Extended Uniform Distribution Accounting for Uncertainty of Uncertainty

James-A. Goulet[1], Ian F.C. Smith[2]

[1]Swiss Federal Institute of Technology (EPFL), IMAC Laboratory, EPFL ENAC IS IMAC, Station 18, CH-1015, Lausanne, Switzerland; James.A.Goulet@gmail.com

[2]Swiss Federal Institute of Technology (EPFL), IMAC Laboratory, EPFL ENAC IS IMAC, Station 18, CH-1015, Lausanne, Switzerland; Ian.Smith@epfl.ch

ABSTRACT

Increasingly, uncertainties are explicitly considered for important engineering tasks. Often, little case-specific information is available for characterizing these uncertainties. Uniform distributions are an easy way to describe errors in absence of more precise information. In many situations, the bounds are fixed based on user experience. The extended uniform distribution (EUD) provides a probability density function that accounts for higher orders of uncertainty (uncertainty of uncertainty) when using a uniform distribution to describe errors. Since the EUD accounts for several orders of uncertainty it is more representative than uniform and curvilinear distributions. The extended uniform distribution helps increase the reliability and robustness of tasks requiring uncertainty combination through better representing incomplete knowledge of parameters.

KEYWORDS

Uncertainties, Extended uniform distribution, System Identification, CMS4SI

Uncertainty is a fundamental part of applied science research. Uncertainty is usually used to describe the distribution of an error through its probability density function (PDF). An error PDF may either be used as itself or it may be combined with other sources of uncertainties. This is called propagation of uncertainties. Several methods are available to propagate uncertainties through models (JCGM 2008a; JCGM 2008b). These methods involve a combination of model-parameter uncertainties into a single probability density function describing the overall uncertainty of model predictions. Propagation methods assume that the uncertainties are adequate representations of the error. In many cases, little information is available for characterizing uncertainties. In absence of knowledge other than the position of minimal and maximal error bounds, the uniform distribution is often chosen according to the principle of maximum entropy (Jaynes 1957). Bounds are usually defined based on user experience. Therefore, an uncertainty can be assigned to the position of minimal and maximal bounds. These three distributions can be combined (the main uncertainty plus the two uncertainties on the minimal and maximal bound position) into a curvilinear probability distribution function (Lira 2008; Raghu and James 2007; Raghu and James 2010). Figure 1 shows an example of this function.

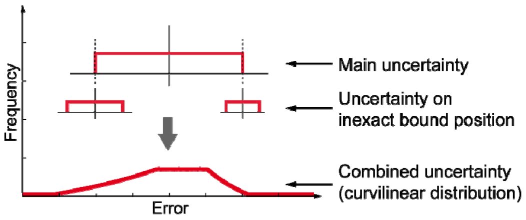

Figure 1 – Example of a Curvilinear Probability Distribution Function

This distribution was adopted in a metrology guideline for expression of uncertainty in measurements as the probability density function to use when the upper and lower limit of an uniform distribution are not known exactly (JCGM 2008a).

The curvilinear distribution represents a more robust representation of uncertainty when compared with the uniform distribution. However, one question remains; are the bounds of the uncertainty on bound positions exactly known? Since the answer is rarely positive, the concept behind the curvilinear distribution needs to be extended in order to account for the inexact position of bounds for higher orders.

This paper introduces the extended uniform distribution (EUD) which overcomes the limitation mentioned above. The first part presents the concepts behind EUD. The second section explains how samples can be drawn from the extended uniform distribution. The third section presents the result of a comparative study between uniform, EUD and the curvilinear distribution and shows how curvilinear distribution is a special case of EUD. In this section, the impact of the number of orders of uncertainty accounted for is studied. Finally, the last section provides a discussion of the results obtained and the use of the extended uniform distribution.

Extended Uniform Distribution

The Extended Uniform Distribution (EUD) accounts for the uncertainty over the bound position for multiple orders of uniform distributions. Figure 2 shows the resulting probability density function obtained using EUD.

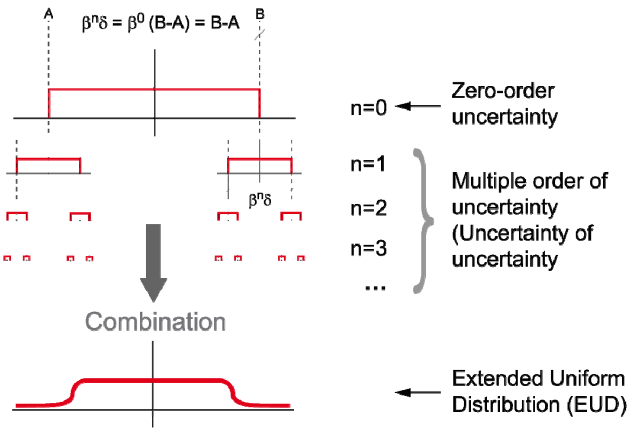

Figure 2 - Extended Uniform Distribution

In this figure, the main uncertainty (n=0) on a given quantity is expressed by a uniform distribution. For illustration purposes, assume that it represents the probability that a modelling error has a certain value. Since errors are never known exactly, the position of the lower and higher bounds (A, B, A≤B) of this distribution are fixed based upon engineering experience. A constant β expresses the uncertainty on the bound positions as a fraction of the zero-order uncertainty varying between zero and one. For the zero-order uncertainty: $\beta^n\delta = \beta^0(B\text{-}A) = B\text{-}A$. The first order of uncertainty (n=1) accounts for incomplete knowledge of the exact position of each bound. For the purpose of illustration, it is assumed that the uncertainty on the lower and upper bounds are the same. The width of the uniform distributions representing the uncertainty on the position of bound A and B is: $\beta^n\delta = \beta\delta = \beta(B\text{-}A)$. The combination of these uncertainties would lead to a curvilinear distribution. However, as mentioned above, the knowledge on the position of the bounds for the order one (n=1) is also incomplete. Therefore, the second order of uncertainty (n=2) accounts for incomplete knowledge of order one. The width of the uniform distributions representing the uncertainty on the position of bound of the first order is: $\beta^n\delta = \beta^2\delta = \beta^2(B\text{-}A)$. Combining general uncertainty orders leads to the extended uniform distribution (EUD). Compared with the uniform distribution, the EUD better represents the lack of knowledge related to the upper and lower-bound positions when user experience is used to fix uncertainties.

Sampling From an Extended Uniform Distribution

The extended uniform distribution may be used in a Monte-Carlo uncertainty combination process as described in the Supplement 1 of JCGM (2008b). In this section a sampling procedure is presented for the case where the uncertainties on the lower and higher bounds are equal (a single value of β is required). These guidelines also allow for unequal uncertainties.

The number of orders of uncertainty has to be fixed initially in order to obtain a stable EUD probability density function. This step is required since it is not possible to sample over an infinite number of uniform PDF in a numerical process. It is possible to obtain a good approximation using a limited number of orders since the influence of each order decreases exponentially. A study of the number of orders is presented in Section 0. The sum of the maximal bound position for an infinite number of orders, as shown in Equation 1, converge to a finite limit. As n becomes larger, the contribution of each order tends to zero

$$\sum_{n=0}^{\infty} \frac{\beta^n (B-A)}{2} \qquad (1)$$

In this equation, A and B are respectively the lower and upper bound of the main uncertainty described by a uniform distribution and where β varies between zero and one.

To generate EUD samples, a sparse matrix M of size [NBO, 2^NBO] is created, where NBO is the number of orders. For each row i=1..NBO, the columns 1 to 2^i are filled with uniformly distributed number between 0 and 1. For one sample, the uncertainty propagation process starts with the last row (highest uncertainty order number) and propagates up to the main uniform distribution (first row). The algorithm that propagates uncertainties is presented below.

```
for i=[NBO:-1:2]
      for j=1:2:(2^i)
            beta_loop=((beta/2)^(i-1))*(B-A)
            M(i,j)=M(i,j)-beta_loop
            M(i,j+1)=M(i,j+1)+beta_loop
            M(i-1,(j-1)/2+1)=M(i,j)+(M(i,j+1)-M(i,j))*rand(1)
      end
end
```

Logically, the upper bound for the error must always remain larger than its lower bound. In the stochastic process for sampling in the EUD, it could be possible to generate uncertainty of uncertainty that do not respect that criterion. Therefore, the second part of the code verifies that at zero-order the upper bound (B) is larger or equal to the lower bound (A). If the condition is fulfilled, it generates a sample for the extended uniform distribution, if falsified, it discards the generation.

```
if M(2,1)>=M(1,1)
     EUD_sample= M(1,1)+M(2,1)-M(1,1)*rand(1)
end
```

Note that for β values smaller or equal to 0.5 the inequality is always satisfied. During Monte-Carlo analysis, the procedure presented above is often repeated several thousand times in order to achieve a target reliability.

A Comparison Between Uniform, EUD and Curvilinear distribution

This section presents a comparative study between the uniform, the EUD and the curvilinear distribution. Figure 3 show the EUD distribution obtained using only one level of uncertainty (n=1). This case is equivalent to the curvilinear distribution. In this figure, the horizontal axis represents the error and the vertical axis is the normalized probability at which each error value should be obtained. The different curves are computed from a Monte-Carlo analysis over 10,000,000 samples by varying the value for β between zero and one. When β is equal to zero, the result obtained is a uniform distribution having lower and upper bounds of -1 and 1.

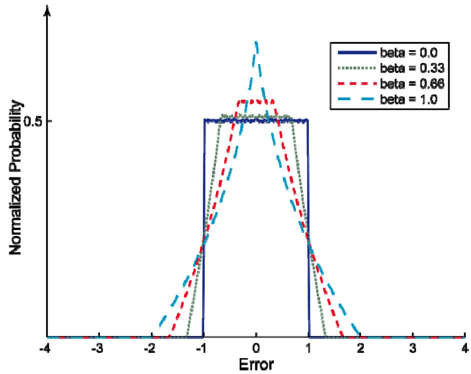

Figure 3 - Extended Uniform Distribution with one Order of Uncertainty (N=1)

The PDF shapes diverge from the uniformly distributed shape as the β value increases; all curves have an integral equivalent to unity. Apart from the small variations due to numerical sampling, the distribution shows sharp edges at their extremities and at the intersection with the constant central portion. Figure 4 shows the PDFs for the same conditions as in Figure 3, except that this time, several order of uncertainty are used.

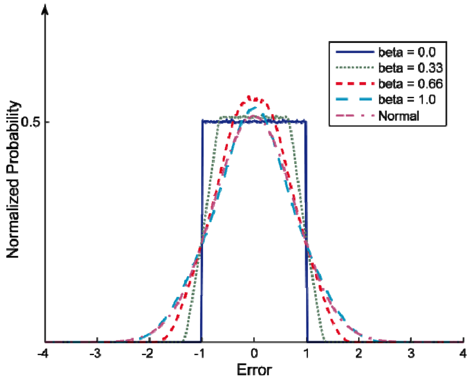

Figure 4 - Extended Uniform Distribution with Several Orders of Uncertainty (N>1)

Again, when β is equal to zero, the result obtained is a uniform distribution having lower and upper bounds of -1 and 1. For values of $\beta>0$, the EUD curves shows smoother transitions than the curvilinear distribution. When $\beta=1$ the shape obtained is close to a normal distribution (also shown in Figure 4) having the same standard deviation (as EUD($\beta=1$)).

Figure 5 presents the evolution of the 95% reliability bounds as a function of the β value and the number of orders taken into account in EUD distribution for zero-order distribution of -1 to 1. In this figure, the β value is plotted on horizontal axis and the bound defining a coverage interval of 95% for each PDF is plotted on the vertical axis. In addition to the EUD distribution using one to four levels of uncertainty, the results obtained with a uniform and a normal distribution (having the same standard deviation as "EUD – 4 orders") are shown. In the case of the uniform distribution and when $\beta=0$ the 95% reliability bound has a value of 0.95. For EUD using any number of uncertainty level, the reliability bound increases when the uncertainty over the exact bound position increases. For small values of β (<0.5) there is no significant change between a curvilinear and EUD distributions. For $\beta\geq0.5$, the discrepancy between the two distributions types (curvilinear and EUD) increases. For $\beta=1$, EUD shows a reliability bound almost equal to a normal distribution having the same standard deviation.

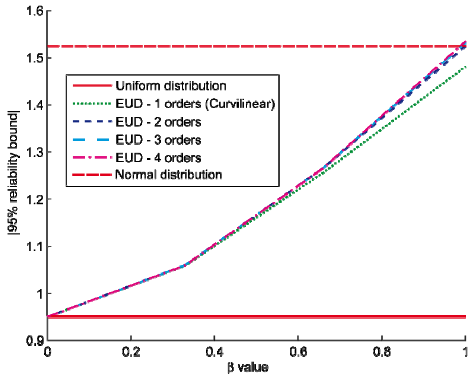

Figure 5 - Evolution of 95% reliability bounds in relation with uncertainty order and beta value for a zero-order distribution from -1 to 1

For $0 \geq \beta \geq 1$ the reliability bounds obtained from extended uniform distributions using a number of levels of uncertainty larger than one for these distributions can all be considered as equivalent for engineering purposes and the PDFs obtained are all similar to the one showed in Figure 4.

DISCUSSION

For engineering applications, two or three orders of uncertainties (as shown in Figure 5) may be sufficient to obtain a close approximation of the extended uniform distribution. Also, in every situation, the EUD distributions provide equal or conservative reliability bounds compared with the curvilinear distributions.

Uniform distributions are often used to describe an error distribution. If the knowledge related to the position of the bounds defining the distribution is incomplete, the extended uniform distribution can be used. For practical applications, when the uncertainty on the exact position of the bounds describing the distribution is large (for instance $\beta = 1$), the EUD distribution can be replaced by a normal distribution having the same mean as the main uniform distribution and a standard deviation equal to 0.39 times the zero-order interval width. This value was determined numerically from 1E7 samples.

The concept of higher order uncertainty and EUD can be used to define the probability density function of errors in system identification tasks. A methodology proposed by Goulet and Smith (2010) explicitly accounts for uncertainty coming from model and measurements in order to identify the behaviour of systems. In such a case, modelling uncertainties may only be quantified through relying on engineering experience. Therefore, the EUD may increase the utility the identification outcome.

CONCLUSIONS

1. The extended uniform distribution (EUD) provides a probability density function that accounts for uncertainty of uncertainty when using uniform distributions to describe errors. EUD accounts for several orders of uncertainty making it more representative than uniform and curvilinear distribution.
2. The extended uniform distribution has the potential to increase the reliability and robustness of decision making that requires the combination of uncertainties through better representing incomplete knowledge of parameters such as modelling uncertainty.

ACKNOWLEDGEMENTS

This research is funded by the Swiss National Science Foundation under contract no. 200020-117670/1.

REFERENCES

Goulet, J.-A., and Smith, I. F. C. (2010). "Structural Identification with Correlated Uncertainties." *Engineering Structures (Submitted for publication)*.

Jaynes, E. T. (1957). "Information Theory and Statistical Mechanics. II." *Physical Review*, 108(2), 171.

JCGM. (2008a). "Evaluation of measurement data - Guide to the expression of uncertainty in measurement." Bureau international des poids et mesures, Paris, France.

JCGM. (2008b). "Evaluation of measurement data - Supplement 1 to the "Guide to the expression of uncertainty in measurement" - Propagation of distributions using a Monte Carlo method." Paris, France.

Lira, I. (2008). "The generalized maximum entropy trapezoidal probability density function." *Metrologia*, 45(4), L17.

Raghu, N. K., and James, F. L. (2007). "Trapezoidal and triangular distributions for Type B evaluation of standard uncertainty." *Metrologia*, 44(2), 117.

Raghu, N. K., and James, F. L. (2010). "Rectangular distribution whose end points are not exactly known: curvilinear trapezoidal distribution." *Metrologia*, 47(3), 120.

