Probabilistic Uncertainty Evaluation Aiding Critical Measurement-Based Decision

G. B. ROSSI, F. CRENNA
University of Genova- DIMEC
I-16145 Genova – Italy

ABSTRACT

Measurements are often an objective support for decision-making. If a measurement result can affect the decision, the measurement uncertainty associated with the result can be useful to manage the risk of a wrong decision. This is applicable to measurements in all fields, scientific, industrial or human-related. In the last case particular care has to be taken when a wrong decision can affect health, life or environment.
Now, measurement-based decisions may be, with fair generality, reduced to a conformity assessment problem. Typically the state of the process under investigation is monitored by a set of measured parameters and their belonging to a given "safe" subset of the parameter space has to be checked.
Decisions are thus strongly affected by measurement uncertainty. Common decision rules are based on expressing measurement uncertainty as intervals of values (expanded uncertainty), which results in an on-off criterion, with an uncertainty region. Due to the importance of the problem, a more sophisticated approach, based on probability, merits investigation, aiming at improving the quantification of the risk associated with each decision. In the article, such an approach is addressed in general terms. The role of uncertainty evaluation is then discussed and an algorithm for evaluating uncertainty in terms of probability density is presented and the advantages of this approach are discussed.

Keywords: measurement vs. decision support, uncertainty of measurement, uncertainty evaluation by probability densities, risk management

1. INTRODUCTION

In order to assess conformity of a parameter characterising some critical process with legal prescriptions or with technical requirements, measurement uncertainty strongly affects the decisional process [1-6].
In order to obtain an explicit evaluation of the risk it is necessary to express the measurement result as a probability density.
So in this paper, we first present a theoretical framework for addressing decisions based on uncertain measurements, in the case of requirements expressed as threshold, which are typical in environment and safety decisions. We will consider both the risk associated to a decision taken on the basis of a single measurement result, and the global risk associated to the continuos monitoring of a process.
Then we briefly address the way measurement results may be expressed by probability distributions.
In particular, after briefly reviewing the hypothesis assumed in the GUM for the evaluation of uncertainty, we present a procedure for expressing the measurement result in terms of a probability density, *under the same hypotheses*.
Algorithms for dealing with more complicated evaluation cases are also briefly mentioned [8].

2. INSIGHT IN MEASUREMENT-BASED DECISIONS

Let us consider the problem of assessing, by measurement, whether a positive-defined parameter (such as an amount of substance) belongs or not to a *safe region*, defined by a threshold condition:

$$x \le a \qquad (1)$$

Let us consider, for simplicity and without loss of generality, the case in which x may take only integer values.
Let $P_X(x)$ be the probability of x taking any specified value.
If we are *able to detect* any violation of condition (1), we may take appropriate actions, i. e. rejecting the item for which (1) is violated, stopping the process producing x, activate a proper feedback or simply taking note of the occurrence of the violation.
On the other hand, if we are unable to detect such a violation, we will suffer its consequences, whatever they may be.
So we will consider here the *risk* R of such an occurrence and *how measurement may reduce it*.
In the particular case *we make no measurement at all*, the (global) risk is simply:

$$R_0 = \text{probability that value of the parameter is larger than the threshold} = \Pr(x > a)$$

For instance, in the case of Fig. 1, the threshold is $a = 1{,}5$, and the global risk is $R_0 = \Pr(x > 1{,}5) = 0{,}1$.

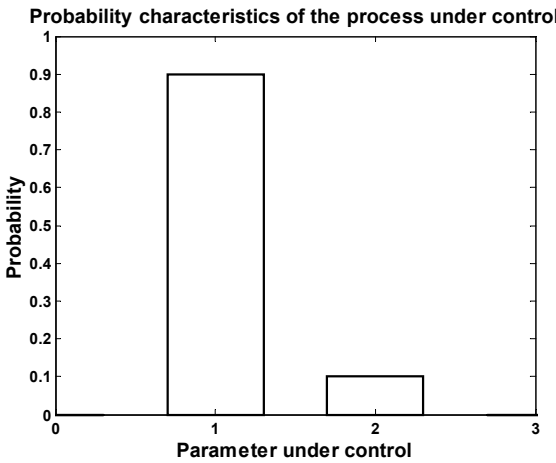

Fig. 1: Probability distribution of the process under control

Alternatively, let us consider the case we measure the parameter x, through a *measurement process* M, providing a (generic) measurement result $\hat{x}$.
Now we are able to implement rule (1), by substituting the (true) value of the parameter x, with the measured value $\hat{x}$, i. e. by actually applying the rule[1]:

$$\hat{x} \le a \tag{2}$$

Adopting (2), we surely reduce the global risk, but *we still have some risk*, since it may happen that $\hat{x} \neq x$ and so *it may happen that (1) is violated, while (2) is not*.
Let us suppose, in our example, that we get $\hat{x} = 1$: in this case we will consider to be in a safe region, but due to the uncertainty of the measurement process, we may actually be outside the safe region.
In this case the risk will be:

$$R_{M,\hat{x}=1} = P_M\left(x > a / \hat{x} = 1\right) \tag{3}$$

Such a risk obviously depends upon the quality of the measurement process M, and refers to the specific case in which we observe the value $\hat{x} = 1$: we will call this *the local risk*, i. e. the risk associated to the specific observation $\hat{x} = 1$, whilst R_0 was to be interpreted as the *global risk,* i. e. *the average risk associated to the management of process X*.

In our example, the event $x > a$, is equivalent to the event $x = 2$, then we have:

$$R_{M,\hat{x}=1} = P_M\left(x = 2 / \hat{x} = 1\right) \tag{4}$$

Now we may assume the measurement process is characterised by a conditional probability distribution $P_M\left(\hat{x} = \hat{x}_j / x = x_i\right)$. For instance, in our example, we may assume such a distribution only depending on the difference $\hat{x} - x$ and having the values presented in Fig. 2.

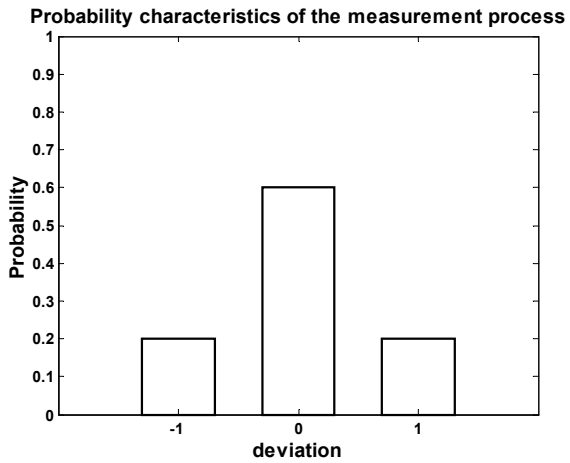

Fig. 2: Probability distribution of the deviation $\hat{x} - x$.

Now in order to evaluate our risk, we actually need the distribution $P_M\left(x / \hat{x}\right)$.
From the theorem of Bayes, we have:

$$P_M\left(x / \hat{x}\right) = \frac{P_M\left(\hat{x} / x\right) P_X\left(x\right)}{P_M\left(\hat{x}\right)} \tag{5}$$

The two involved conditional probability distributions are presented in Fig. 3.

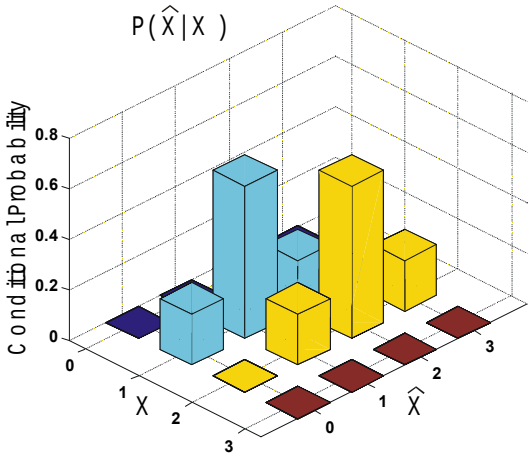

Fig. 3,a $P_M\left(\hat{x} / x\right)$

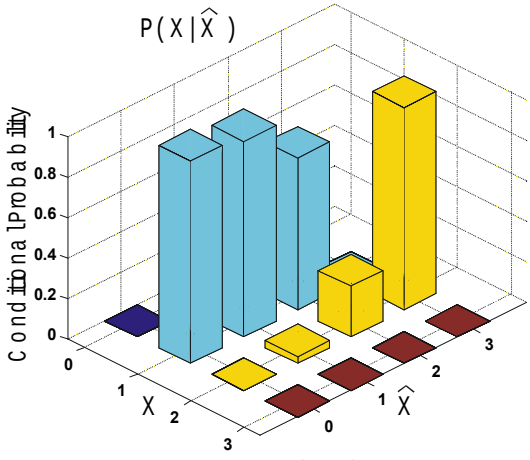

Fig 3,b $P_M\left(x / \hat{x}\right)$

[1] The alternative, more cautionary rule $\hat{x} \le b$, with $b < a$ will be discussed later on.

It is important to note that in this case we need to know the distribution of x, or anyway to suitably assume it, otherwise the evaluation of the risk is strongly influenced.

On the basis of $P_X(x)$ and $P_M(\hat{x}/x)$ we may evaluate $P_M(\hat{x})$, which in our example is presented in Fig. 4.

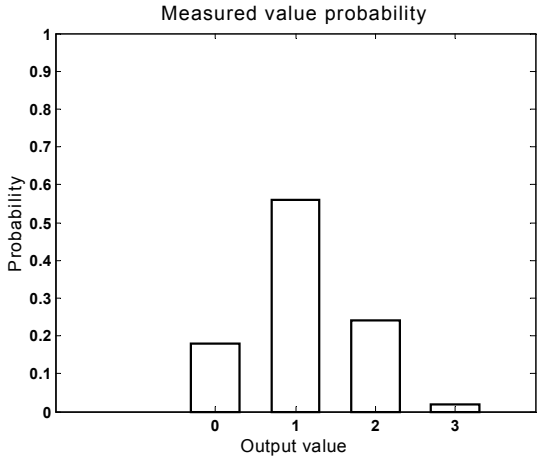

Fig. 4: Probability distribution of the measured values.

Finally the risk is:

$$R_{M,\hat{x}=1} = P_M(x = 2/\hat{x} = 1) = 0,036 \quad (6)$$

In the case we assume a uniform distribution for $\hat{x}$, we would have obtained:

$$R'_{M,\hat{x}=1} = P_M(x = 2/\hat{x} = 1) = 0,25 \quad (7)$$

which is a rough overestimation of the risk. Actually, the evaluation of $P_M(\hat{x})$ requires the evaluation of $P_X(x)$.
So the evaluation of the distribution of the process x is a necessary step in risk evaluation!

Now if we want to turn to the global risk, we have to consider all the cases in which (2) applies and we have make use of the theorem of total probability.
In our case we have:

$$\begin{aligned} R_M &= \\ &= P_M(x = 2/\hat{x} = 0)P_M(\hat{x} = 0) + \\ &+ P_M(x = 2/\hat{x} = 1)P_M(\hat{x} = 1) = \\ &= 0 \cdot 0,18 + 0,036 \cdot 0,56 = 0,02 \end{aligned} \quad (8)$$

So with measurement we have obtained a reduction of the global risk of 80% *!*
It may be noted that an evaluation of the risk performed through R'_M would have led to an evaluated global risk of 0,14, which means a risk greater than the risk associated to no measurement at all, which is obviously nonsense.
Actually the risk considered so far is the risk of not taking appropriate actions, in the case the parameter is outside the safe region. This kind of risk is sometimes called *consumer's risk*.
There is also the complementary risk of taking rescue actions when the parameter actually lays inside the safe region, but the measured value is not. This kind of risk, sometimes called *producer's risk*, is given by:

$$\begin{aligned} R^*_M &= \\ &= P_M(x \le a/\hat{x} = 2)P_M(\hat{x} = 2) + \\ &+ P_M(x \le a/\hat{x} = 3)P_M(\hat{x} = 3) \end{aligned} \quad (9)$$

which means:

$$\begin{aligned} R^*_M &= \\ &= P_M(x \le a/\hat{x} = 2)P_M(\hat{x} = 2) + \\ &+ P_M(x \le a/\hat{x} = 3)P_M(\hat{x} = 3) = \\ &= P_M(x = 1/\hat{x} = 2)P_M(\hat{x} = 2) + \\ &+ P_M(x = 1/\hat{x} = 3)P_M(\hat{x} = 3) = 0,18 \end{aligned} \quad (10)$$

If we now consider an alternative measurement system, M', more accurate, characterised by a probability distribution of the deviation $\hat{x} - x$, as in Fig. 5.

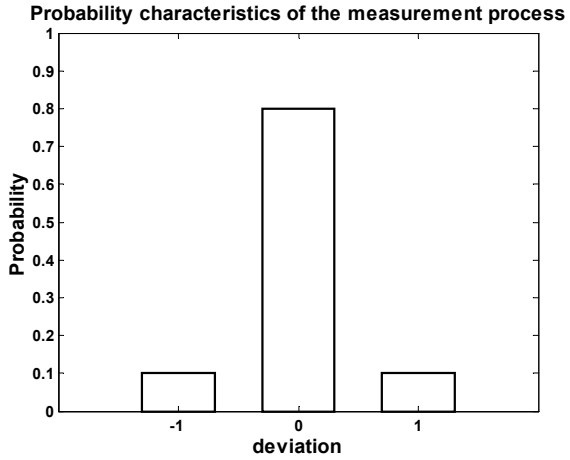

Fig. 5: Probability distribution of the deviation $\hat{x} - x$, for a more accurate measurement process

Performing the same calculation as above, we obtain a global (consumer's) risk:

$$R_{M'} = 0,01 \quad (11)$$

Moreover, for the producer's risk, we get:

$$R^*_{M'} = 0,09 \quad (12)$$

So *the reduction of uncertainty in the measurement process causes a reduction of both consumer's and producer's risk*, which may be evaluated in a quantitative fashion, and may be compared to the presumable increase in cost associated to a more accurate measurement process.

3. GENERAL FORMULAS

Let us now establish some general formulas.
We will consider both the discrete and the continuous case.
In the discrete case, let us assume we have a process X, characterised by a probability distribution $P_X(x)$, which

safe operation region is $x \le a$.
Let us consider a measurement process M, producing measurement values $\hat{x}$, characterised by a conditional probability distribution $P_M(\hat{x}/x)$.
In the examples above, we have considered a decision (acceptance) rule of the kind $\hat{x} \le a$. In more general terms, we may consider a decision rule $\hat{x} \le b$, where $b = \kappa a$, where $\kappa \le 1$ is a guard factor, i. e. we define an *acceptance region* which is included in the *safe region*.

Then, the *consumer's risk* for a specific measurement result $\hat{x}$, falling in the acceptance region, i. e. $\hat{x} \le b$, is:

$$R_{M,P_X}(\hat{x}) = \Pr(x > a/\hat{x}) = \sum_{x>a} P_M(x/\hat{x}) \quad \text{with } \hat{x} \le b \tag{13}$$

where:

$$P_M(x/\hat{x}) = \frac{P_M(\hat{x}/x)P_X(x)}{P_M(\hat{x})} = \frac{P_M(\hat{x}/x)P_X(x)}{\sum_x P_M(\hat{x}/x)P_X(x)} \tag{14}$$

It is apparent the role of the probability distribution of the process. Formulas above include the special case we have no information on such a distribution, which may be treated by assuming a uniform distribution over a suitable range, as shown in one of the examples above.

The related *producer's risk* may also be evaluated, in the case we get a value outside the acceptance region, i. e. $\hat{x} > b$. The risk is given by the probability of the value of the parameter actually falling within the safe region:

$$R^*_{M,P_X}(\hat{x}) = \Pr(x \le a/\hat{x}) = \sum_{x \le a} P_M(x/\hat{x}) \tag{15}$$

where again $P_M(x/\hat{x})$ is given by the formula above.
Let us now consider the *global consumer's risk*. This is given by the theorem of total probability:

$$R_{M,P_X} = \sum_{\hat{x} \le b} R_{M,P_X}(\hat{x})P_M(\hat{x}) = \sum_{\hat{x} \le b}\sum_{x>a} P_M(x/\hat{x})P_M(\hat{x}) \tag{16}$$

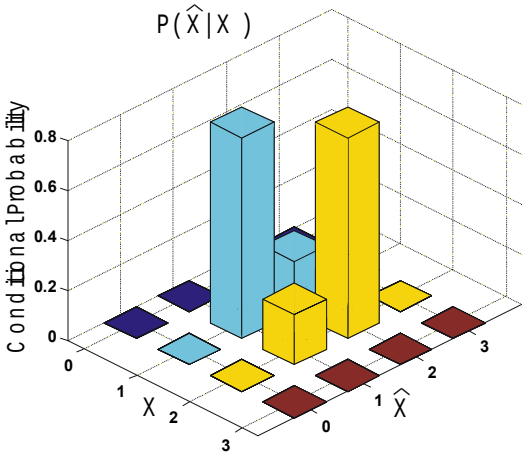

Fig. 6: Conditional probability distribution for the measurement process, in a yes-no case

Analogously, the *global producer's risk*, is given by:

$$R^*_{M,P_X} = \sum_{\hat{x}>b} R^*_{M,P_X}(\hat{x})P_M(\hat{x}) = \sum_{\hat{x}>b}\sum_{x \le a} P_M(x/\hat{x})P_M(\hat{x}) \tag{17}$$

On the other hand, if we consider the process X and the observation process M that are best represented in terms of continuous variables and related probability densities, we have simple to replace the probability distributions with their continuous counterpart. A summary of main useful formulas is presented in tables 1 and 2.

Kind of risk	*Consumer's risk*	*Producer's risk*
Specific risk	$R_{M,P_X}(\hat{x}) = \Pr(x > a/\hat{x}) = \sum_{x>a} P_M(x/\hat{x})$ with $\hat{x} \le b$	$R^*_{M,P_X}(\hat{x}) = \Pr(x \le a/\hat{x}) = \sum_{x \le a} P_M(x/\hat{x})$ for $\hat{x} > b$
Global risk	$R_{M,P_X} = \sum_{\hat{x} \le b} R_{M,P_X}(\hat{x})P_M(\hat{x}) = \sum_{\hat{x} \le b}\sum_{x>a} P_M(x/\hat{x})P_M(\hat{x})$	$R^*_{M,P_X} = \sum_{\hat{x}>b} R^*_{M,P_X}(\hat{x})P_M(\hat{x}) = \sum_{\hat{x}>b}\sum_{x \le a} P_M(x/\hat{x})P_M(\hat{x})$

Table 1 Summary of risk evaluation formulas for discrete processes.

Kind of risk	*Consumer's risk*	*Producer's risk*
Specific risk	$R_{M,p_X}(\hat{x}) = \Pr(x > a/\hat{x}) = \int_{x>a} p(x/\hat{x})dx$ with $\hat{x} \le b$	$R^*_{M,p_X}(\hat{x}) = \Pr(x \le a/\hat{x}) = \int_{x \le a} p(x/\hat{x})dx$ for $\hat{x} > b$
Global risk	$R_{M,p_X} = \int_{\hat{x} \le b} R_{M,p_X}(\hat{x})p_M(\hat{x})d\hat{x} = \int_{\hat{x} \le b}\int_{x>a} p(x/\hat{x})p_M(\hat{x})dxd\hat{x}$	$R^*_{M,p_X} = \int_{\hat{x}>b} R_{M,p_X}(\hat{x})p_M(\hat{x})d\hat{x} = \int_{\hat{x}>b}\int_{x \le a} p(x/\hat{x})p_M(\hat{x})dxd\hat{x}$

Table 2 Summary of risk evaluation formulas for continuous processes.

4. YES-NO DECISIONS

In the field of environmental analysis the importance of developing metrological criteria for measurements giving rise to a yes-no output has been recently recognised and underlined [9].
Now the examples we have presented above may be also considered as preparatory for the establishment of such criteria.
Let us for instance consider the same distribution for the process as presented in Fig. 1. Since in that distribution only two values have been considered, we may interpret them as characterising the two situation: process in the safe region vs. process outside the safe region.
In this case, for the measurement process, we may more properly consider a conditional probability distribution $P_M(\hat{x}/x)$ as in Fig. 6. This is perhaps a quite unusual distribution for measurement processes, and this may be

one of the reasons explaining why the field of yes/no decisions appears to be so troublesome.
Actually there is no special problem in obtaining such a distribution, for instance by, respectively lower and upper saturation of the probability distribution presented in Fig. 5. Now the risk evaluation formulas, for yes/no decision, are here presented.
It may be noted that now the difference between the specific and the global risk, is very subtle.
For the specific consumer's risk, we have:

$$R_{M,P_X}(\hat{x}) = \Pr(x > a/\hat{x} < a) = P_M(x = 2/\hat{x} = 1) = 0,027$$

While for the global risk, we have:

$$R_{M,P_X} = P_M(x = 2/\hat{x} = 1)P_M(\hat{x} = 1) = 0,02$$

5 EXPRESSING UNCERTAINTY IN TERMS OF PROBABILITY

5.1 The hypotheses

Till now we have assumed to express the final result of a measurement by a probability distribution over the possible values of the measurand.
This is of course a non-trivial task and the whole matter is, at present, object of investigation.
Anyway this task should not be overestimated. In this nse we present now a procedure which allows to assign a probability density to the result of measurement, on the basis of the same information required to evaluate measurement uncertainty according to the GUM.
Such hypotheses may be summarised as follows:

1. it is possible to identify the set of all quantities having a non negligible effect on the final measurement result (the estimate of the measurand), let denote it, using a slightly different notation, as $v = [v_1, v_2, ..., v_n]$;
2. it is possible to express the effect of such quantities on the final measurement result as a linear function of the variations in the quantities: $\delta x = \sum_{i=1}^{n} c_i \delta v_i$, with known coefficients c_i;
3. for each of the input quantities, the following information is available[2]:
   a) either an estimated standard deviation $\hat{\sigma}_i$, with ν_i degrees of freedom, which may be finite or infinite;
   b) or an estimate of the range of variability, say $[v_{0i} - \Delta v_i, v_{0i} + \Delta v_i]$, with limit either certain or subject to a relative uncertainty, say $\alpha_i = \Delta\Delta v_i / \Delta v_i$;
4. it is possible to identify a subset of mutually uncorrelated input quantities, say, without loss of generality, $v' = [v_1, v_2, ..., v_m]$, $0 \le m \le n$, for which the corresponding correlation matrix is known $R = [r_{ij}]$;
5. for the evaluation of expanded uncertainty according to the Welch Satterwaite method it is assumed that all input quantities are uncorrelated.

Let us briefly comment about these hypotheses.
Hypothesis one is quite natural. Indeed it may be considered as a minimum requirement for any uncertainty evaluation to take place. The second hypothesis is not so general since relationships are not necessarily linear, anyway some suitable linearisation may in many cases be performed.
Hypothesis three is again very general; hypothesis four is a natural consequence of hypothesis two.
Hypothesis five is clearly in contrast with hypothesis three and is a clear limitation of the method.
So hypothesis one trough four form together a set of near-minimum requirements for virtually any measurement task: so we will assume them in our procedure, while dropping hypothesis five.

5.2 The procedure

Actually, under the assumed hypotheses, $\hat{x}$ turns out to be a position parameter, so that $p(x/\hat{x}) = p_{\Delta v}(x - \hat{x})$, where $\hat{x}$ is the estimated value of the measurand, corrected for any known systematic effect. So we have actually to calculate only $p_{\Delta v}(x)$, for a suitable set of possible values x of the measurand.
In turn $p_{\Delta v}(x)$ may be expressed as the composition, through a convolution rule, of the density accounting for the global influence of the subset of uncorrelated input quantities (if not null), $p_{\Delta v,uncorr}(x)$, with the density accounting for the effect of complementary subset of mutually correlated input quantities (if not null), $p_{\Delta v,corr}(x)$. So we have:

$$p_{\Delta v}(x) = p_{\Delta v,uncorr}(x) * p_{\Delta v,corr}(x) = \int p_{\Delta v,uncorr}(x - \xi) p_{\Delta v,corr}(\xi) d\xi \qquad (18)$$

So we have to evaluate separately these two densities.
The density associated to the subset of non correlated input variables is again the convolution of the density associated to each variable describing its effect on the final result. Every such density should then account for available knowledge, which includes a dispersion figure (either a standard deviation or a variation range), a sensitivity coefficient and in some cases an index of uncertainty on the knowledge of the dispersion figure (either a limited number of degrees of freedom or a relative uncertainty on the width of the range of variability).
So we adopt the following rules, for calculated single densities:

1. If we known the standard deviation σ_i with infinite degrees of freedom, we assign a gaussian density, with standard deviation $|c_i|\sigma_i$.
2. If we known the standard deviation σ_i with a finite number of degrees of freedom ν_i, we assign a t-

[2] We do not consider here, for simplicity, the case of a triangular distribution, which may however be treated in essentially a similar way as case b.

Student density, with ν_i degrees of freedom, scaled by a factor $|c_i|\sigma_i$.

3. If we know the variation range Δv_i, with no uncertainty on its value, we assign a rectangular density over the scaled range $|c_i|\Delta v_i$.
4. If we know the variation range Δv_i, with a relative uncertainty α_i on its value, we assume a numerically evaluated density (not to be detailed here).

The above rules allow assigning a density accounting for the effect of each of the m variables belonging to the subset of mutually uncorrelated variables. By convolving these density we obtain $p_{\Delta \mathbf{v},\text{uncorr}}(x)$. We have now to calculate $p_{\Delta \mathbf{v},corr}(x)$.

For doing so we consider a gaussian density resulting from the sum of n-m variables, with correlation matrix equal to R, and standard deviations to be calculated with the following rules.

1. If we known the standard deviation σ_i with infinite degrees of freedom, the standard deviation will be $|c_i|\sigma_i$.
2. If we known the standard deviation σ_i with a finite number of degrees of freedom ν_i, the standard deviation will be $\nu_i |c_i|\sigma_i / (\nu_i - 2)$.
3. If we know the variation range Δv_i, with now uncertainty on its value, the standard deviation will be $|c_i|\Delta v_i / \sqrt{3}$.
4. If we know the variation range Δv_i, with a relative uncertainty α_i on its value, the standard deviation will be calculated numerically through the density as in Fig. 1.

In this way we may account at the same time, in a reasonable, although approximate, way of both the correlation between the variables an of the uncertainty on dispersion parameters.

5.3 The software package

A software package for the computation of the final probability density is under development. From the user point of view it is important to have a graphical interface which offers a guidance during the input of the parameters for the measurement process under investigation. Two main input windows are available for correlated and uncorrelated influence quantities. In both of them it is possible to select the proper shape for the probability distribution, among a set of available possibilities, to establish a value for the dispersion figure together with its uncertainty expressed either as a number of degrees of freedom or as a relative uncertainty on the distribution parameter. If there are correlated variables, the correlation parameters are requested. During the computation, preliminary results are presented, such as the distributions for the influence quantities, obtained considering the uncertainty on their parameter. Final results are presented graphically by plotting the final probability distribution, together with the indication of standard and expanded uncertainty. Standard and expanded uncertainties evaluated according to GUM are also available for reference.

6. CONCLUSIONS

In the paper the risk reduction which may be achieved by measuring critical parameters in a potentially risky process has been discussed. General formulas for the evaluation of both local risk, associated to a single measurement result, and global risk, related to the continuous monitoring of a process, have been derived and illustrated by examples. The important case of yes-no decisions has also been considered. Since evaluation of the risk requires the measurement result to be expressed in terms of probability distributions, a procedure for doing so has been presented. The procedure assumes the same kind of information required by the GUM, so it is of wide application.

7. ACKNOWLEDGEMENTS

The present study is a part of the activities that our laboratory carries on within the Thematic Network METROPOLIS: Metrology in Support to Precautionary Sciences and Sustainable Development Policies, EU funded within Framework Program 5.
The research is also partially funded by The Italian Ministery of Education and Research (MIUR-PRIN 2001)

8 . REFERENCES

[1] G. B. Rossi, F. Crenna, Probabilistic evaluation of uncertainty as a support to measurement-based decisions, European Conference "Environment, health, safety, a challenge for measurement, Paris, June, 2001
[2] Guide to the Expression of Uncertainty in Measurement, ISO, Geneva, 1995
[3] Geometrical Product Specification (GPS),Part 1: Decision rules for proving conformance or non-conformance with specification, ISO 14253 – 1: 1998.
[4] CENELEC- Draft prEN 50222 (1997) Standard for the evaluation of measurement results taking measurement uncertainty into account
[5] IEC CISPR/A/204/CD (1997) Accounting for measurement uncertainty when determining compliance with a limit.
[6] K. -D. Sommer, M. Kochsiek, W. Schulz, Error limits and measurement uncertainty in legal metrology, Proc. of the XVI IMEKO World Congress, Vienna, 2000
[7] I. Lira, A bayesian approach to the consumer's and producer's risk in measurement, Metrologia, 36, 397-402, 1999
[8] G. B. Rossi, The potentials of probability density for the evaluation of measurement uncertainty, IMEKO TC8 Int. Symposium, Torino, 1998
[9] M. Valcarcel, S. Cardenas: Metrology of the binary yes/no response in environmental Analysis, European Conference "Environment, health, safety, a challenge for measurement, Paris, June, 2001