XVI IMEKO World Congress
Measurement - Supports Science - Improves Technology - Protects Environment ... and Provides Employment - Now and in the Future
Vienna, AUSTRIA, 2000, September 25-28

A COMPUTATIONAL METHOD IN THE CONSTRUCTION OF A PROBABILITY REGION OF CONFIDENCE

G. Iuculano [1]***, A. Zanobini*** [1] ***and G. Pellegrini*** [2]
[1] Dipartimento di Ingegneria Elettronica, Facoltà di Ingegneria and
[2] Dipartimento di Matematica Applicata, Facoltà di Ingegneria
Università di Firenze, Via di S. Marta 3 50139 Firenze, Italy

Abstract: The construction of a region of confidence or "confidence belt", that is the expression of an expanded uncertainty with a given coverage probability, for measurements performed on the same measurand by different processes, is a critical task in the perspective to establish a measurement traceability on a global basis. An alternative procedure to the analytical one suggested in the Guide to the Expression of Uncertainty in Measurement [1], strictly related to the validity of the Central Limit Theorem, is desirable, in order to cover a larger range of situations not included in the normal distribution cases. In this work statistical approach is proposed based on the use of Monte Carlo approximation technique and on the bootstrap resampling iteration. An experimental model is examined to check the validity of the proposed method.
Keywords: Confidence Belt, Monte Carlo-Bootstrap Approximation

1 INTRODUCTION

A measurement result is complete only when accompanied by a quantitative statement of its uncertainty. A measurement is of little use unless there is some way of estimating the associated uncertainty, so an experimental result should always be accompanied by such an estimate. This policy requires that a uniform approach to expressing measurement uncertainty be followed by the metrological community.

Many researches have been developed on the subject in the recent years to ensure that the quantitative statements of uncertainty produced by different methodologies are consistent with each other and agree with those recommended by the Major International Committee for Weights and Measures. A contribution has been provided by the so called ISO Guide [1], as a result of an international effort to coordinate the whole subject within a unifying framework.

The Guide defines uncertainty (of measurement) as: "a parameter , associated with the result of a measurement, that characterises the dispersion of the values that could reasonably be attributed to the measurand" and from the operative point of view distinguishes two explicitly parameters:

The standard uncertainty of the result of a measurement, σ: that corresponds to the standard deviation associated with the result.

The expanded uncertainty, u: corresponding to a confidence interval obtained by multiplying the combined standard uncertainty by a coverage factor k. The intended purpose of $u = \pm k\sigma$ is to provide an interval about the result of a measurement that may be expected to encompass a large fraction of the distribution of values that could reasonably be attributed to the measurand.

The method that implicitly the Guide suggests, for evaluating uncertainty, is essentially limited to the so called Gaussian "propagation law" for uncertainty, while the calculation of expanded uncertainty is based on an approximation method.

It is reasonable to look for generalising methods, robust with respect different situations no strictly related to the validity of the Central Limit Theorem and including multivariate cases no mentioned in the Guide.

The main purpose of this research is the estimation of the probability region or "confidence belt", as an expression of the expanded uncertainty with a given coverage probability, by using Monte Carlo technique and bootstrap resampling iterations.

2 BASIC DEFINITIONS

A multivariate n-dimensional "measure" is defined as a random vector: $\underline{M} = (M_1,...,M_n)$, whose components $M_1,...,M_n$ could represent the results performed on the same measurand by different processes, or to be the results of different measurands inside the same measurement process.

The variability of $\underline{M}$ may be summarised into a probability region say $C^{(n)} \subset R^n$, with an assigned coverage probability equal to p that is:

$$P\{\underline{M} \in C^{(n)}\} = p \tag{2.1}$$

Introducing the n-th dimensional joint probability density: $f_{\underline{M}^T}(\underline{m})$ with $\underline{m} = (m_1, m_2, ... m_n)$, it is possible to evaluate the (2.1) by the following relation:

$$P\{\underline{M} \in C^n\} = \int_{C^n} \cdot \int f_{\underline{M}}(\underline{m}) \, d\, m_1 \cdot \cdot d\, m_n = p \tag{2.2}$$

with

$$f_{\underline{M}}(\underline{m}) = \lim_{\underline{h} \to \underline{0}} \frac{P\{m_1 < M_1 \le m_1 + h_1, ..., m_n < M_n \le m_n + h_n\}}{h_1 .....h_n} = \frac{\partial^n F(\underline{m})}{\partial m_1 \partial m_2 ... \partial m_n} \tag{2.3}$$

where $\underline{h} = (h_1, h_2, ... h_n)$, $\underline{0}$ is the zero vector and

$$F(\underline{m}) = \int_{-\infty}^{m_1} \int_{-\infty}^{m_2} ... \int_{-\infty}^{m_n} f_{\underline{M}}(x_1, ... x_n) dx_1 ... dx_n$$

is the probability distribution function.

In the hypothesis of normal distribution the n-th dimensional joint probability density assumes the form:

$$f_{\underline{M}}(\underline{m}) = (2\pi)^{-\frac{n}{2}} \left(det\, \underline{\underline{D}}\right)^{-\frac{1}{2}} exp\left[-\frac{1}{2}(\underline{m} - \underline{\mu})\underline{\underline{D}}^{-1}(\underline{m} - \underline{\mu})^T\right] \tag{2.4}$$

with $\underline{\mu} = E\{\underline{M}\} = (E\{M_1\}, ..., E\{M_n\})$ and $\underline{\underline{D}} = E\{(\underline{M} - \underline{\mu})^T (\underline{M} - \underline{\mu})\}$ is the $n \times n$ dispersion matrix.

The symmetric property of the dispersion matrix guaranties the existence of a non singular matrix $\underline{\underline{Q}}$ such that

$$\underline{\underline{Q}}^T \underline{\underline{D}}\, \underline{\underline{Q}} = \underline{\underline{\ddot{E}}} \tag{2.5}$$

where $\underline{\underline{\ddot{E}}}$ is the diagonal matrix with the nonull elements equal to the eigenvalues of the matrix $\underline{\underline{D}}$.

Introducing the linear transformation

$$L(\underline{M}) = \underline{W} \tag{2.6}$$

defined by the relation:

$$\underline{W} = (\underline{M} - \underline{\mu})\underline{\underline{Q}} \tag{2.7}$$

the (2.2) taking into account the (2.4), assumes the following form:

$$P\{\underline{M} \in C^n\} = P\{\underline{W} \in B^n\} = \int_{B^n} \cdot \int (2\pi)^{-\frac{n}{2}} \left(det\, \underline{\underline{D}}\right)^{\frac{1}{2}} exp\left[-\frac{1}{2}\underline{W}\underline{\underline{\ddot{E}}}\, \underline{W}^T\right] \left|det\left(\underline{\underline{Q}}\right)\right| d\, w_1 \cdot \cdot d\, w_n = p \tag{2.8}$$

where $B^n = L(C^n)$ is the transformed domain according to (2.6) and (2.7).

3 THE MONTE CARLO METHODS

The class of algorithms that solve problems probabilistically are known by the name of Monte Carlo methods (MCM). At the present MCM are the most competitive approaches in many computational problems when the numerical analytic algorithms become too heavy as in the multiple integral computation, or in the optimization and functional approximation in multidimensional spaces.

Here the purpose is to look at using Monte Carlo methods to evaluate the confidence region in a measurement process.

Let's briefly outline the basic idea associated with these methods. The numerical value of a given entity J can be evaluated as described in the following steps:

A random variable X with mean value J, such that $E(X) = J$, is chosen, in order to provide an unbiased estimation of J. The variable X is supposed to assume the subsequently values $X_1, X_2, ..., X_n$ in a series of repeated trials.

The sample mean: $Y_N = \frac{1}{N}\sum_{i=1}^{N} X_i$, whose value will probably be very close to J, according to the law of large numbers, provides another unbiased estimator of J and will follow a normal distribution when N tends to infinity. So that, as soon as N is large enough, by the central limit theorem, $E(Y_N) = J$.

The variance can be estimated by the following relations:

$$V(X) = s^2 \cong \frac{\sum_{i=1}^{N}(X_i - Y_N)^2}{N-1} \tag{3.1}$$

and thereafter

$$V(Y_N) = \frac{s^2}{N} \tag{3.2}$$

A statistically conditioned error estimate can be written as follows:

$$\text{Prob}\left(\left|Y_N - J\right| \le \frac{ts}{\sqrt{N}}\right) = f(t) \tag{3.3}$$

where $f(\tau)$ is increasingly monotone, goes to 1 for $\tau \to \infty$. The error estimate is controlled by the desired confidence level by varying τ.

In particular Monte Carlo integration is easily understood by considering the following n-dimensional integral:

$$J = \int_{Q^n} \cdot \int f(x_1,...x_n)\,dx_1 \cdot\cdot dx_n \quad \text{where} \quad Q^n = (a_1,b_1)\times(a_2,b_2)\times....\times(a_n,b_n) \tag{3.4}$$

By defining a function $g(\underline{x}) = g(x_1,...,x_n)$ as:

$$g(\underline{x}) = \begin{cases} 1 & \text{if } \underline{x} \in Q^n \\ 0 & \text{otherwise} \end{cases}$$

the integral (1.4) can be written:

$$J = \frac{1}{V}\int_{Q^n} \cdot \int f(x_1,...x_n)\,g(x_1,...x_n)\,V dx_1 \cdot\cdot dx_n \quad \text{with} \quad V = (b_1 - a_1)(b_2 - a_2)\cdot\cdot(b_n - a_n) \tag{3.5}$$

The (1.5) can be interpreted as the expectation of the function $h(\underline{x}) = f(\underline{x})g(\underline{x})V$ for the n-dimensional random variable $\underline{x} = (x_1,...,x_n)$ which is uniformly distributed within the domain Q^n .

This then gives an approximate procedure:

$$J \approx \frac{1}{N}\sum_{i=1}^{N} h(\underline{x}_i) = \frac{V}{N}\sum_{i=1}^{N} f(\underline{x}_i) \tag{3.6}$$

The quantity j may be regarded as a random variable whose standard deviation decreases as $N^{-1/2}$ independently of the dimension n of the integral, other numerical integration methods have errors decreasing as $N^{-1/2}$. This property make the Monte Carlo method more efficient from computational point of view as the dimension n increases.

4 BOOTSTRAP METHOD

The bootstrap, as a computational device, was introduced by Efron (1979) as a quite intuitive and simple way of finding approximations of quantities that are very hard, or even impossible to compute analytically. The bootstrap is a method for estimating the distribution of an estimator or test statistic by resampling one's data.

The bootstrap is often more accurate in finite samples than first order asymptotic approximations, thus, it can provide a practical method for improving upon first order approximation results, reducing or eliminating finite-sample distortions of the levels of statistical tests.

The bootstrap has been the object of much research in statistics since its introduction. The results of this research are synthesised in the books by Efron and Tibshirani (1993), Hall (1992) and Mammen (1992).

The purpose of this paper is to test the usefulness of the bootstrap for improving upon first order asymptotic approximations in contexts of interest in metrology to uncertainty evaluation.

The basic idea in bootstrap algorithms is remarkably simple.

Given the n sample values x_1, ... , x_n, an empirical frequency function $\hat{f}$ which has values 1/n at each x_i and zero elsewhere may be defined. The bootstrap estimate of the sampling variance of a statistic is obtained by computing a sample of size n with replacement from the population described by $\hat{f}$ and computing the estimate $\boldsymbol{Q}^*$ from this sample; this can be repeated M times, giving the set of bootstrap samples: $\boldsymbol{Q}_j^*$ with j=1, ... , M.

The bootstrap estimator is $q^* = \frac{1}{M}\sum_{j=1}^{M} q_j^*$ and the corresponding estimator of the sampling variance is $\hat{V}_B(\theta^*) = \frac{1}{M-1}\sum_{j=1}^{M}(\theta_j^* - \theta^*)^2$. The convergence of θ^* and $\hat{V}_B(\theta^*)$ to θ and to the sampling variance $V(\theta^*)$, respectively, as n and M tend to infinity is the justification for this approach and in practice $M \geq 50$ appears to be suffice.

One of the most important applications of the bootstrap methods is the confidence intervals estimation especially in presence of asymmetric distributions: Davison and Hinkley (1997), Efron and Tibshirani (1993), Hall(1992), Mooney and Duval (1993), Shao and Tu (1995).

5 APPLICATIONS AND EXAMPLE

Only in recent years computer-intensive techniques as Monte Carlo and Bootstrap methods has been proposed for metrological applications, for a review on the subject see [8].

A theoretical tri-dimensional model with a normal distribution concerning the estimation of a confidence region in a measurement process is examined. An analytical expression relating the characteristic parameters of the model and the level of confidence of the correspondent probability region may be gained by the usual techniques of multidimensional integral calculus and vector analysis.

An experimental situation is examined and computational results by the Monte Carlo and Bootstrap techniques are presented and compared with the analytical solution to test the validity of the proposed methods.

5.1 A tri-dimensional model

The three random variables M_i i=1, 2, 3 are supposed to have normal distribution M_i=N (μ_i, $\mathbf{s}_i^2$), where μ_i and $\mathbf{s}_i^2$ is the expectation and the variance respectively of M_i. The standard deviation $\sigma_i = \sqrt{\mathrm{Var}\{M_i\}}$ $i = 1,2,3$ can be used to evaluate the expanded uncertainty, u_i of the measure M_i as specified in (GUM) [1], being $\mu_i \pm u_i = \mu_i \pm k_i \mathbf{s}_i$ i=1, 2, 3 where k_i is a coverage factor.

The correlation coefficients are:

$$r_{ij} = corr(M_i M_j) = \frac{E\{M_i M_j\} - m_i m_j}{s_i s_j} \quad i, j = 1,2,3 \text{ with } |\rho_{ij}| \leq 1$$

The variability of the vector $\underline{M} = (M_1, M_2, M_3)^T$ may be summarised into a probability region say $C^{(3)} \subset R^3$, with an assigned coverage probability equal to p that is:

$$P\{\underline{M} \in C^{(3)}\} = \iiint_{C^3} f_{\underline{M}}(\underline{m}) d\,m_1\, d\,m_2\, d\,m_3 = p$$

with $f_{\underline{M}}(\underline{m}) = (2p)^{-\frac{3}{2}} (\det \underline{\underline{D}})^{-\frac{1}{2}} \exp\left[-\frac{1}{2}(\underline{m} - \underline{m})\underline{\underline{D}}^{-1}(\underline{m} - \underline{m})^T\right]$ and

$$\underline{\underline{D}} = \begin{bmatrix} s_1^2 & r_{12}s_1 s_2 & r_{13}s_1 s_3 \\ r_{12}s_1 s_2 & s_2^2 & r_{23}s_2 s_3 \\ r_{13}s_1 s_3 & r_{23}s_2 s_3 & s_3^2 \end{bmatrix}$$

according to (2.2) and (2.4).

Assuming $C^3 = I_1 x I_2 x I_3$ where $I_i = [\mu_i \pm k_i \mathbf{s}_i]$ i=1, 2, 3 and following the procedure described in the section 2, it is possible to obtain an analytical relation in terms of the characteristic parameters of the model and the level of confidece p of the region C^3 as follows:

$$(\det \underline{\underline{D}})^{-1/2} \left|\det \underline{\underline{Q}}\right| \prod_{i=1}^{3} \frac{1}{\sqrt{l_i}} erf\left(\sqrt{\frac{l_i}{2}}\, k_i^*\right) = p \qquad (5.1)$$

where l_i $i = 1,2,3$ are the eigenvalues of the matrix $\underline{\underline{D}}$, and $\underline{\underline{Q}} = (q_{ij})$ $i, j = 1,2,3$ is the matrix satisfying the (2.5) associated to the linear transformation (2.5), while $k_i^* = \sum_{j=1}^{3} q_{ij} k_j$ $i = 1,2,3$ are the transformed coverage factors, erf(-) is the error function.

An experimental model is examined whose characteristic parameters are reported in table (1), the correlations r_{ij} $\mathrm{i, j} = 1{,}2{,}3$ following the relation: $1 + 2 r_{12} r_{13} r_{23} \geq r_{12}^2 r_{13}^2 r_{23}^2$

The level of confidence is evaluated for different values of the coverage factors assuming that $k_i^* = k^*$ $\mathrm{i} = 1{,}2{,}3$.

The computational results, based on the Monte Carlo and Bootstrap technique application, are compared with the analytical ones as reported in table (2):

Table 1. characteristic parameter of the model

ì $_1 = 4{,}999\ 57$ V	ì $_2 = 19{,}661\ 01$ mA	ì $_3 = 1{,}044\,46$ rad
$\sigma_1 = 0{,}003\,2$ V	$\sigma_2 = 0{,}009\,5$ mA	$\sigma_3 = 0{,}0007\,5$ rad
$\rho_{32} = -0{,}36$	$\rho_{33} = 0{,}86$	$\rho_{23} = -0{,}65$

Table 2. p_a = analytical value of the level of confidence
p_c = Bootstrap-Montecarlo numerical value of the level of confidence

k^*	p_a	p_c
0.002	0.9110	0.9090
0.003	0.9892	0.9801
0.004	0.9993	0.9981

6 COMMENTS AND CONCLUSIONS

The present research has to be considered as a first approach to understand the applicability of the numerical thecniques and to ceck the potentiality of the proposed method. The Monte Carlo numerical approximation has been pointed out with N=5x10^3. A "population" of size n=10^5 has been generated and the bootstrap resampling estimation has been calculated on M=5x10^4 samples by drawing at random n=100 values from the entire population.

The results summarised in Table 2 are appreciable enough to suppose that the method could deserve further attention inside more complicated models.

REFERENCES

[1] *Guide to the Expression of Uncertainty in Measurement*, first edition, (1993), corrected and reprinted (1995), International Organisation for Standardisation (Geneva, Switzerland).

[2] *International Standard ISO 3534-1 Statistics*-Vocabulary and Symbols-Part I: Probability and General Statistical Terms, first edition, (1993), International Organisation for Standardisation (Geneva, Switzerland).

[3] Davison A.C., Hinkley D.V. (1997) *Bootstrap Methods and their Application,* Cambridge, Cambridge University Press.

[4] Efron B.D., Tibshirani R.J. (1993) *An introduction to the Bootstrap* , New York, Chapman & Hall.

[5] Hall P. (1992) *The Bootstrap and Edgeworth Expansion,* New York, Springer-Verlag.

[6] Mooney C.Z., Duval R.D. (1993) *Bootstrapping. A Nonparametric Approach to Statistical Inference,* Newbury Park, CA, Sage Publications.

[7] Shao J., Tu D. (1995) *The Jacknife and the Bootstrap* , New York, Springer-Verlag.

[8] P.Ciarlini, G.Regoliosi and F.Pavese, *Non parametric bootstrap with application to metrological data,* Series on advances in mathematics for applied sciences, Vol.16, 1994.

[9] G. Iuculano, S. D'Emilio, G. Pellegrini "*Uncertainty Interpretation by Defining a Probability Interval with a High Level of Confidence*" IMEKO Technical Committee 8 "Traceability in Metrology" Torino 10-11-Settembre 1998.

[10] G. Iuculano, E. Arri, G. Pellegrini "*Measurement Compatibility: A novel Approach Based on the Definition of Probability Intervals*" IMTC/99 ,Venice (Italy), May 1999.

[11] G. Iuculano, S. D'Emilio, G. Pellegrini "*Using Probability Interval For Measurement Uncertainty Evaluation*" IMEKO-XV ,Osaka (Japan), June 1999.

AUTHORS: G. IUCULANO, A. ZANOBINI and, Dipartimento di Ingegneria Elettronica, Facoltà di Ingegneria, Università di Firenze, Via di S. Marta 3 50139 Firenze, Italy and G. PELLEGRINI, Dipartimento di Matematica Applicata, Facoltà di Ingegneria, Università di Firenze, Via di S. Marta 3 50139 Firenze, ItalyE-Mail: Iuculano@ingfi1.ing.unifi.it