XVIII IMEKO WORLD CONGRESS
Metrology for a Sustainable Development
September, 17 – 22, 2006, Rio de Janeiro, Brazil

PROPAGATION OF MEASUREMENT UNCERTAINTY EXPRESSED BY A POSSIBILITY DISTRIBUTION WITH COVERAGE-INTERVAL-BASED SEMANTICS

Gilles MAURIS [1]

[1] LISTIC University of Savoie, Annecy, France, mauris@univ-savoie.fr

Abstract: The main purpose of this paper is to present a possibility theory-based generalization of conventional interval propagation to coverage intervals. Indeed, the whole set of coverage intervals for all the probability levels stacked on top of one another constitutes a possibility distribution. Thus by slight modifications of Zadeh's extension principle, we will prove that it is possible to compute the coverage intervals of an indirect measurement from many other measurements (possibly dependent) by a known non decreasing relationship.

Keywords: uncertainty propagation, coverage intervals, possibility theory.

1. INTRODUCTION

Uncertainty is a key concept for the measurement expression [1][2][3]. Indeed, in many application domains, take the measurement uncertainties into account is important. This will allow to define, around the measurement result, an interval which will contain an important part of the distribution of the measured values, that is, a coverage interval [4]. Such an interval further allows to define decision risks, as for example the risk to accept a defective lot, the risk to exceed an alarm threshold, etc. Moreover, when the measurement uncertainty is defined, it is also sometimes necessary to propagate it. For example, the uncertainty associated to an environment temperature measurement will have to be propagated to define the uncertainty associated to the sound speed in this environment.

A main tool to deal with sensor measurement uncertainty is statistics [5]. This tool requires a mathematical support to be used, especially to propagate uncertainties. Two main theories are considered : the interval calculus [6] and the probability theory [5]. Although the interval calculus allows simple calculations, the resulting model is very imprecise. Moreover, it only supplies the coverage interval of the 100% coverage probability. Thus, the use of the probability theory seems to be necessary to supply coverage intervals, but to handle the whole sets of coverage intervals is quite complex by a probability approach. And choosing a particular coverage probability (e.g. 95% which means a 0.05 probability for the value to be out of the interval) is rather arbitrary. Thus a possibility approach has been proposed in [7][8][9] and further developed by a few authors in a measurement context [10] [11][12][13][14].

This paper further explores the connection between a possibility distribution and the coverage intervals. It particularly addresses the propagation of uncertainty in measurement. In section 2, we explain how a possibility distribution can be built from coverage intervals. In the third section, we present the main contribution of the paper, i.e. how to propagate coverage intervals by a formula closely related to Zadeh's extension principle [15] proposed for possibility distributions. Some concluding remarks point out the interest of the approach and some future developments.

2. COVERAGE INTERVALS AND POSSIBILITY DISTRIBUTION

Let us assume that the considered variable x under measurement is continuous and is guaranteed to lie within an interval $\left[x_{\min}, x^{\max}\right]$.

2.1. Coverage interval definition

For every possible coverage probability $\beta \in [0,1]$, the corresponding coverage interval is defined as an interval that contains x with a probability exactly equal to β. In other words, a coverage interval of coverage probability level β (denoted $I_\beta = \left[x_{low}(\beta), x^{up}(\beta)\right]$) is defined as an interval for which the probability p_{out} of being outside this interval I_β is exactly $\alpha \overset{def}{=} 1-\beta$ [16].

Note that some authors consider the closely related notion of *statistical coverage interval* for which it can be stated with a given coverage probability that it contains *at least* (instead of exactly) a specified proportion of the population [16]. Here, we will consider the notion of *statistical coverage intervals* but we will refer to them as *coverage intervals*.

An important point is that the notion of coverage intervals relies only on the theory of probability without importing any principle of inference, as is the case for the notions of confidence, credible or fiducial intervals.

In fact, there are two possibilities for x to be outside the interval $\left[x_{low}(\beta), x^{up}(\beta)\right]$:

- when x is smaller than the lower bound $x_{low}(\beta)$ and
- when x is larger than the upper bound $x^{up}(\beta)$.

Thus, the probability p_{out} that x is outside the confidence interval is equal to the sum of two probabilities:

- the probability that x is smaller than the lower bound $x_{low}(\beta)$ and,
- the probability that x is larger than the upper bound $x^{up}(\beta)$.

2.2. Different types of coverage intervals

For the same statistical distribution and for the same coverage probability, we can have different types of coverage intervals. Indeed, we can impose that the coverage intervals are defined around a same point x^*, generally a "one-point" estimation of the "true" value, for example the mean, the median or the mode of the statistical distribution.

For example, we can set

- $x_{low}(\beta) = x_{\min}$ and find the value $x^{up}(\beta)$ for which the probability $\Pr\left[x \le x^{up}(\beta)\right]$ that $x \le x^{up}(\beta)$ is equal to β,
- or $x^{up}(\beta) = x^{\max}$ and find the value $x_{low}(\beta)$ for which the probability $\Pr\left[x \ge x_{low}(\beta)\right]$ that $x \ge x_{low}(\beta)$ is equal to β,
- or also we can impose the symmetry, i.e. to find $x^{sym}_{low}(\beta)$ for which the probability that $x \le x^{sym}_{low}(\beta)$ is equal to $\alpha/2 = (1-\beta)/2$ and $x^{up}_{sym}(\beta)$ for which the probability that $x \ge x^{up}_{sym}(\beta)$ is also equal to $\alpha/2$.

There are many other options, but in practice the most used situations is the last one because it leads to symmetric intervals $\left[x^{sym}_{low}(\beta), x^{up}_{sym}(\beta)\right]$ around the median. Moreover, for symmetric distributions, these intervals are the coverage intervals which have the shortest length. Note also that the different possibilities express the same information, albeit in different forms, but we can reconstruct one from another for continuous distribution as presented hereafter.

If we know the value $x^{up}(\beta)$ for all β, we can reconstruct :

- $x_{low}(\beta)$ as $x^{up}(1-\beta)$, and
- $x^{sym}_{low}(\beta)$ and $x^{up}_{sym}(\beta)$ as, respectively $x^{up}((1-\beta)/2)$ and $x^{up}(1-(1-\beta)/2)$.

If we know the value $x_{low}(\beta)$ for all β, we can reconstruct :

- $x^{up}(\beta)$ as $x_{low}(1-\beta)$, and
- $x^{sym}_{low}(\beta)$ and $x^{up}_{sym}(\beta)$ as, respectively: $x_{low}(1-(1-\beta)/2)$ and $x_{low}((1-\beta)/2)$.

If we know the value $x^{sym}_{low}(\beta)$ and $x^{up}_{sym}(\beta)$ for all β, then :

- when $\beta \ge 0.5$, we can reconstruct $x_{low}(\beta)$ as $x^{sym}_{low}(2\beta-1)$ and $x^{up}(\beta)$ as $x^{up}_{sym}(2\beta-1)$,
- when $\beta \le 0.5$, we can reconstruct $x_{low}(\beta)$ as $x^{up}_{sym}(1-2\beta)$ and $x^{up}(\beta)$ as $x^{sym}_{low}(1-2\beta)$.

Thus it is sufficient to consider only one of these representations, e.g. the representation by intervals $\left[x_{low}(\beta), x^{\max}\right]$.

2.3. Possibility representation

Obviously, the coverage intervals built around the same point x^* are nested. It has been proven in [16] that stacking coverage intervals on top of one another leads to a possibility distribution (denoted π^* having x^* as modal value). In fact, this way, the α-cuts of π^*, i.e. $A_\alpha = \{x / \pi^*(x) \ge \alpha\}$ are identified to the coverage intervals I^*_β of coverage probability $\beta = 1-\alpha$ around the nominal value x^*.

Thus, the possibility distribution π^* encodes the whole set of coverage intervals in its membership function. Moreover, this possibility distribution satisfies:

$$\forall \mathrm{A} \subset \mathrm{R}, \Pi^*(\mathrm{A}) \ge \mathrm{P}(\mathrm{A})$$

with Π^* and P the possibility and probability measures associated respectively to π^* and p (the underlying probability density function of the statistical measurements).

For example, if we consider $x^{\max}$ as the nominal point to build the coverage intervals, then we obtain for every $x_{low}(\beta)$ belonging to $\left[x_{\min}, x^{\max}\right]$, $\pi^{x^{\max}}(x) = \beta / x_{low}(\beta) = x$. In particular we have $\pi^{x^{\max}}(x_{\min}) = 0$ and $\pi^{x^{\max}}(x^{\max}) = 1$. Note that in this case, if the variable is described by a probability density function p, then the possibility distribution collecting all the coverage intervals is related to the probability distribution function F corresponding to p by:

$$\pi^{x^{\max}}(x) = F(x).$$

If F is symmetric, the stacking up of the symmetric coverage intervals around the distribution mode (which is equal to the median and to the mean) leads to the following possibility distribution expression:

$$\text{for } x \le x_m, \pi^{x_m}(x) = 2F(x)$$

$$\text{for } x \ge x_m, \pi^{x_m}(x) = 2(1-F(x))$$

For example for a uniform probability distribution having $\left[x_{\min}, x^{\max}\right]$ as support, we find a triangular possibility distribution (see figure 1):

$$\text{for } x \le x_m, \pi^{x_m}(x) = \frac{x - x^{\min}}{x_m - x^{\min}},$$

for $x \geq x_m, \pi^{x_m}(x) = \frac{x - x^{\max}}{x_m - x^{\max}}$.

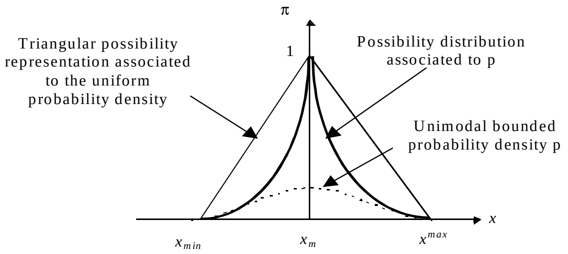

Figure 1: Example of possibility distributions

Note that the triangular possibility distribution dominates all the possibility distributions associated to all unimodal probability distributions having the same support $\left[x_{\min}, x^{\max}\right]$ and the same mode [17].

3. PROPAGATION OF POSSIBILITY

As some quantities y are difficult to measure directly, we often measure quantities $x_1,...,x_n$ related to y by a known relationship $y = f(x_1,...,x_n)$. When x_i is described by coverage intervals gathered in the associated possibility distribution, it is thus required to produce the resulting coverage intervals for y (and the associated possibility distribution). Indeed, to get a coverage interval that contains y with probability 85% for example, we cannot simply apply interval computations to the intervals that contain respectively x_1 and x_2 with probability 85%; the resulting interval would include y with the probability of the joint event that is $< 85\%$.

We will see hereafter that the formulas for computing such propagated coverage intervals are closely related with the formulas for processing possibility distributions by using Zadeh's extension principle [15]. In this paper, we will assume that the function f is a non-decreasing function of all its variables.

We consider first the case of independent variables and then the case of possibly dependent variables and we use the coverage interval representation $\left[x_{low}(\beta), x^{\max}\right]$, i.e. $\pi^{x^{\max}}(x) = F(x)$ for demonstrating the proposed results.

3.1. Result for the independent case

Let us set α and try to find for $y = f(x_1,...,x_n)$, the value $y_{low}(\alpha)$ for which the probability $\Pr\left[y \leq y_{low}(\alpha)\right]$ is equal to α.

Let us remark that for every tuple $(\alpha_1,...,\alpha_n)$, we have:

- $x_1 \notin \left[x_{1low}(\alpha_1), x_1^{\max}\right]$ i.e., $x_1 \leq x_{1low}(\alpha_1)$ with probability α_1;
- $x_2 \notin \left[x_{2low}(\alpha_2), x_2^{\max}\right]$ i.e., $x_2 \leq x_{2low}(\alpha_2)$ with probability α_2;
- …
- $x_n \notin \left[x_{nlow}(\alpha_n), x_n^{\max}\right]$ i.e., $x_n \leq x_{nlow}(\alpha_n)$ with probability α_n;

Since the probabilities are independent , we can therefore conclude that with the probability $\alpha_1 \times ... \times \alpha_n$ all n inequalities hold, i.e., $x_1 \leq x_{1low}(\alpha_1)$, …, and $x_n \leq x_{nlow}(\alpha_n)$. Since the function $f(x_1,...,x_n)$ is non decreasing, we can conclude with the coverage probability $\alpha_1 \times ... \times \alpha_n$ that we have:

$$y = f(x_1,...,x_n) \leq f(x_{1low}(\alpha_1),...,x_n low(\alpha_n)).$$

Thus we have a bound that bounds y from above with the probability $\alpha_1 \times ... \times \alpha_n$. In particular, if we select α_i for which the product $\alpha_1 \times ... \times \alpha_n$ is equal to α, we get the bound corresponding to the given α.

Our goal is to find the smallest of such bounds, so we can take:

$$y_{low}(\alpha) = \min_{\alpha = \alpha_1 \times ... \times \alpha_n} f(x_{1low}(\alpha_1),...,x_{nlow}(\alpha_n)).$$

Let us reformulate the above equation in terms of corresponding possibility distributions $\pi_i(x_i)$ and $\pi(y)$.

Note that by definition, $\alpha = \pi(y_{low}(\alpha))$ and $\alpha_i = \pi_i(x_{ilow}(\alpha_i))$. Thus, $y_{low}(\alpha)$ and $x_{ilow}(\alpha_i)$ are the quasi-inverse functions of $\pi(y_{low}(\alpha))$ and $\pi_i(x_{ilow}(\alpha_i))$. Note that these latter functions are non decreasing ones though they are probability distribution functions. Thus, according to the duality theorem of Frank and Schweizer [18]:

$\min_{\alpha = \alpha_1 \times ... \times \alpha_n} f(x_{1low}(\alpha_1),...,x_{nlow}(\alpha_n))$ is the quasi inverse function of $\max_{f(x_1,...,x_n) = y}(\alpha_1 \times ... \times \alpha_n)$.

Therefore:

$$\pi(y) \geq \max_{f(x_{1low},...,x_{nlow}) = y} \pi_1(x_{1low}) \times ... \times \pi_n(x_{nlow}).$$

This latter equation is in fact Zadeh's extension principle where the operator *min* has been replaced by the operator *product*.

Let us consider for example the sum of two uniform probability distributions having as support $[-1,+1]$, then the resulting probability distribution is a triangular one with support $[-2,+2]$. Using the associated possibility representation leads to triangular possibility distributions for the inputs and to the parabolic one for the resulting sum by applying Zadeh's extension principle. Moreover, this parabolic possibility distribution correspond to the triangular probability distribution.

3.2. Result for the possibly dependent case

Let us now consider the case when the variables are possibly dependent. The dependence can be expressed under the form of a copula C that relates the marginal probability distribution function to the joint probability distribution [19]: $F_y(x_1, x_2) = C(F_1(x_1), F_2(x_2))$

Thus following the same development as in the preceding paragraph, all n inequalities i.e., $x_1 \le x_{1low}(\alpha_1)$, …, and $x_n \le x_{nlow}(\alpha_n)$. hold with a probability $C(\alpha_1,...,\alpha_n)$. Since the function $f(x_1,...,x_n)$ is non decreasing, we can conclude with probability $C(\alpha_1,...,\alpha_n)$ that we have:

$$y = f(x_1,...,x_n) \le f(x_{1low}(\alpha_1),...,x_{nlow}(\alpha_n)) .$$

Thus we have a bound that bounds y from above with probability $C(\alpha_1,...,\alpha_n)$. In particular, if we select α_i for which the value $C(\alpha_1,...,\alpha_n)$ is equal to α, we get the bound corresponding to the given α.

Our goal is to find the smallest of such bounds, so we can take:

$$y_{low}(\alpha) = \min_{\alpha=\alpha_1\times...\times\alpha_n} f(x_{1low}(\alpha_1),...,x_{nlow}(\alpha_n)) .$$

According to the same theorem of Frank and Schweizer [17]:

$\min_{\alpha=\alpha_1\times...\times\alpha_n} f(x_{1low}(\alpha_1),...,x_{nlow}(\alpha_n))$ is the quasi inverse function of $\max_{f(x_1,...,x_n)=y} C(\alpha_1,...,\alpha_n)$. Therefore:

$$\pi(y) \ge \max_{f(x_{1low},...,x_{nlow})=y} C(\pi_1(x_{1low}),...,\pi_n(x_{nlow})) .$$

This latter equation is in fact Zadeh's extension principle where the operator min has been replaced by the Copula representing the dependence between the variables.

In fact it has been demonstrated that when no information is available about dependence, we have [19]:

$$\max(a+b-1,0) \le C(a,b) \le \min(a,b)$$

Let us pursue the preceding example of the sum of two uniform probability distributions but in the case where no information about dependence is available. In this case we have:

$$\min_{f(x_1,x_2)=y} C(\pi_1(x_1)+\pi_2(x_2)-1,0) \le \pi(y)$$

$$and\, \pi(y) \le \min_{f(x_1,x_2)=y} C(\pi_1(x_1),\pi_2(x_2))$$

In fact the two bounds are triangular possibility distributions which bracket the parabolic possibility distribution obtained for the independent case. (see figure 2).

3.3. Summary

In summary, the computation of coverage intervals of an indirect measurement related to direct measurements by a non decreasing function can be done by using Zadeh's extension principle, in which the *min* operator is replaced by a copula operator representing the dependence between the variables. For instance, independence is represented by the *product* operator. When no information about the relationship between the variables x_i is available, the result of propagation is bounded by the results obtained by using the extreme copulas $C(a,b) = \max(a+b-1,0)$ and $C(a,b) = \min(a,b)$ in the possibility propagation.

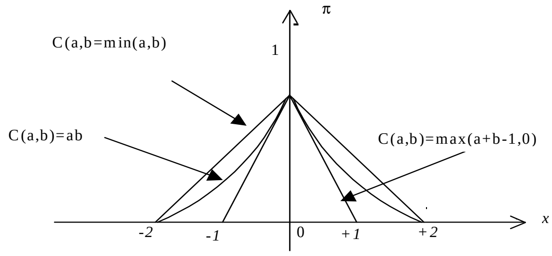

Figure 2: Possibility distribution for the sum of two uniform probability distributions for different dependence case

3.4. Discussion

In practice, dependence between the variables is often given under the form of correlation coefficients (Kendall, Spearmann,…). Thus to apply the proposed approach, a link between the correlation and the copula must be identified. Some studies have already been made to tackle this problem [19]. Another approach consists in deriving the copula directly from the data.

4. CONCLUSION

The presented relationship between nested coverage intervals and a possibility distribution is an interesting bridge between probability and possibility theories. They provide us a with tool to propagate measurement uncertainty expressed under the form of a possibility distribution by applying Zadeh's extension principle albeit with a different t-norm, i.e. a copula which represents the dependence between the variables. For instance, only propagation by a non decreasing function has been considered (e.g. addition). Future development will concern other functions.

REFERENCES

[1] L. Mari, G. Zingales, "Uncertainty in measurement science", in Proceedings of the Int. Workshop on Advances of measurement science, Kyoto, Japan, pp. 177-189, 1999.

[2] L. Mari, "Notes on the uncertainty of measurement", in Proceedings of the XV IMEKO World Congress, Osaka, Japan, Vol. II, pp. 89-92, 1999.

[3] L. Reznik, K.P. Dabke, "Evaluation of uncertainty in measurement: a proposal for application of intelligent methods, in Proceedings of the XV IMEKO World Congress, Osaka, Japan, Vol. II, pp. 93-100, 1999.

[4] Guide for the expression of uncertainty in measurement, ISO 1993, 99 pages, 1993.

[5] M. Kendall M., A. Stuart, *The advanced theory of statistics*, Ed. Griffin and Co., 1977.

[6] R. Moore, *Interval Analysis*, Prentice-Hall, Englewood Cliffs, N.J., 1966.

[7] D. Dubois, H. Prade,"On several representations of an uncertain body of evidence", in: *Fuzzy Information and Decision Processes* (M.M. Gupta, E. Sanchez, eds.), North-Holland, pp. 167-181, 1982.

[8] S. Moral, "Constructing of possibility distribution from a probability distribution", in : *Fuzzy sets theory and Applications*, A. Jones, A. Kauffmann, H.J. Zimmermann, Eds, pp. 51-60, 1986.

[9] D. Dubois, H. Prade, S. Sandri, "On possibility/probability transformations", in*: Fuzzy Logic*, R. Lowen, M. Roubens, Eds, pp. 103-112, 1993

[10] G. Mauris, V. Lasserre, L. Foulloy, "Fuzzy modeling of measurement data acquired from physical sensors", IEEE Trans. on Measurement and Instrumentation, Vol. 49, No. 6, pp. 1201-1205, 2000.

[11] G. Mauris, V. Lasserre, L. Foulloy, "A fuzzy approach for the expression of uncertainty in measurement", International Journal of Measurement, Vol. 29, No 3, pp. 165-177, 2001.

[12] M. Urbanski, J. Wasowki J., "Fuzzy approach to the theory of measurement inexactness", International Journal of Measurement, Vol. 34, pp. 67-74, 2003.

[13] A. Ferrero, S. Salicone, "A t-norm based fuzzy approach to the estimation of measurement uncertainty", IEEE Trans. On Instrumentation and Measurement, Vol. 53, No 5, pp. 1370-1377, 2004.

[14] C. De Capua, E. Romeo, "The random-fuzzy variables: a new approach to the expression of uncertainty", IEEE Technology Conference on Instrumentation and Measurement, Como, Italy, pp. 229-233, 2004.

[15] L.A. Zadeh, "Fuzzy sets as a basis for a theory of possibility, Fuzzy Sets and Systems", Vol.1, No1, pp. 3-28, 1978.

[16] R. Willink., "Coverage intervals and statistical coverage intervals", Metrologia, No 41, pp. I.5-I.6, 2004.

[17] D. Dubois, L. Foulloy, G. Mauris, H. Prade, "Probability-possibility transformations, triangular fuzzy sets and probabilistic inequalities", Reliable Computing, Vol. 10, No 4, pp. 273-297, 2004.

[18] M. Franck, B. Schweizer, "On the duality of generalized infimal and supremal convolutions", Rend. Mat. Ser., Vol. VI, No 12, pp. 1-23, 1979.

[19] R. Nelsen, *An introduction to Copulas*, Springer, New York, 1999.