# New *Mathematica* functions for Uncertainty Analysis

**M. D. Mikhailov[1], V. Y. Aibe[2] and I. Yasuda[3]**
**National Institute of Metrology, Quality and Technology**
**Av. N. Sra. das Graças, Xerém, Duque de Caxias, Rio de Janeiro, Brazil**
[1]m.d.mikhailov@gmail.com, [2]vyaibe@inmetro.gov.br, [3]ivanyasuda@gmail.com

## I - SUMMARY

The Guide to the Expression of Uncertainty in Measurement (GUM) [1] describes the standard approach for assessing uncertainty. By considering the law of uncertainty propagation, it tells us how to compute variance of functions of several random variables. Although the method is straightforward, it can became rather involved depending on the complexity of the model. In this work we develop a package for *Mathematica* [1] software system which aims to render the procedure of calculating uncertainty simple and transparent for the end user.

In this paper we present two functions of our package. The first function, called UncertaintyGUM function, follows the guidelines established in GUM, in that all calculations are performed after Taylor expanding integrands. It also takes into account second order corrections, as well as correlation between variables.

Because the law of uncertainty propagation relies on Taylor series approximation, GUM method only is exact for linear models. In order to circumvent this limitation, the Supplement of the above guide [2] suggests a complementary method based on Monte Carlo simulations. The second function of our package, called Uncertainty function, addresses this type of problem by using the method of propagation of distributions. It finds the desired solution both analytically, whenever possible, and numerically, when the variables follow prescribed distributions.

One special application of the Uncertainty function concerns a problem treated in the book by B. Christianson and M. Cox [3] for which computation of uncertainty by Monte Carlo method would be problematic. The question raised by those authors is how one computes uncertainty of $exp(x^2)$ when $x$ is normally distributed. Here we show instead that the problem admits perfectly consistent solutions provided that the standard deviation of $x$ lies in a well defined interval.

The paper is organised as follows. In the next section we give preliminary information concerning the documentation of the functions. In section III we illustrate the use of the package through both symbolic and numerical calculations. In section IV we consider the problem of finding real solutions for expected value and uncertainty of the measurand $exp(x^2)$ when $x$ is normally distributed. We conclude in section V.

1 See http://www.wolfram.com

## II - PRELIMINARIES

Before considering concrete examples, it is useful to reproduce the documentation of both functions:

### UncertaintyGUM Function

**UncertaintyGUM** [ $f[x]$, $x == \mu \pm \sigma$ ] or
**UncertaintyGUM** [ $f[x_1,...,x_n]$,{$x_1==\mu_1 \pm \sigma_1$,..., $x_n==\mu_n \pm \sigma_n$}, *case*]
gives first order series approximation of $f$ (expectation ± standard deviation) .
*case* = 1 or IdentityMatrix [ $n$ ] (uncorrelated variables),
*case* = 2 nonlinear $f$ (more series terms included in case 1),
*case* = 3 or $n$ x $n$ constant matrix $r_{i,j}$ = 1 (fully correlated variables),
case = $n$ by $n$ matrix: $r_{i,j} = r_{j,i}$, $r_{i,i} = 1$ , $-1 \leq r_{i,j} \leq 1$ .

### Uncertainty Function

**Uncertainty**[*expr*, *case*] in domain of symbols, rationals or reals attempts to calculate uncertainty (expectation ± standard deviation) of *expr*.
**Uncertainty**[*expr*, *case*, *n*] in domain of reals calculates uncertainty by using *n* pseudorandom variates.
If *case* is $x \sim dist$ then $x$ follows the probability distribution *dist*.
If *case* is {$x_1 \sim dist_1$, ... , $x_n \sim dist_n$} then $x_1$ to $x_n$ are independent and follow the distributions $dist_1$ to $dist_n$ .
If *case* is {$x_1$ , ... , $x_n$} ~ *dist* then $x_1$ *to* $x_n$ follows the multivariate distributions *dist*.

## III - THE *MATHEMATICA* FUNCTIONS

We begin by considering examples which both will give us insight on the applications of the uncertainty functions and will make clearer their ability to find exact solutions.

It is convenient to treat the functions separetely.

### UncertaintyGUM Function

Consider the following cubic equation:

$$y\left(x_1, x_2\right) = x_1^3 + x_2^3\,. \qquad \text{(Equation 1)}$$

We wish to calculate the combined uncertainty of the model $y$, $u(y)$, as outlined in GUM [1]. To begin with, we first consider the case where the input quantities $x_1$ and $x_2$ are noncorrelated. The case of correlated variables will be considered in the sequel.

The UncertaintyGUM function gives us the following solution for *u(y)*:

$$\text{UncertaintyGUM}\left[x_1^3+x_2^3,\{x_1==\bar{x}_1\pm\sigma_{x_1},\,x==\bar{x}_2\pm\sigma_{x_2}\},2\right]$$

$$\left(\bar{x}_1^3+\bar{x}_2^3\right)\pm 3\sqrt{\sigma_{x_1}^2\bar{x}_1^4+\sigma_{x_2}^2\bar{x}_2^4+4\sigma_{x_1}^4\bar{x}_1^2+4\sigma_{x_2}^4\bar{x}_2^2}$$

(Equation 2)

Here, $\bar{x}_1$ and $\bar{x}_2$ are the expected values of the input quantities $x_1$ and $x_2$ , respectively, with uncertainties $\sigma_{x_1}$ and $\sigma_{x_2}$ .

Let us comment on the above result, Eq. (2). First of all, we clarify the use of the arguments. The first one is the model, while the second argument specifies the mean values of the input quantities, as well as their corresponding uncertainties. The last argument tells the function to compute the combined uncertainty up to second order in the standard deviations. This explains the presence of the last two terms under square root sign (we could ommit those terms by telling the function to retain dominant terms only, in which case integer number 2 in the last argument should be replaced by 1.)

Secondly, since the law of uncertainty propagation relies on Taylor series expansion, and that the measurand, Eq. (1), is nonlinear, the above result is not exact. Moreover, it only holds for small uncertainties. By using the Uncertainty function, based on the method of propagation of distributions, we will show below how to compute the combined uncertainty of *y* exactly (provided that the solution exists.)

Before considering the case of exact solutions, let us compute the combined uncertainty of *y,* this time by taking into account correlation between variables. As outlined in GUM [1], statistical dependence between input quantities are expressed in terms of the matrix of correlation coefficients. For our purposes, it acts as the third argument of UncertaintyGUM function, which gives:

$$\text{UncertaintyGUM}\left[x_1^3+x_2^3,\{x_1==\bar{x}_1\pm\sigma_{x_1},\,x==\bar{x}_2\pm\sigma_{x_2}\},\{\{1,c_{12}\},\{c_{21},1\}\}\right]$$

$$\left(\bar{x}_1^3+\bar{x}_2^3\right)\pm 3\sqrt{\sigma_{x_1}^2\bar{x}_1^4+2c_{12}\sigma_{x_1}\sigma_{x_2}\bar{x}_1^2\bar{x}_2^2+\sigma_{x_2}^2\bar{x}_2^4}$$

(Equation 3)

Here, $c_{12}$ is the correlation coefficient of variables $x_1$ and $x_2$ . As previously mentionned, the above solution is not exact. This motivates us to consider the problem of looking for exact solutions for the combined uncertainty of systems with two correlated random variables.

## Uncertainty Function

We take the model of Eq. (1), and suppose that the correlated variables $x_1$ and $x_2$ follow a multinormal distribution. Making use of the Uncertainty function, we readily obtain the following exact solution:

$$\text{Uncertainty}\left[x_1^3+x_2^3,\{x_1,x_2\}\approx\text{MultinormalDistribution}\left[\{\bar{x}_1,\bar{x}_2\},\left\{\left\{\sigma_{x_1}^2,\sigma_{x_1}\sigma_{x_2}\right\},\left\{\sigma_{x_2}\sigma_{x_1},\sigma_{x_2}^2\right\}\right\}*\{\{1,c_{12}\},\{c_{12},1\}\}\right]\right]$$

$$\begin{aligned}&\bar{x}_1^3+\bar{x}_2^3+3\left(\bar{x}_1\sigma_{x_1}^2+\bar{x}_2\sigma_{x_2}^2\right)\pm\\&\sqrt{3}\left\{\left[3\left(\sigma_{x_1}^2\bar{x}_1^4+2c_{12}\sigma_{x_1}\sigma_{x_2}\bar{x}_1^2\bar{x}_2^2+\sigma_{x_2}^2\bar{x}_2^4\right)+\right.\right.\\&+12\left(\sigma_{x_1}^4\bar{x}_1^2+c_{12}^2\sigma_{x_1}^2\sigma_{x_2}^2\bar{x}_1\bar{x}_2+\tfrac{c_{12}}{2}\left(\sigma_{x_1}\sigma_{x_2}^3\bar{x}_1^2+\sigma_{x_1}^3\sigma_{x_2}\bar{x}_2^2\right)+\sigma_{x_2}^4\bar{x}_2^2\right)+\\&\left.\left.+5\left(\sigma_{x_1}^6+\sigma_{x_2}^6\right)+2c_{12}\sigma_{x_1}^3\sigma_{x_2}^3\left(3+2c_{12}^2\right)\right]\right\}^{\frac{1}{2}}\end{aligned}$$

(Equation 4)

It is easy to verify that Eq. (4) reduces to Eq. (3) in the limit of very small uncertainties.

The above examples show us how one can use UncertaintyGUM and Uncertainty functions to obtain the combined uncertainty. They are generic, in the sense that they deal with independent as well as correlated variables. The difference is that in the former the result is valid for small standard deviations, while in the latter exact solutions can eventually be computed. This is what one should expect, since in the first case uncertainty is computed by means of Taylor series expansion of moments, while in the second case propagation of distributions is used instead. Finally, we mention that the previous computations can be performed in real or rational domains.

Uncertainty Function and Monte Carlo simulations

Now we go to the second application of Uncertainty function, namely, computation of expected value and combined uncertainty by using random numbers generation [2].

We take the following model as example:

$$y\left(x_1,x_2\right)=x_1^3\cos^2\left(\frac{x_1}{x_2}\right)+x_2^3\sin^2\left(\frac{x_1}{x_2}\right) \qquad \text{(Equation 5)}$$

We suppose that the correlated input quantities are given by $x_1=1.000\pm0.001$, $x_2=1.000\pm0.001$ and follow multinormal distribution. The correlation coefficients are $c_{12}=0.1=c_{21}$. Telling Uncertainty function to perform $10^6$ random variates, one obtains:

$$\text{Timing}\left[\text{Uncertainty}\left[x_1^3\text{Cos}\left[\frac{x_1}{x_2}\right]^2+x_2^3\text{Sin}\left[\frac{x_1}{x_2}\right]^2,\{x_1,x_2\}\approx\text{MultinormalDistribution}\left[\{1,1\},\left\{\left\{0.001^2,0.001^2\right\},\left\{0.001^2,0.001^2\right\}\right\}*\{\{1,.1\},\{.1,1\}\},10^6\right]\right]\right]$$

$$\{0.516032,\ 0.999996\pm0.00237918\}$$

(Equation 6)

The output (the last line of Eq. (6)) shows a list in which the first element denotes the time (in seconds) taken by our personal Intel i7 platform computer to perform the random variate. The second element is the computed result. Obviously, repeated computations do not give the same result, for Monte Carlo cycles are supposed to be random. We can easily verify this assertion by generating a table with repeated $10^6$ random numbers generation. We have:

$$\text{Timing}\left[\text{Table}\left[\text{Uncertainty}\left[x_1^3\text{Cos}\left[\frac{x_1}{x_2}\right]^2 + x_2^3\text{Sin}\left[\frac{x_1}{x_2}\right]^2, \{x_1, x_2\} \approx \text{MultinormalDistribution}[\{1,1\}, \{\{0.001^2, 0.001^2\}, \{0.001^2, 0.001^2\}\} * \{\{1, .1\}, \{.1, 1\}\}, 10^6]\right], \{6\}\right]\right]$$

$$\{5.776361, \{0.999998 \pm 0.0023781, 0.999998 \pm 0.00237883, 0.999997 \pm 0.00237706, 0.999997 \pm 0.00237665, 1. \pm 0.00237792, 0.999998 \pm 0.00237956\}\} \quad \text{(Equation 7)}$$

Again, in the output of Eq. (7), the first element is the time taken by the function to generate 6 cycles of random numbers, each with $10^6$ elements. As previously argued, the results in the second element of the output list are not equal.

It is instructive to compare values of Eqs. (6) and (7) with results we obtain using the UncertaintyGUM function. We have:

$$\text{N}\left[\text{UncertaintyGUM}\left[x_1^3\text{Cos}\left[\frac{x_1}{x_2}\right]^2 + x_2^3\text{Sin}\left[\frac{x_1}{x_2}\right]^2, \{x_1 == 1. + .001, x_2 == 1. + .001\}, \{\{1, .1\}, \{.1, 1\}\}\right]\right]$$

$$1.00000000 \pm 0.00237726 \quad \text{(Equation 8)}$$

If we compare Eqs. (7) and (8), then we immediately see that both results are compatible for one part in $10^5$.

In the next section we will consider a special Uncertainty function application which will lead us to impose value constraints on standard deviation of normal distributed variables.

## IV - σ Constraints for $y(x) = exp(x^2)$

In a chapter of their book [3], B. Christianson and M. Cox consider an extreme problem for which "Monte Carlo techniques are computationally problematic" [3]. The question is to find the combined uncertainty of the following measurand:

$$y(x) = exp(x^2), \quad \text{(Equation 9)}$$

with $x$ being a normal distributed random variable, and

$$\sigma = 1$$
$$\mu = 0 \quad \text{(Equation 10)}$$

By using Uncertainty function, the problem as evoked in [3] becomes trivial, for when the measurand is given by Eq. (9), we have the following analytical solution:

$$\text{Uncertainty}\left[\text{Exp}\left[x^2\right], x \approx \text{NormalDistribution}[\mu, \sigma]\right]$$

$$\frac{e^{\frac{\mu^2}{1-2\sigma^2}}}{\sqrt{1-2\sigma^2}} \pm \sqrt{\frac{e^{\frac{2\mu^2}{1-4\sigma^2}}}{\sqrt{1-4\sigma^2}} - \frac{e^{\frac{2\mu^2}{1-2\sigma^2}}}{1-2\sigma^2}} \quad \text{(Equation 11)}$$

Replacing σ and μ by their corresponding values of Eq. (10), we obtain a complex solution:

$$u(y) = \frac{\sqrt{-i\left(1+i\sqrt{3}\right)}}{3^{\frac{1}{4}}} \quad \text{(Equation 12)}$$

Therefore, when the measurand is given by Eq. (9), it is not sufficient for σ to be positive. This means that, in order to obtain consistent solutions, we must impose additional constraints on the real, positive values of σ. Also, note that the expected value of *y(x)* (see Eq. (11)) becomes a complex number when σ and μ assume the values of Eq. (10). With those knowledges on hand, as will be shown below, computation of combined uncertainty of *y* by Monte Carlo method becomes ordinary.

Before doing this, let us deduce the real domain of σ. The first obvious condition is that σ must be positive definite. In addition, it follows from Eq. (6) that, in order to have real solutions, we must impose:

$$1 - 4\sigma^2 > 0 \quad \text{and} \quad \text{(Equation 13)}$$
$$1 - 2\sigma^2 > 0, \quad \text{(Equation 14)}$$

from which we deduce, together with $\sigma > 0$, that:

$$0 < \sigma < \tfrac{1}{2} \quad \text{(Equation 15)}$$

In order to verify that real solutions exist whenever $\sigma \in \left(0, \frac{1}{2}\right)$, we plot contour lines of both expected value and uncertainty of *y(x)*.

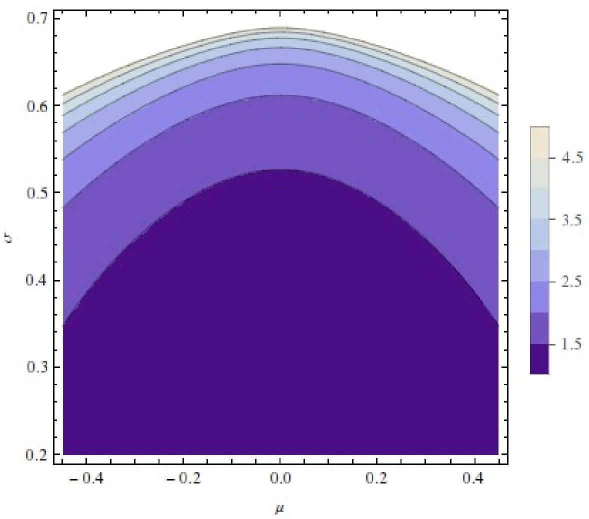


Figure 1: *y(x)* expected value contour lines. The solutions are real provided that $0 < \sigma < 1/2^{1/2}$.

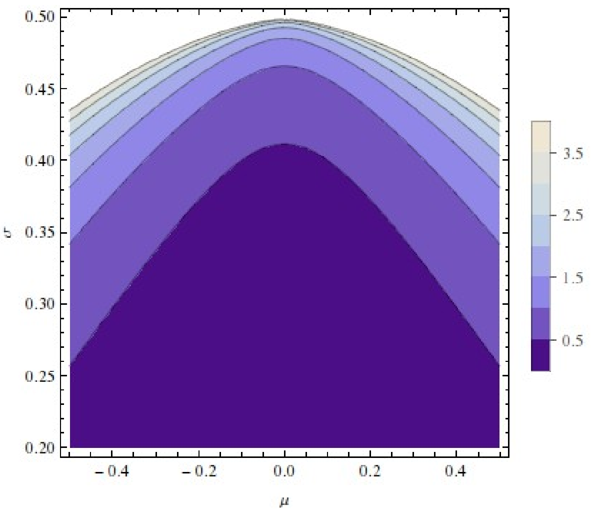


Figure 2: *y(x)* uncertainty contour lines. The solutions are real provided that $0 < \sigma < 1/2$.

Having established the feasible region for σ, we are ready to compute uncertainty of *y* in real domain both exactly and numerically.

For this purpose, we take μ = 0 and σ = 0.4. Uncertainty function gives us

$$\text{Timing}\left[\text{Uncertainty}\left[\text{Exp}\left[x^2\right], x \approx \text{NormalDistribution}[\mu, \sigma]\right]\right]$$

$$\{0.312019,\ 1.21268 \pm 0.442807\} \qquad \text{(Equation 16)}$$

As usual, the first element denotes time taken for the CPU to perform the computation. The second element is the desired exact solution in real domain.

After employing Monte Carlo method, we obtain Table 1 below for less than a second, when combined uncertainty is computed using 100, 1000, 10000, 100000 and 1000000 random numbers generations. As we can see, the results smoothly approache the exact solution of Eq. (16) as the number of variates increases.

| Random Variates | Expected Value ± Combined Uncertainty |
|---|---|
| 100 | 1.14894 ± 0.228561 |
| 1000 | 1.20845 ± 0.390797 |
| 10000 | 1.21245 ± 0.418098 |
| 100000 | 1.21134 ± 0.436378 |
| 1000000 | 1.21324 ± 0.447601 |

Table 1: Computed combined uncertainty by Monte Carlo method. The results smoothly approache the exact solution of Eq. (16) as the number of variates increases.

## IV – CONCLUSION

We developed a *Mathematica* package for uncertainty analysis. The aim is to compute expected values and combined uncertainty of models with several random variables.

The functions follow GUM guidelines in that law of uncertainty propagation [1] as well as Monte Carlo cycles [2] are evoked. In relationship to the first method, the function UncertaintyGUM deals with correlated as well as noncorrelated variables. In addition, it computes combined uncertainty with higher order corrections.

One advantage of using the second method is that one is able to compute uncertainty for nonlinear models. The function Uncertainty computes combined uncertainty by generating multivariate random numbers. Moreover, by using the method of propagation of distribution, it also computes exact solutions whenever possible.

One interesting application of exact computation is that for which the model is given by Eq. (5). We have seen that real solutions exist provided that the standard deviation of the input quantity lies in a well defined interval. This treatment completes the argument presented in reference [3].

Our *Mathematica* package is a work in progress. Many other uncertainty analysis functions, as well as uncertainty calculators for CDF player [2], have been implemented [4, 5, 6, 7, 8, 9]. Among them we may cite another *Mathematica* function which allows us to compute coverage intervals. We will present them in future occasion. Concerning the demos for CDF player, the interested reader may have acces to them from the web page of Wolfram Demonstration Project [3, 4, 5, 6, 7, 8, 9].

## REFERENCES

[1] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML, Evaluation of Measurement Data – Guide to the Expression of Uncertainty in Measurement, JCGM 100:2008.

[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, JCGM 101:2008.

[3] B. Christianson and M. Cox, Automatic Propagation of Uncertainties, Springer, 2005, ch. 3, pp. 47 - 58.

[4] V. Y. Aibe and M. D. Mikhailov, Solving Matrix Systems with Real Interval or Uncertain Elements, in Wolfram Demonstration Project, http://demonstrations.wolfram.com/SolvingMatrixSystemsWithRealIntervalOrUncertainElements/

[5] V. Y. Aibe and M. D. Mikhailov, Solving a Linear System with Uncertain Coefficients, in Wolfram Demonstration Project, http://demonstrations.wolfram.com/SolvingALinearSystemWithUncertainCoefficients/

2 See the website of this project: http://www.wolfram.com/cdf-player

[6] V. Y. Aibe and M. D. Mikhailov, Area Perimeter and Diagonal of a Rectangle with Uncertain Sides, in Wolfram Demonstration Project, http://demonstrations.wolfram.com/AreaPerimeterAndDiagonalOfARectangleWithUncertainSides/

[7] V. Y. Aibe and M. D. Mikhailov, Fit Interpolation or Polynomial Interpolation in Uncertain Calculus, in Wolfram Demonstration Project, http://demonstrations.wolfram.com/FitInterpolationOrPolynomialInterpolationInUncertainCalculus/

[8] M. D. Mikhailov and V. Y. Aibe, GUM *versus* Exact Uncertainty of Sin(x) and Cos(x) when x is Normally Distributed, in Wolfram Demonstration Project, http://demonstrations.wolfram.com/GUMVersusExactUncertaintyOfSinXAndCosXWhenXIsNormallyDistrib/

[9] M. D. Mikhailov, V. Y. Aibe and I. Yasuda, Percentiles of Continuous Probability Distributions, in Wolfram Demonstration Project, http://demonstrations.wolfram.com/PercentilesOfContinuousProbabilityDistributions/