 + |x_2|^2}{3}} \tag{16}$$

If, instead, the two measures are strongly correlated, the worst case and the following variability interval must be taken into consideration:

$$-\left[|x_1| + |x_2|\right] \le M \le \left[|x_1| + |x_2|\right] \tag{17}$$

the corresponding standard uncertainty in the case of an uniformly distributed variable is associated with measures is:

$$u_M = \frac{|x_1| + |x_2|}{\sqrt{3}} \tag{18}$$

which results greater than Eq. (16).

II A.2 Second example

For a better explanation of the proposed approach, let's now consider M_1 and M_2 as independent random variables, each distributed respectively with uniform density $f_1(m_1)$ in the interval *[-a, +a]*, with a>0, and with normal density $f_2(m_2)$ around the origin. So that it can be supposed that:

$$f_1(m_1) = \begin{cases} \dfrac{1}{2a} = \dfrac{1}{2\sqrt{3}u_1} & with \quad -a \le m_1 \le a \\ 0 & otherwise \end{cases} \tag{19}$$

with $E(M_1) = 0$, $Var(M_1) = u_1^2 = \dfrac{a^2}{3}$. For the M_2 random variable:

$$f_2(m_2) = \frac{1}{\sqrt{2\pi}u_2} e^{-\left(\frac{m_2}{2u_2}\right)^2} \tag{20}$$

with $E(M_2) = 0$, $Var(M_2) = u_2^2$.
On the basis of Eq. (15) the density function $f_M(m)$ of the random variable $M = M_1 + M_2$ is:

$$f_M(m) = \frac{1}{\sqrt{2\pi}u_2}\frac{1}{2a}\int_{-a}^{a} e^{-\frac{(m-x)^2}{2u_2^2}} dx = \frac{1}{\sqrt{2\pi}}\frac{1}{2a}\int_{\frac{m-a}{u_2}}^{\frac{m+a}{u_2}} e^{-\frac{x^2}{2}} dx = \frac{1}{2\sqrt{2\pi}}\frac{1}{a}\left[\Phi\left(\frac{m+a}{u_2}\right) - \Phi\left(\frac{m-a}{u_2}\right)\right] \tag{21}$$

We will also have $E(M) = 0$, $u_M = \sqrt{Var(M)} = \sqrt{\dfrac{a^2}{3} + u_2}$.

II B. Product of measures

The uncertainty interval of the product of measures, $Z = XY$, is defined as:

$$E\{XY\} \pm k\sqrt{Var\{XY\}} \tag{22}$$

with expected value:

$$E\{XY\} = \mu_X \mu_Y + \rho\sigma_X\sigma_Y \tag{23}$$

and variance:

$$Var\{XY\} = E\{X^2Y^2\} - E^2\{XY\} \tag{24}$$

where the first term of Eq. (24) is:

$$E\{X^2Y^2\} = \mu_X^2 \mu_Y^2 E\left\{\left[1 + \frac{X - \mu_X}{\mu_X}\right]^2 \left[1 + \frac{Y - \mu_Y}{\mu_Y}\right]^2\right\} \tag{25}$$

and the second term of Eq. (24) can be evaluated according to Eq. (23):

$$E\{X^2Y^2\} = \mu_X^2 \mu_Y^2 E\left\{1 + \left[\frac{X - \mu_X}{\mu_X}\right]^2 + \left[\frac{Y - \mu_Y}{\mu_Y}\right]^2 + 2\left[\frac{X - \mu_X}{\mu_X} + \frac{Y - \mu_Y}{\mu_Y}\right] + 4\left[\frac{X - \mu_X}{\mu_X}\right]\left[\frac{Y - \mu_Y}{\mu_Y}\right]\right\} \tag{26}$$

Remembering that the expected value is a linear operator, taking into account Eq. (5) and considering that $E\{X - \mu_X\} = E\{Y - \mu_Y\} = 0$, Eq. (26) can be simplified as:

$$E\{X^2Y^2\} = \mu_X^2 \mu_Y^2 + \mu_Y^2 \sigma_X^2 + \mu_X^2 \sigma_Y^2 + 4\mu_X \mu_Y Cov\{X, Y\} \tag{27}$$

For the variance evaluation, considering negligible the term $\rho^2 \sigma_X^2 \sigma_Y^2$ for the assumed hypothesis, we deduce:

$$Var\{XY\} \approx \mu_Y^2 \sigma_X^2 + \mu_X^2 \sigma_Y^2 + 2\rho \mu_X \mu_Y \sigma_X \sigma_Y \tag{28}$$

which, in the particular case of $\rho = \pm 1$, is reduced to:

$$Var\{XY\} \approx [\mu_Y \sigma_X \pm \mu_X \sigma_Y]^2 \tag{29}$$

Considering Eq. (3), it is interesting to note the case of two independent measures where $E\{XY\} = \mu_X \mu_Y$, therefore:

$$E\{X^2Y^2\} = E\{X^2\}E\{Y^2\} = (\sigma_X^2 + \mu_X^2)(\sigma_Y^2 + \mu_Y^2) \tag{30}$$

By applying Eq. (24) yet again:

$$Var\{XY\} = \mu_Y^2 \sigma_X^2 + \mu_X^2 \sigma_Y^2 + \sigma_X^2 \sigma_Y^2 \tag{31}$$

and assuming negligible $\sigma_X^2 \sigma_Y^2$ we have:

$$Var\{XY\} \approx \mu_Y^2 \sigma_X^2 + \mu_X^2 \sigma_Y^2 \tag{32}$$

It is interesting to observe that Eq. (32) coincides with the Eq. (28) for $\rho = 0$. In this case the final expression of the uncertainty is obtained as:

$$u_M = \sqrt{Var\{XY\}} \approx \sqrt{\mu_Y^2 \sigma_X^2 + \mu_X^2 \sigma_Y^2} \tag{33}$$

II B.1 Third example

In the case of independent measures M_1 and M_2, with a density of $f_1(m_1)$ and $f_2(m_2)$ respectively, it can be proved that the density $f(m)$ of the product ($M_1 \cdot M_2$) is obtained by:

$$f(m) = \int_{-\infty}^{+\infty} \frac{1}{|x|} f_1(x) \; f_2\left(\frac{m}{x}\right) dx \tag{34}$$

As an example, let's suppose that M_1 and M_2 are both uniformly distributed in the interval [1, 1.01] with a range 10^{-2} and constant density equal to $f_1(m_1) = f_2(m_2) = 10^2$. Obviously the product is contained in the interval [1, 1.01^2] with density:

$$f(m) = \begin{cases} 10^4 \ln(m) & 1 \le m \le 1.01 \\ 10^4 \left[2\ln(1.01) - \ln(m)\right] & 1.01 \le m \le 1.01^2 \end{cases} \quad (35)$$

It is a matter of asymmetric distribution unlike the distribution of M_1 and M_2 which are symmetric around the respective expected values equal to $E\{M_1\} = E\{M_2\} = 1.005$.
In addition the normalization property of the density of the product must be satisfied, thus it can be verified that:

$$\int_1^{1.01^2} f(m)\, dm = 1 \quad (36)$$

The expected value of the product M is equal to $E\{M\} = 1.005^2$. Considering that the variance of both measures M_1 and M_2 is represented by $Var\{M_1\} = Var\{M_2\} = \frac{\{Range\}^2}{12} = \frac{10^{-4}}{12}$, the variance of product M can be so obtained and, therefore by applying the approximate formula in Eq. (32), the corresponding standard uncertainty, therefore as:

$$u_M{}' = \sqrt{2\left[\frac{10^{-4}}{12}\frac{1}{1.005^2}\right]} \approx 10^{-2}(0.41) \quad (37)$$

Then, the final results concerning the product M can be expressed as:

$$M = M_1 M_2 = 1.005^2\left(1 \pm 10^{-2}(0.41)\right) \quad (38)$$

III. Conclusions

In this paper is explained an alternative approach to the Monte Carlo method applying a bivariate analysis in the model of sum and product of two related and uncorrelated random variables.
Three practical examples are also developed to find the standard uncertainties in the case of sum and product of two measures with uniform and normal distributions.

References

[1] BIPM, IEC, IFCC, ISO, IUPAC, IUPAP and OIML, *Guide to the Expression of Uncertainty in Measurement*, Geneva, Switzerland, 1995, International Organization for Standardization ISBN 92-67-10188-9.
[2] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, and OIML, Evaluation of measurement data — *Supplement 1 to the "Guide to the expression of uncertainty in measurement" —Propagation of distributions using a Monte Carlo method*, Joint Committee for Guides in Metrology, Bureau International des Poids et Mesures, JCGM 101: 2008.
[3] W. Navidi, *Statistics for Engineers and Scientists*, McGraw Hill, NY, 2007.
[4] Papoulis, A., *Probability and Statistics*, Englewood Cliffs, NJ: Prentice-Hall, 1990.