Metrol. Meas. Syst., Vol. XVII (2010), No. 2, pp. 195-204

PAN POLSKA AKADEMIA NAUK

METROLOGY AND MEASUREMENT SYSTEMS

Index 330930, ISSN 0860-8229

www.metrology.pg.gda.pl

UNCERTAINTY INTERVAL EVALUATION USING THE CHI-SQUARE AND FISHER DISTRIBUTIONS IN THE MEASUREMENT PROCESS

Marcantonio Catelani, Andrea Zanobini, Lorenzo Ciani

University of Florence, Department of Electronics and Telecommunications, Via di Santa Marta 3, 50139, Florence, Italy
(✉ marcantonio.catelani@unifi.it, +39 55 479 6377, *andrea.zanobini@unifi.it, lorenzo.ciani@unifi.it)*

Abstract

Referring to the Guide to the Expression of Uncertainty in Measurement (GUM), the paper proposes a theoretical contribution to assess the uncertainty interval, with relative confidence level, in the case of n successive observations. The approach is based on the Chi-square and Fisher distributions and the validity is proved by a numerical example. For a more detailed study of the uncertainty evaluation, a model for the process variability has been also developed.

Keywords: measurement uncertainty, confidence level, statistical distributions.

1. Introduction

As it is known, a random variable M characterizing the measurement process can be associated with a measurement interval and, consequently, with the quality of results, therefore the measure. We introduce the confidence level that can be attributed to the occurrence of each single event associated with the variable M in the space of all possible measurement results $S = \{ m_{\min} \le M \le m_{\max} \}$.

So, it is possible to assign the highest confidence level, equal to one by convention, when we have the certainty that M belongs to S; vice versa, the confidence level is minimum, equal to zero by convention, when the values of M do not belong to S.

Considering a subinterval $[m_a, m_b]$ of S, it is possible to assign a probability to the confidence level associated with the occurrence of M in $[m_a, m_b]$.

From these assumptions, the random variable M is characterized by a probability distribution, that is a function of random events that represent the probability that the measurement belongs to one of the possible subintervals of S. The probability distribution associated with M is all that is known in the measurement interval.

According to the GUM [1, 2] we introduce:

$$P\{|M - E\{M\}| \le k\, u_M\} = P\{E\{M\} - k\, u_M \le M \le E\{M\} + k\, u_M\} = p. \tag{1}$$

Eq. (1) represents the probability that the measure M is between its expected value $E\{M\}$ plus or minus a quantity given by the product of the standard uncertainty u_M and the coverage factor k. The parameter p, denoted as confidence level, should tend to one to have a high value of the occurrence of an event.

The interval:

$$E\{M\} - k\, u_M \le M \le E\{M\} + k\, u_M \tag{2}$$

Article history: received on Dec. 23, 2009; received in revised form on May 26, 2010; accepted on Jun. 7, 2010; available online on Jun. 00, 2010.

represents the confidence interval and it can be interpreted as that interval able to guarantee a high probability that it contains a large number of possible values of *M*. Hence a rise of the value of *p* leads to an increase of the number of events in which *M* is within the interval.

If the probability density function $f_M(m)$ of *M* is known, it is possible to evaluate the confidence level by means of the following expression:

$$p = \int_{E\{M\}-k\,u_M}^{E\{M\}+k\,u_M} f_M(m)\,dm. \tag{3}$$

It is now possible to indicate, explicitly, the measurement result as "uncertainty interval" associated with a measurand with an assigned confidence level *p*.

So, if we suppose to know the probability density, its distribution function $F_M(m)$ is also known, given by its integral. Therefore the uncertainty interval with confidence level *p* is defined by the equation:

$$P\left\{m_\alpha \le M \le m_{p+\alpha}\right\} = \int_{m_\alpha}^{m_{p+\alpha}} f_M(m)\,dm = F_M\left(m_{p+\alpha}\right) - F_M\left(m_\alpha\right) = p, \tag{4}$$

where α is an appropriate value in the range [0, 1]. The extremes of the interval within which *M* is enclosed takes the name of quantiles of the distribution function F_M, and we have the following relationship:

$$F_M\left(m_\alpha\right) = P\left\{M \le m_\alpha\right\} = \alpha. \tag{5}$$

2. Application of the Chi-square and Fisher distribution to the estimation of the uncertainty interval

As introduced in [3], taking again into account *n* independent successive observations ($o_1, o_2, \cdots, o_n$) and assuming each observation as a normally distributed random variable with expected value m_o and standard uncertainty u_o, the chi-square distribution with (*n-1*) degrees of freedom can be represented by:

$$\chi^2_{n-1} = \frac{\sum_{i=1}^{n}\left(o_i - \bar{o}\right)^2}{u_0^2} \tag{6}$$

being the mean of such variables $\bar{o} = \frac{\sum_{i=1}^{n} o_i}{n}$ also normally distributed with mean value m_o and reduced variance $\frac{u_0^2}{n}$ [4].

The uncertainty interval can be introduced by considering the Chi-square probability distribution with associated ν degrees of freedom. With the pre-arranged confidence level *p*, this interval is defined as:

$$P\left\{\chi^2_\alpha \le \chi^2_\nu \le \chi^2_{p+\alpha}\right\} = F_\nu\left(\chi^2_{p+\alpha}\right) - F_\nu\left(\chi^2_\alpha\right) = p, \tag{7}$$

where α is a value in the range from zero to (1–*p*); the extremes of the interval χ^2_α and $\chi^2_{p+\alpha}$ are, respectively, the α- and ($p+\alpha$)-quantiles of the distribution function of χ^2_ν, whose cumulative distribution is given by:

$$F_\nu(m) = P\{\chi^2_\nu \le m\} = \int_0^m f_\nu(z)\,dz, \tag{8}$$

where: $f_\nu(z) = \dfrac{z^{\frac{\nu}{2}-1}}{2^{\frac{\nu}{2}}\Gamma\left(\frac{\nu}{2}\right)} e^{-\left(\frac{z}{2}\right)}, \quad 0 \le z < +\infty.$

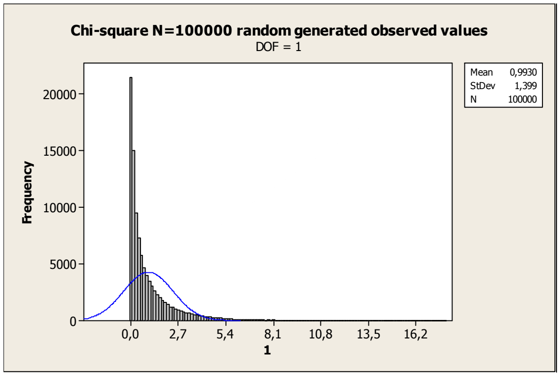
A β-quantile is an m_β value so that $F_\nu(m_\beta) = \beta$. Such quantiles are tabulated for different values of degrees of freedom ν corresponding to the respective β but they can be obtained more efficiently by means of specific statistic software.

Table 1 summarizes the results concerning the amplitude of the uncertainty interval with α = 0.025 ÷ 0.005 and ν = 1 ÷ 100 according to Eq. (7). Consequently, four histograms of 10^5 random generated observed values for different degrees of freedom fitted with Gaussian distribution can be obtained, as shown in Fig. 1. For each case, the mean and the standard deviation are also computed [3, 6].

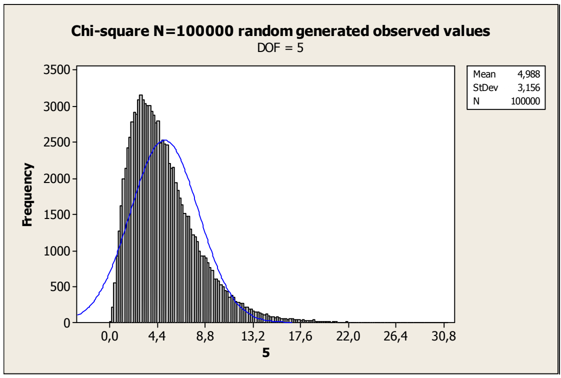
Table 1. Uncertainty interval amplitude in function of ν and α.

ν	$\chi^2_{0.025}$	$\chi^2_{0.975}$	$\chi^2_{0.005}$	$\chi^2_{0.995}$
1	0.000982	5.024	0.0000393	7.879
2	0.0586	7.378	0.01	10.597
5	0.831	12.832	0.412	16.750
10	3.247	20.483	2.156	25.188
20	9.951	34.170	7.434	39.997
50	32.357	71.420	27.991	79.490
100	74.222	129.561	47.328	140.169

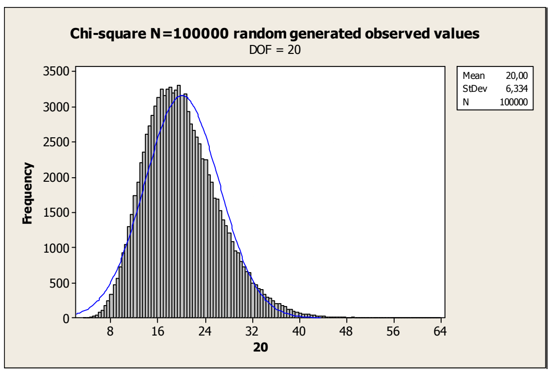
a) DOF = 1



b) DOF = 5



c) DOF = 20



d) DOF = 100

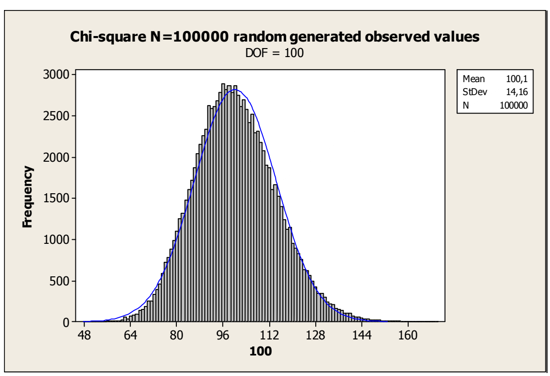


Fig. 1. Four histograms of 10^5 random generated observed values for different degrees of freedom (DOF) fitted with Gaussian distribution. Mean and standard deviation is also computed in each case.

The ratio of two independent chi-square variables, each divided by its respective degrees of freedom, is a random variable F_{ν_1,ν_2} defined as follows:

$$F(\nu_1,\nu_2)=\frac{\chi^2_{\nu_1}/\nu_1}{\chi^2_{\nu_2}/\nu_2}\ , \qquad 0\le F_{\nu_1,\nu_2}<+\infty. \tag{9}$$

The probability density function of F_{ν_1,ν_2} can be represented by:

$$f(m;\nu_1,\nu_2)=\frac{\Gamma\left[(\nu_1+\nu_2)/2\right]}{\Gamma(\nu_1/2)\Gamma(\nu_2/2)}\nu_1^{\nu_1/2}\nu_2^{\nu_2/2}\frac{m^{(\nu_1/2)-1}}{(\nu_1 m+\nu_2)^{(\nu_1+\nu_2)/2}}. \tag{10}$$

It can be observed that the distribution is asymmetric and, in this case, the β-quantiles $m_\beta(\nu_1,\nu_2)$ defined as:

$$P\{F(\nu_1,\nu_2)\le m_\beta(\nu_1,\nu_2)\}=\int_0^{m_\beta(\nu_1,\nu_2)} f(m,\nu_1,\nu_2)\,dm=\beta. \tag{11}$$

This can be verified, considering that $\chi^2_{\nu_1+\nu_2}=\chi^2_{\nu_1}+\chi^2_{\nu_2}$ and $\chi^2_{\nu_2-\nu_1}=\chi^2_{\nu_2}-\chi^2_{\nu_1}$. Therefore it is possible to write the following expression:

$$m_{(1-\beta)}(\nu_2,\nu_1)=\frac{1}{m_\beta(\nu_1,\nu_2)}. \tag{12}$$

3. Numerical examples

The numerical example presented in this section takes into consideration the evaluation of the number of wrong words transmitted in an automatic measurement system. In Table 2 the number of wrong words acquired in six different acquisition phases for two qualified error levels, equal to 8 and 9 LSB respectively, is shown.

Table 2. Experimental wrong words measurement .

# Test	**Wrong words in 10^7 samples (8 LSB)**	**Wrong words in 10^7 samples (9 LSB)**
1	77	2
2	76	4
3	64	3
4	86	2
5	71	8
6	61	4
Total number of wrong words	**435**	**23**

The following values for both the mean and variance have been calculated:

Variable	Mean	Variance
wrong words 8LSB	72,50	84,30 (s_1^2)
wrong words 9LSB	3,833	4,967 (s_2^2)

The idea behind this example is that if the standard deviation of the population is unknown in the calculation of the confidence interval for the wrong words mean of a high-dimension sample, it can be replaced with the sample standard deviation [3–5]. Therefore it can be very useful to determine confidence intervals for the variance and standard deviation, because in many practical applications the interval estimation for the variance σ^2 and standard deviation σ of the population are based on the sample variance s^2 and the sample standard deviation s.

So if we assume a normal distribution for the population of random variables that represent the wrong words transmitted, by extracting samples of size n (with n = 6) it is possible to write the Chi-square distribution, considering Eq. 6, as $\chi^2 = \frac{(n-1)s_1^2}{\sigma_1^2}$.

This is an important hypothesis because it leads to deal confidence intervals in non-symmetrical distribution. Using distributions with the same tail areas and indicating with $\frac{\alpha}{2}$ the area of each tail (see Fig. 2), the confidence interval for the variance of the population, with a confidence level, as percentage, equal to $(1-\alpha)\cdot 100\%$, is defined as:

$$\frac{(n-1)s_1^2}{\chi^2_{\frac{\alpha}{2}}} < \sigma^2 < \frac{(n-1)s_1^2}{\chi^2_{1-\frac{\alpha}{2}}}. \tag{13}$$

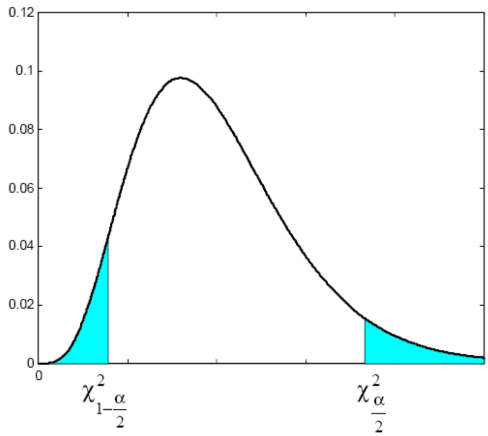


Fig. 2. A detail of the Chi-square distribution and its tails, used in this example.

For a confidence level equal to 95% and a degree of freedom $\nu = 6 - 1 = 5$, we obtain $\chi^2_{1-\frac{\alpha}{2}} = \chi^2_{0.975} = 0.831$ and $\chi^2_{\frac{\alpha}{2}} = \chi^2_{0.025} = 12.832$. Consequently, recalling Eq. 13, we deduce the following confidence interval for the population variance in terms of wrong words and confidence level of 95% as: $32.85 < \sigma_1^2 < 505.42$; $5.73 < \sigma_1 < 22.48 \cong 6 < \sigma_1 < 22$.

From the results so obtained it is possible to observe that the interval, with the above mentioned confidence level, is quite wide: this is due to the fact that the sample dimension, for this particular experiment, can never be too high, which obviously would allow us to reduce the interval size. A possible solution is to decrease the confidence level to 90% with the aim to find a compromise between the interval dimension and the correlated confidence level.

Another frequent situation is represented by two populations with variances unknown. However, if the sample variances are known, it is possible to compare the variances of two populations, always assuming a normal distribution for the two populations and that the samples, with size equal to n_1 and n_2 respectively, can be extracted independently. Introducing the sample variances as s_1^2 and s_2^2 respectively, with $s_1^2 > s_2^2$, the Eq. (9) can be written as $F = \dfrac{s_1^2 / \sigma_1^2}{s_2^2 / \sigma_2^2}$, that is an F distribution with parameters $\nu_1 = n_1 - 1$ e $\nu_2 = n_2 - 1$.

Using, also in this case, distributions with the same tail areas and denoting with $\frac{\alpha}{2}$ the area of each tail, the confidence interval for the ratio between the variances of each population, with a confidence level equal to (1 – α)·100%, is defined as:

$$\frac{s_1^2}{s_2^2}\frac{1}{F_{\frac{\alpha}{2}}} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{s_1^2}{s_2^2}\frac{1}{F_{1-\frac{\alpha}{2}}}. \tag{14}$$

For a confidence level equal to 95% and two degrees of freedom respectively $\nu_1 = 6 - 1 = 5$ and $\nu_2 = 6 - 1 = 5$, respectively, we obtain $F_{1-\frac{\alpha}{2}} = F_{0.975} = 0.140$ and $F_{\frac{\alpha}{2}} = F_{0.025} = 7.146$.

From Eq. 14, it is possible to evaluate the confidence interval for the ratio of the variances concerning two populations in terms of wrong words with a confidence coefficient of 95%:

$$2.375 < \frac{\sigma_1^2}{\sigma_1^2} < 121.23; \quad 1.54 < \frac{\sigma_1}{\sigma_2} < 11.01 \quad \cong \quad 2 < \frac{\sigma_1}{\sigma_2} < 11.$$

Also in this case, the interval dimension could be further reduced, not increasing the sample size, which proves quite complex in this particular experimental condition, but decreasing the coefficient to 90%, in order to find an optimal relationship between interval dimension and confidence level.

4. A statistical model for the process variability

As an application of F_{ν_1,ν_2} distribution, the paper takes into consideration the example presented in the GUM par. H5, [1].

Let us consider a set of n repeated observations throughout each day and suppose that such a set is reproduced in the following m days. We denote as v_{jk} the random variable associated with k-observation throughout the j-day.

The model adopted can be represented as:

$$\begin{cases} V_{jk} = m_0 + G_{jk} + T_j & j = 1,\cdots,m \;; \quad k = 1,\cdots,n \\ \overline{V_{j.}} = \sum_{k=1}^{n} V_{jk} / n = m_0 + \overline{G_{j.}} + T_j, & \text{with } \overline{G_j}. = \sum_{k=1}^{n} G_{jk} / n \\ \overline{\overline{V}}.. = \sum_{j=1}^{m} \overline{V_{j\cdot}} / m = m_0 + \overline{\overline{G}}.. + \overline{T}. & \text{with } \overline{\overline{G}}.. = \sum_{j=1}^{m} \overline{G_{j.}} \Big/ m; \;\; \overline{T}. = \sum_{j=1}^{m} T_j / m \end{cases} \tag{15}$$

G_{jk} and T_j denote the random errors, with expected values zero, which distinguish respectively the variability within a day (within variability) and the variability between days (in periods of time such as, for example, weeks, months, years – between variability). We hypothesize as normal the distribution of the model, so that:

$$\begin{cases} G_{jk} = N\left(0,\sigma_G^2\right);\ \overline{G}_{j.} = N\left(0,\sigma_G^2/n\right);\ T_J = N\left(0,\sigma_T^2\right); \overline{G}_{j\cdot} + T_J = N\left(0,\sigma_G^2/n + \sigma_T^2\right) \\ \overline{\overline{G}}_{..} = N\left(0,\sigma_G^2/nm\right);\ \overline{T}_{.} = N\left(0,\sigma_T^2/m\right) \\ \overline{\overline{G}}_{..} + \overline{T}_{.} = N\left(0,\sigma_G^2/nm + \sigma_T^2/m\right) \end{cases} \tag{16}$$

Due to the hypothesized independence among the observations, the random errors of the model are independent also in their mutual behavior.

Consequently, we deduce the following property:

$$\begin{cases} \left(V_{jk} - \overline{V_{j\cdot}}\right) \text{ is independent from } \overline{V_{j\cdot}} \text{ and therefore from } \overline{G_{j\cdot}} & (17.1) \\ \left(\overline{V_{j\cdot}} - \overline{\overline{V}}_{..}\right) \text{ is independent from } \overline{\overline{\mathrm{V}}}_{..} \text{ and therefore from } \overline{\overline{\mathrm{G}}}_{..} + \overline{T}_{.} & (17.2) \end{cases}$$

Now it is possible to consider the following equation:

$$\sum_{k=1}^{n} \frac{\left(V_{jk} - \overline{V_{j\cdot}}\right)^2}{\sigma_G^2} = \sum_{k=1}^{n} \frac{G_{jk}^2}{\sigma_G^2} - \frac{\overline{G}_{j\cdot}}{\sigma_G^2/n} = \chi_{j,n}^2 - \chi_{j,1}^2 = \chi_{j,n-1}^2 \tag{18}$$

with the Chi-square associated with j day and ν degrees of freedom.

Summing up Eq. (18) with respect to j, bearing for the property $\chi_{\nu_1+\nu_2}^2 = \chi_{\nu_1}^2 + \chi_{\nu_2}^2$ that $\sum_{j=1}^{m} \chi_{j,n-1}^2 = \chi_{m(n-1)}^2$ due to the independence from day to day and dividing by the degrees of freedom $m(n\text{-}1)$, we assume:

$$\frac{\sum_{j=1}^{m} \sum_{k=1}^{n} \left(V_{jk} - \overline{V_{j\cdot}}\right)^2}{m\,(n-1)} = \sigma_G^2 \frac{\chi_{m(n-1)}^2}{m\,(n-1)}. \tag{19}$$

Taking into account that $E\left\{\chi_\nu^2\right\} = \nu$, we can observe that the first member of Eq. (19) is an unbiased estimator $\tilde{\sigma}_G^2$ of σ_G^2, being $E\left\{\tilde{\sigma}_G^2\right\} = \sigma_G^2$.

Again, we introduce another important quantity, that is:

$$\frac{\sum_{j=1}^{m} \left(\overline{V}_{j\cdot} - \overline{\overline{V}}_{..}\right)^2}{\sigma_G^2/n + \sigma_T^2} = \frac{\sum_{j=1}^{m} \left(\overline{G}_{j\cdot} + T_j\right)^2}{\sigma_G^2/n + \sigma_T^2} - \frac{\left(\overline{\overline{G}}_{..} + \overline{T}_{.}\right)^2}{\left[\sigma_G^2/n + \sigma_T^2\right]/m} = \chi_m^2 - \chi_1^2 = \chi_{m-1}^2, \tag{20}$$

where the property 17.2 has been taken into account.

By analogy with eq. (19) we can introduce the quantity:

$$\frac{\sum_{j=1}^{m}\left(\overline{V}_{j\cdot}-\overline{\overline{V}}_{\cdot\cdot}\right)^2}{m-1}=\left[\frac{\sigma_G^2}{n}+\sigma_T^2\right]\frac{\chi_{m-1}^2}{m-1} \tag{21}$$

affirming that the first member represents an unbiased estimator of $\left[\frac{\sigma_G^2}{n}+\sigma_T^2\right]$.

Considering the estimators of Eqs. (19) and (21), we can also introduce an unbiased estimator $\tilde{\sigma}_T^2$ for σ_T^2, as:

$$\tilde{\sigma}_T^2=\frac{\sum_{j=1}^{m}\left(\overline{V}_{j\cdot}-\overline{\overline{V}}_{\cdot\cdot}\right)^2}{m-1}-\frac{\sum_{j=1}^{m}\sum_{k=1}^{n}\left(V_{jk}-\overline{V}_{j\cdot}\right)^2}{m(n-1)}. \tag{22}$$

Recalling the variable F_{ν_1,ν_2} defined in Eq. (9) and considering the unbiased estimators of Eq. (19) and (21), the random variable F_{ν_1,ν_2} can be represented as:

$$F\left(m[n-1],m-1\right)=\frac{\frac{\sigma_G^2}{n}+\sigma_T^2}{\sigma_G^2}\frac{\tilde{\sigma}_G^2}{\frac{\tilde{\sigma}_G^2}{n}+\tilde{\sigma}_T^2}. \tag{23}$$

In the particular case where the contribution of between-group variability (from day to day) is null, therefore $\sigma_T^2=0$, Eq. (23) can be simplified as:

$$F\left(m[n-1],m-1\right)=\frac{\tilde{\sigma}_G^2}{\tilde{\sigma}_G^2+n\tilde{\sigma}_T^2}. \tag{24}$$

Equation (24) can also be useful to test the hypothesis of the insignificance of the variability from day to day ($\sigma_T^2=0$). For this purpose, in Eq. (24), the estimators are substituted by the corresponding values obtained in any specific measurement. If the value of $F_{\nu_1,\nu_2}(\cdot,\cdot)$ so obtained is superior to 0,95-quantile of the F_{ν_1,ν_2} distribution (for example), it allows to reject the hypothesis and therefore to maintain that the variability from day to day is statistically significant with a risk of 5%.

5. Conclusions

The aim of this paper is to estimate the uncertainty interval, in the case of inherent variability of the measurement process, using Chi-square and Fisher distributions, that have not yet found a role in the GUM [1] as well as the supplement [2].

Simulations with a software that generates random values observed for different degrees of freedom and some practical examples were also developed in order to prove the theoretical approach.

In addition, the cases of within and between variability have been also studied, assuming a model for the process variability associated with the observations in different days. A test to assess the significance of the daily variability through the use of the distributions introduced in this paper complete this work.

References

[1] BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, OIML: *Guide to the Expression of Uncertainty in Measurement*. International Organization for Standardization, Geneva, Switzerland, 1995,

[2] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, 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*. Joint Committee for Guides in Metrology, Bureau International des Poids et Mesures, JCGM 101, 2008.

[3] M. Catelani, A. Zanobini, L.Ciani: "Introduction to the t and chi-square distribution for a more accurate evaluation of the measure of the Word Error Rate in Analog-to-Digital Converters". *Metrol. Meas. Syst.*, no. 4, vol. XV, 2008, pp. 483−488.

[4] G.E.P. Box, W.G. Hunter, J.S. Hunter: *Statistics for experimenters*. John Wiley, New York, 1978.

[5] M. Catelani, A. Zanobini, L.Ciani: "Statistical Analysis Of The Word Error Rate Measurement In Analog-To-Digital Converters". *Proc. of XIX IMEKO World Congress Fundamental and Applied Metrology*, Lisbon, Portugal, September 6−11, 2009, pp. 694–697.

[6] M. Kendall: *Kendall's Advanced Theory of Statistics*. Griffin & Company Limited, London, vol. 1, 2, 3, 1987, 1979, 1983.

[7] J.S. Bendat, A.G. Piersol: *Random data. Analysis and Measurament Procedures*. John Wiley & Sons, New York, 1986.

[8] C.F. Dietrich: *Uncertainty, Calibration and Probability*. The Statistics of Scientific and Industrial Measurament. Adam Higler, Bristol, 1991.

[9] W. Navidi: *Statistics for Engineers and Scientists.* McGraw Hill, New York, 2007.

[10] A. Papoulis: *Probability and Statistics.* Englewood Cliffs, Prentice-Hall, 1990.