

## GENERAL PROBLEMS OF METROLOGY AND MEASUREMENT TECHNIQUE

### BAYESIAN ESTIMATES OF SYSTEMATIC ERRORS OF MEANS OF MEASUREMENT

**N. A. Burmistrova, A. V. Stepanov,  
and A. G. Chunovkina**

UDC 53.081.1

*The problem of calculating the expanded uncertainty of measurements during calibration is examined. Two sources of measurement error are identified: the uncertainty of the reference value obtained from a standard and the dispersion in the readings from the measuring instrument that is being calibrated. A Bayesian approach is used to determine the dependence of the coverage factor on the number of repeated measurements and the relationship of these uncertainties.*

**Keywords:** *systematic error, expanded uncertainty, coverage factor, Bayesian estimate.*

The concept of measurement uncertainty [1–4] is used in metrology along with error estimates [5]. The question of which approach is preferable in these metrological problems remains open. Error calculations are, however, widely used in the calibration of means of measurement. During calibration, the systematic error of a measuring instrument is often determined as the deviation of the readings from a reference value of the measured quantity that is reproduced or obtained from a suitable standard.

The following are taken into account when calculating measurement uncertainty: uncertainty in the calibration characteristics of the standard measurement instruments or in the current values of the standard gauges, as well as their instability; nonlinearity of the calibration characteristic of the measurement instruments; random errors in the standards, the instrument being calibrated, and the calibration technique; and uncertainty in the corrections for the deviations from standard conditions specified for each type of measurement and in the method for transfer of the unit of the measured quantity.

This article discusses a simplified model for the systematic error  $B$  of a measuring instrument that is being calibrated:

$$B = X - X_{\text{ref}}, \quad (1)$$

where  $X$  is the average reading of the instrument being calibrated for a number of repeated measurements that approaches infinity, and  $X_{\text{ref}}$  is the reference value of the measured quantity.

For a specific form of measurement and calibration technique, model (1) becomes more complicated with the introduction of corrections for influence factors and the approach discussed below can be generalized for any particular case.

For measurements with the aid of a standard, an estimate of the measured quantity is obtained with the corresponding uncertainty  $\{x_{\text{ref}}, u(x_{\text{ref}})\}$  or with an indication of the limits of error  $\{x_{\text{ref}}, \theta\}$ . In the former case, information on the measured quantity can be represented by a normal density distribution  $p_{\text{ref}}(x)$  with mathematical expectation  $x_{\text{ref}}$  and standard deviation  $u(x_{\text{ref}})$  and in the latter case, by a uniform distribution with center  $x_{\text{ref}}$  and symmetric boundaries  $\pm\theta$  [6].

Usually, repeated measurements  $x_1, \dots, x_n$  are made during calibration in order to evaluate the combined random error owing to the standards, the instruments to be calibrated, and the calibration technique. Here it is assumed that the unit readouts of the measurement instrument are distributed normally with unknown parameters  $\{X_{\text{ref}} + B, \sigma\}$ , where  $\sigma$  is a precision (repeatability) parameter for the measurement results during calibration. Thus, in this problem there are three unknown parameters  $X_{\text{ref}}, B$ , and  $\sigma$ . The joint probability distribution for these parameters is found using Bayes' theorem on the basis of the available *a priori* information, as well as the information obtained during the measurements. We assume that these parameters are independent. We also assume that there is no *a priori* information on the bias of the readings from the measurement instrument or the precision of the measurements and that these are described by an uninformative *a priori* density distribution  $p(b, \sigma) \propto \sigma^{-1}$  for these parameters (the symbol  $\propto$  denotes equality to within a normalization factor).

Let us consider the first of these cases. Here the *a posteriori* joint distribution is specified by

$$p(x, b, \sigma | x_1, \dots, x_n, x_{\text{ref}}, u(x_{\text{ref}})) \propto \sigma^{-(n+1)} \exp\left\{-\frac{1}{2\sigma^2} \sum (x_i - b - x)^2\right\} \times \\ \times \exp\left\{-(x - x_{\text{ref}})^2 / [2u^2(x_{\text{ref}})]\right\}.$$

Integrating with respect to the interfering parameter makes it possible to find an *a posteriori* distribution for the systematic error:

$$p(b | x_1, \dots, x_n, x_{\text{ref}}, u(x_{\text{ref}})) = \int p(x, b, \sigma | x_1, \dots, x_n, x_{\text{ref}}, u(x_{\text{ref}})) d\sigma dx.$$

After some simple calculations, we obtain:

$$p(b) = \frac{1}{2^{(n-1)/2} \pi} \frac{((n-1)S^2)^{(n-1)/2}}{u(x_{\text{ref}})} \frac{\Gamma(n/2)}{\Gamma(n-1/2)} \frac{1}{\sigma^{n+1}} \times \\ \times \int_{-\infty}^{\infty} (n(b+x-\bar{x})^2 + (n-1)S^2)^{-n/2} \exp\left\{\frac{-(x-x_{\text{ref}})^2}{2u^2(x_{\text{ref}})}\right\} dx; \\ S^2 = \frac{1}{n-1} \sum_1^n (x_i - \bar{x})^2, \quad \bar{x} = \frac{1}{n} \sum_1^n x_i, \quad (2)$$

where  $\Gamma(n/2)$  and  $\Gamma(n-1/2)$  are gamma functions, and  $\bar{x}$  is the mean value for a finite sample.

Therefore, in the Bayesian approach the systematic error is described by the random quantity  $B$  in the distribution (2), for which the mathematical expectation and standard deviation are estimates of the error and standard uncertainty, respectively. An estimate for the systematic error can be obtained by substituting the estimates of the input quantities in Eq. (1):

$$\hat{b} = \bar{x} - x_{\text{ref}}.$$

According to [1], the corresponding standard uncertainty is calculated using the formula

$$u^2(b) = u_A^2(\bar{x}) + u_B^2(x_{\text{ref}}) = S^2 / n + u^2(x_{\text{ref}}). \quad (3)$$

The Bayesian approach, however, yields another formula for the standard uncertainty [7]:

$$\tilde{u}^2(b) = (n-1)S^2 / [(n-3)n] + u^2(x_{\text{ref}}). \quad (4)$$

Equation (4) is meaningful only for  $n > 3$ . It is interesting to note that in [8], repeated measurements are also taken to have  $n > 3$ . Equation (4) decreases monotonically and approaches Eq. (3) with increasing  $n$ . For sufficiently large  $n$ , Eqs. (3) and (4) are essentially the same.

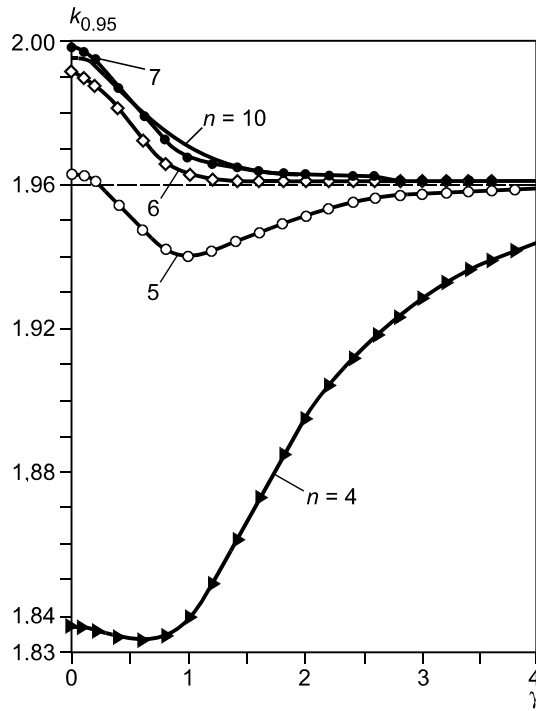


Fig. 1.  $k_{0.95}(\gamma, n)$  for the case in which there is a reference value and a corresponding uncertainty.

The purpose of this article is to calculate the expanded uncertainty of measurements for a probability (confidence level)  $P$ ; it is given by the product of the combined standard uncertainty and the coverage factor  $k$ :

$$U_p(b) = k\tilde{u}(b).$$

An expression for  $k$  with  $P = 0.95$  has been found for the cases discussed above ( $B$  and  $\sigma$  independent, with no *a priori* information on their possible values). In this case, it is generally assumed that  $k = 2$ . Here we analyze the differences in the estimated uncertainty for  $k = 2$  and for its exact value.

We consider the transformed random variable:

$$\tilde{B} = [B - (\bar{X} - X_{\text{ref}})]\sqrt{n} / S.$$

The distribution of  $\tilde{B}$  is determined by transforming Eq. (2) for  $B$ :

$$p(\tilde{b}) = \frac{1}{\pi\sqrt{2(n-1)}} \frac{\Gamma(n/2)}{\Gamma((n-1)/2)} \int_{-\infty}^{\infty} \left( \frac{1}{n-1} \left( z \frac{\sqrt{nu}(x_{\text{ref}})}{S} + \tilde{b} \right)^2 + 1 \right)^{-n/2} \exp\left\{-\frac{z^2}{2}\right\} dz. \quad (5)$$

Equation (5) is the convolution of the normal distribution and the Student scaling distribution with  $n - 1$  degrees of freedom, which depends on the parameter  $\gamma = n^{1/2}u(x_{\text{ref}})/S$ . As  $\gamma \rightarrow 0$ , the distribution (5) converges to the Student distribution, while for  $\gamma \rightarrow \infty$  it approaches the normal distribution. The parameter  $\gamma$  is the ratio of the standard uncertainty resulting from the standard that is used to the sample standard deviation of the averaged result, which characterizes the random error of the measurements during transfer of the unit of the quantity. Evidently, the values of this parameter are always limited in practice.

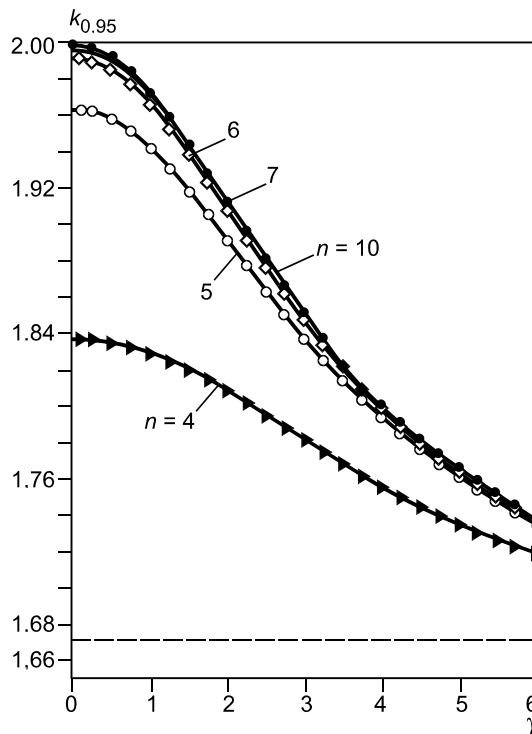


Fig. 2.  $k_{0.95}(\gamma, n)$  for the case in which there is a reference value and a limit on the permissible error.

We have calculated the percentage points  $\alpha_{0.95}$  of the (5) distribution for different values of  $\gamma$  and  $n$  and for symmetric probability intervals for  $P = 0.95$ . Accordingly, the expanded uncertainty of the initial quantity  $B$  is calculated using the formula

$$\Delta_{0.95} = \alpha_{0.95} S / \sqrt{n}.$$

The final goal was to obtain the functions  $k(\gamma, n)$ , which makes it possible to proceed from the standard to the expanded uncertainty of the measurement. When  $n > 3$ , Eq. (4) applies for the standard uncertainty, so the coefficient  $k$  is given by

$$k_{0.95} = \alpha_{0.95} [\gamma^2 + (n-1)/(n-3)]^{-1/2}.$$

Figure 1 is a plot of  $k_{0.95}$  as a function of  $\gamma$  for different values of  $n$ .

$k = 2$  is an upper bound estimate for  $P = 0.95$  if the standard uncertainty of the measurements is calculated using Eq. (4); for  $n > 5$ ,  $k$  differs from 2 by less than 2%. Similar calculations of  $k$  were done for the case when estimates of the reference value and the limit on its permissible error  $\{x_{\text{ref}}, \theta\}$  are known. Here the information on the value of the measured quantity is formalized as a uniform distribution with center  $x_{\text{ref}}$  and symmetric boundaries at  $\pm\theta$ . Values of  $k$  for calculating the expanded uncertainty were also obtained:

$$U_p(b) = k\tilde{u}(b), \quad \tilde{u}^2(b) = (n-1)S^2 / [(n-3)n] + \theta^2 / 3. \quad (6)$$

$k_{0.95}$  is plotted as a function of  $n$  and  $\gamma = n^{1/2}\theta/S$  for  $P = 0.95$  in Fig. 2. As opposed to the previous case, the values of  $k$  are significantly different for different  $\gamma$ . However, it should be noted that for  $P = 0.95$  and  $n > 4$ , the  $k(\gamma)$  curves are essentially the same for different  $n$ .

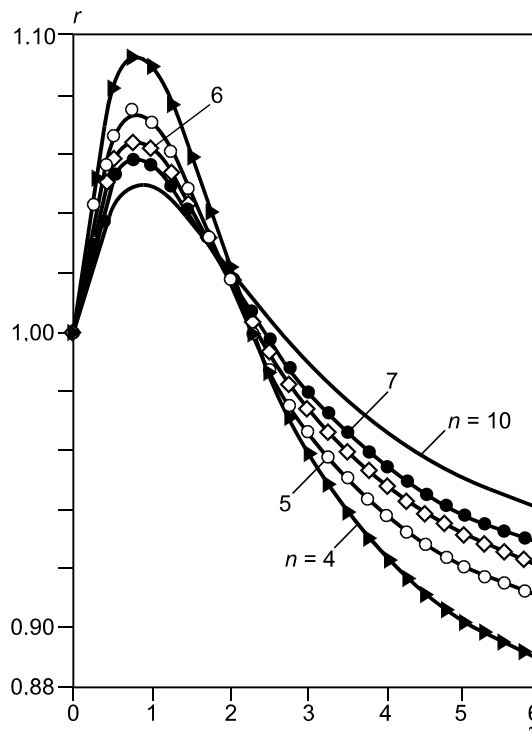


Fig. 3. The ratio  $r$  of the expanded measurement uncertainty calculated using Eq. (6) to the confidence limits on the error calculated using Eq. (7) as functions of  $n$  and  $\gamma = n^{1/2}\theta/S$  for  $P = 0.95$ .

There is some interest in comparing these estimates for the expanded uncertainty with the confidence limits on the error given in [8]:

$$\Delta_p = \sqrt{\frac{S^2}{n} + \frac{\theta^2}{3}} \left[ t_p(n-1) \frac{S}{\sqrt{n}} + \theta \right] \left[ \frac{S}{\sqrt{n}} + \frac{\theta}{\sqrt{3}} \right]^{-1}. \quad (7)$$

We now denote  $r$  the ratio of the expanded measurement uncertainty given by Eq. (6) to the confidence limits on the error given by Eq. (7). Figure 3 shows plots of  $r$  as a function of  $n$  and  $\gamma = n^{1/2}\theta/S$  for  $p = 0.95$ . Figure 4 shows that the difference between the two estimates of the accuracy can approach 10%. For  $0 < \gamma < 2$ , Eq. (7) gives a low estimate of the accuracy compared to the Bayesian estimate (6), and for  $\gamma > 2$  the situation is the opposite.

We conclude with an example of estimating the measurement uncertainty during calibration of a liquid thermometer. The following measurement data for the temperature of the thermometer to be calibrated: {40.01; 40.018; 40.020; 40.017; 40.014; 40.014; 40.011; 40.016; 40.011; 40.017},  $\bar{x} = 40.016$ . A standard was used to find the measured value  $x_{\text{ref}} = 39.999$  with a permitted error limit of  $\theta = 0.01$ . The systematic error of the thermometer was calculated using the formula  $b = \bar{x} - x_{\text{ref}}$  and found to be  $b = 0.017$ . Formal application of [1] gives

$$u_A^2(\bar{X}) = S^2 / n = 0.00307^2 / 10 = 9.4 \cdot 10^{-6};$$

$$u_B^2(\bar{X}) = \theta^2 / n = 0.01^2 / 3 = 3.33 \cdot 10^{-5};$$

$$u_{\Sigma}^2(b) = u_A^2(\bar{X}) + u_B^2(\bar{X}) = 3.4 \cdot 10^{-5};$$

$$U_{0.95}^{[1]} = 2\sqrt{u_A^2(\bar{X}) + u_B^2(\bar{X})} = 0.012.$$

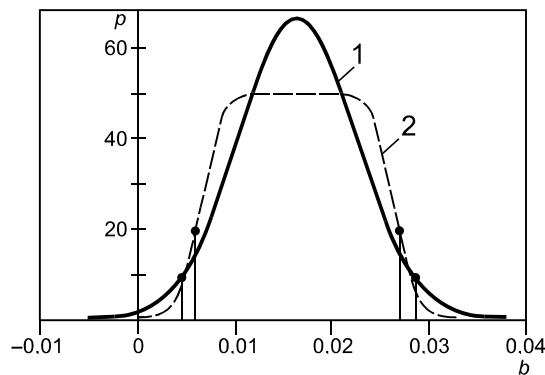


Fig. 4. The probability distributions for a normally distributed systematic error  $b$  (1) and for the Bayesian approach (2).

When the approach with  $k = 2$  was used, a coverage interval of (0.005; 0.29) was found for  $b$ . The choice of  $k = 2$  corresponds to the assumption of a normal distribution for the systematic error of the instrument. Figure 4 shows a normal distribution (curve 1). When a Bayesian estimate is used for the systematic error, the coverage interval is (0.006; 0.027). Curve 2 of Fig. 4 corresponds to the probability distribution for the systematic error. The points on curves 1 and 2 indicate the corresponding coverage intervals. The approach of Ref. 1 yields a wider coverage interval. It can be obtained using the curve of Fig. 3 for  $\gamma = 10.29$  and  $n = 10$ . For these values,  $k = 1.7$  and the corresponding coverage interval is (0.006; 0.027), in agreement with the interval obtained using the Bayesian approach. Finally, using Eq. (7) gives an interval of (0.006; 0.028) for the systematic error of the thermometer. At the same time, formal application of Ref. 1 yields an excessive estimate of the measurement uncertainty.

In this article, we have examined the problem of estimating measurement uncertainty during calibration, when the measurement equation is of the form (1). We have considered repeated measurements and two ways of representing the data on the reference value of the measured quantity. The coverage factor  $k$  and the expanded measurement uncertainty have been calculated using a Bayesian approach. The results have been compared with those obtained using [1] and the calculations of [8]. The conditions for using  $k = 2$  for the standard combined uncertainty according to the Bayesian approach have been formulated.

These calculations of the expanded uncertainty in a determination of the systematic error of a measuring instrument during calibration can be used to calculate the measurement uncertainty in a situation where the following measurement data are available: the results of more than three repeated readouts from the instrument; *a priori* information on the accuracy of this instrument, which can be represented by the instrumental uncertainty  $u_B$  or the measurement error limit  $\theta$ . In this case, it is necessary to find the distribution of the measured quantity, which can be obtained in a way similar to the calculations of [9].

In this situation with two sources of uncertainty, Eq. (7) for the confidence limits of the combined error [8] is within 10% of the Bayesian estimate for the expanded measurement uncertainty (6).

## REFERENCES

1. GOST R 54500.3-2011, ISO/IEC 98-3:2008 Guide. *Measurement Uncertainty. Part 3. Guide to the Expression of Uncertainty in Measurement.*
2. GOST R 54500.1-2011, ISO/IEC 98-1:2009 Guide. *Measurement Uncertainty. Part 1. Introduction to the Guide to the Expression of Uncertainty in Measurement.*
3. GOST R 54500.3.1-2011, ISO/IEC 98-3:2008/Supplement 1:2008. *Measurement Uncertainty. Part 3. Guide to the Expression of Uncertainty in Measurement. Supplement 1. Transformation of Distributions by a Monte-Carlo Method.*

4. GOST R 54500.3.2-2013, *ISO/IEC 98-3:2008/Supplement 2:2011. Measurement Uncertainty. Part 3. Guide to the Expression of Uncertainty in Measurement. Supplement 2. Generalization to the Case of an Arbitrary Number of Input Variables.*
5. RMG 29-2013, *GSI. Metrology. Basic Terms and Definitions.*
6. W. Wöger, “Information on a measured quantity as the basis for forming a probability density function,” *Izmer. Tekhn.*, No. 9, 3–8 (2003).
7. I. Lira and W. Wöger, “Comparison between the conventional and Bayesian approaches to evaluating measurement data,” *Metrologia*, **43**, 249–259 (2006).
8. GOST R 8.736-2011, *GSI. Repeated Direct Measurements. Methods of Data Processing. Basic Assumptions.*
9. A. Stepanov, N. Burmistrova, and A. Chunovkina, “Calculation of coverage intervals: Some study cases,” *Advanced Mathematical and Computational Tools in Metrology and Testing: Conf. Proc.*, 2014, St. Petersburg (2015), pp. 132–139.