Measurement Techniques, Vol. 53, No. 4, 2010

GENERAL PROBLEMS OF METROLOGY AND MEASUREMENT TECHNIQUE

AN ALTERNATIVE APPROACH TO THE ESTIMATION OF THE UNCERTAINTY OF MEASUREMENTS

E. A. Golubev

UDC 620.1.08

An approach to the estimation of measurement uncertainty based on assumptions that are adopted in the design of information and measurement systems is presented. Arguments that attest to the possibility of increasing the precision of estimation and its simplification by comparison with the approach adopted in the Guide to the Expression of Uncertainty in Measurement are presented.

Key words: *statistical model, random error, storage factor, uncertainty indices, preparatory stage, invariance of scatter, added uncertainty.*

In analyzing the approach proposed in the *Guide to the Expression of Uncertainty in Measurement* [1] from the standpoint of mathematical statistics, questions arise related to the need to improve the quality of the estimate of measurement uncertainty. In point of fact, the approach is based on an expansion of the indirect measurement transfer function into a Taylor series in degrees of increments of the input variables and and on a representation of the scatter that characterizes the uncertainty of the output variable, in the form of a weighted sum of scatters and covarances of the output variables. As follows from the subsequent presentation, far from all the parameters of this expansion and the required input variables may be expressed quantitatively, and for this reason it is not possible to learn to what extent the values of the uncertainty obtained will correspond to its calculated values [1, 2]. Therefore, it is essential to find new procedures that would be more adequate and more well-founded. The objective of the present study is to set forth a different approach to estimation of measurement uncertainty and to present the advantages of this approach.

Below, an attempt will be made to look into the procedural questions involved in estimating measurement uncertainty from the standpoint not of metrology, but instead of mathematical statistics. In view of the fact that the participation of specialists in statistics in questions related to estimation of uncertainty is clearly inadequate, there has arisen a serious element of incompatibility between the assumptions adopted in the basic documents (handbooks) on the expression of uncertainty and certain principles of statistics. Meanwhile, the majority of specialists in both areas (in metrology and in mathematical statistics) are not at all aware of this incompatibility.

As a start, let us turn our attention to the fact that the statistical approach and mathematical statistics itself necessarily entails an extension of this application domain to human beings' information needs and to important measurement problems, chiefly in the areas of detection and ranging and in communication. These two scientific disciplines emerged out of these problems and have achieved great success not only in their own internal development, but also in practical applications. It is precisely the statistical approach that has provided a methodological foundation for the creation of analogous systems, since

All-Russia Research Institute of Metrological Service (VNIIMS), Moscow, Russia; e-mail: goledu@mail.ru.
Translated from Izmeritel'naya Tekhnika, No. 4, pp. 35–41, April, 2010. Original article submitted December 30, 2009.

0543-1972/10/5304-0402

procedures for finding the best structure of different systems on the basis of previously developed optimality criteria may be formalized from the standpoint of the statistical approach. In their simplest form, such problems have now arisen in the area of methodology as well, as problems of optimization of decision rules in conformity estimation. But in relation to measurement uncertainty itself, it is important to bear in mind that, on the one hand, the statistical approach does not only refer to the use of probabilistic methods, and on the other hand, any detector is a measurement tool, moreover, the methods used to determine the characteristics of the variability of its readings should not be distinguished from the corresponding methods used in metrology. However, in the form in which they are set forth in the Guide [1] and its appendices, these methods unfortunately differ and, in fact, quite substantially.

Statistical Models. The process of finding a parallel between modern metrology and mathematical statistics leads not only to the discovery of a difference in approaches adopted to the estimation and application of the statistical characteristics of variable quantities, but also the understanding of the potential for important borrowings and generalizations. For example, in the construction of statistical models in metrology, most specialists divide the set of influencing factors (sources of uncertainty) into two parts, factors that induce random measurement errors and those that induce systematic errors. By [1], a finite (non-infinite) number of observations in a series is assumed to be the only source of random errors in both output and input variables, whereas the sample mean and standard deviation of the results of *observations*, respectively correspond to the measured value and the "random" component of uncertainty.

As regards the "systematic" component of uncertainty, it incorporates the contributions of the remaining nine out of ten sources of uncertainty enumerated in [1]. These sources characterize the variability of measurement conditions on both the current and all preceding stages (elements) of the measurement process, beginning with calibration of secondary standards. Moreover, it is not directly stated in [1] that uncertainty develops in each set, whether it is the result of the overall influence of factors that are inherent exclusively to the elements of a unified chain of traceability to the primary standard, or an indicator characterizing the entire set of such chains and copies of the tool used to measure the given variable on all stages. The second type of uncertainty, which is widely used, and unjustifiably so, at the present time in practical calculations of uncertainty may be ten times as great as the first type.

From the standpoint of mathematical statistics, the model used to describe the effect of influencing factors on measured values (noise model) that is usually applied in metrology can hardly be called a statistical model. (The exception here are the models used in the ISO 5725 standard [30]). It is more proper to refer to them as interpretations of noise, where it is not so much a detailed study of the noise that is based on the model as it is the ease in obtaining and processing the results of observations. In particular, it is believed that in indirect measurements performed in series, it is sufficient in every case to perform n observations and then average the results of these observations in order to achieve an n-fold reduction in the scatter of the random error. In the statistical approach, the requirement of independence of repeated observations in serial observations means that noise must be wideband to the extent required and that if the repeatability conditions are satisfied, this will determine its degree of required stationarily. But since an appropratie study of the properties of noise is generally not undertaken, we are not justified assuming that the scatter will decrease by a factor n upon averaging.

Another feature of the "metrological" model of noise is that for all types of measurement, the influencing factors and the input variables that depend on these factors are divided into two classes, with the variability of the classes being characterized by very different time constants. For the first class (variables estimated according to type A), the time constants (correlation intervals) of the variations are much less than the length of the series of observations, while for the second class (variables estimated according to type B), the opposite is the case, that is, the time constants are much greater than the length of the series. It is assumed that there are no intermediate influencing factors whatsoever, and even that all the type B input variables (in particular, those that depend on external measurement conditions) are random but only on a different time scale and not on the set of repeated observations in the series but rather on the set of the actual measurements, in particular, single (isolated) measurements. Moreover, from the standpoint of the statistical approach, the interpretation of questions related to a benchmark relative to which these discrepancies (errors) are calculated also changes. Most metrological specialists would like to see in a benchmark value the true value of a measurable variable, and now in view of the adoption of the VIM-3 dictionary [2], a set of values of the measured variables that satisfy its definition. But here specialists in mathematical statistics would be unanimous in identifying this set with the overall mean of the measurable variable, i.e., the mean value that is

obtained in the course of a lengthy period of time (measurement cycle) on a set of several series of repeated observations or a set of isolated measurements. And indeed there are very weighty reasons for this conclusion.

In the applied statistical disciplines (these include statistical physics, statistical radiophysics, statistical radio engineering, statistical communication theory, as well as game theory, statistical decision theory, and others), a general concept of noise (i.e., one that does not depend on its origin) is introduced, given as the manifestation of variability of the conditions under which measurements are performed and the overall effect of any type of influencing factors. The basic point here is the universality of noise and a clearly expressed tendency to study the properties of noise, directed not at the mechanisms involved in the appearance or effect of individual influencing factors, but rather in a focused investigation by means of statistical methods. As regards estimation of the uncertainty of indirect measurement, this means:

1) the rejection of any study of the characteristics of individual input variables and individual influencing factors and the study of the properties of only a single (output) variable on a set of realizations of this variable; and

2) full account of all the influencing factors (including the unknown factors), independently of their time (correlation) properties.

In numerous studies devoted to the subject of uncertainty, beginning with the Guide [1], it has been repeatedly claimed that uncertainty and error are two different concepts, though such an assertion does not require any special rationale. First, error, understood as the difference between a measured and a measurable value, serves to quantitatively characterize a single measurement, whereas uncertainty is an aggregate indicator. Second, the pivotal values differ in the representation of the concepts of error and uncertainty. Third, from the standpoint of probability theory, there exists at least two distinct indicators characterizing the two properties of a set of measured values in a set of errors of measurements of a single specimen, the bias and the scatter, and by definition, uncertainty can be related solely to the latter. This means that of the two components of error (systematic error and random error), only the second may constitute a basis of uncertainty. However, it is not the type of random error that is assumed in [1].

In the interpretations of realizations of a random error as noise, a transition from a "metrological" model of noise to a statistical model requires an adjustment of the concept of a random error. First, the inclusion in the concept of changes in the values of the output variable induced by "slower factors" that generate type B uncertainty presupposes a transition to a significantly greater volume of the experimental data used to estimate the scatter coefficient. Further, a transition to a statistical model entails a replacement of the set of pivotal values corresponding to the sample mean for several series by a single pivotal value (i.e., an overall mean for an entire cycle of measurements that includes a set of several series or set of single measurements). Moreover, noise (for example, standard deviation) must grow due to the inclusion of new factors that correspond to all the known and unknown sources of uncertainty the effect of which varies *nonmonotonically* over time throughout this entire cycle (variable factors). And, finally, the constraint related to the division of sources of noise into two types is lifted and noise is now considered over time throughout the cycle as a random process that automatically incorporates any (in terms of length) physical effects of the measurement conditions on the measured value, and not just very brief and very lengthy effects. The new standard deviation now includes types of scatter that are induced by such factors as "imprecision in the introduction of corrections into known systematic effects" or "nonexcluded systematic errors." It is precisely this type of random error that underlies the concept of uncertainty set forth in [1, 2], and it is precisely its statistical characteristics (variance, standard deviation, confidence probability) obtained on a set of measured values that are the measures of uncertainty.

Note that in the absence of additional information, only the mean value (sample mean when processing repeated observations or overall mean when processing measured values), which is the best estimator of a measurable value, can be a measurable variable. It is precisely the mean value that assures the least discrepancy in the correction and the lowest nonexcluded error caused by the the same nonmonotonic factors as well as finiteness (non-infiniteness) of the number of observations in each series. In subtracting the mean value from a measured value, the "excluded systematic error" will also be subtracted.

Preparatory Stage. The required stage of measurements and may be performed and the statistical characteristics of the noise may be found on a special preparatory stage, which must be preceded by "working" measurements that are performed independently of these measurements and are based on repeated measurements of several consistent specimens. These specimens must encompass the required range of measurements and the preparatory stage itself must consume an intermediate length of time sufficient for reproduction of different natural conditions and effects of different influencing factors in their

natural combinations. As for the number of measurements performed, this number will not play an important role in estimating the required length of the preparatory stage.

We wish to emphasize once again that, in particular, in estimating the statistical properties of a measured output variable, like any random variable it is sufficient to rely on *its* measured values and that the use of input variables and influencing factors only complicates this estimation. But, unfortunately, for indirect measurements the methods of estimating the uncertainty of an output variable that are proposed in the Guide [1] and its appendices are based precisely on the use of input variables and influencing factors, even though no such proof whatsoever demonstrating the need for such an approach is given. Thus, we may recall in this connection that input variables and influencing factors are not present in the definition of uncertainty. These are also absent from the ordinary formula for the sample variance of an output variable y:

$$s^2(y) = \frac{1}{n-1}\sum_z (y_z - \bar{y})^2, \tag{1}$$

where $\bar{y}$ is the sample mean of n *measured* values y_z.

The propagation law [1] used to estimate the uncertainty $u_c(y)$:

$$u_c^2(y) = \sum\nolimits_i (\partial f / \partial x_i)^2 u^2(x_i) + \sum\nolimits_i \sum\nolimits_j (\partial f / \partial x_i)(\partial f / \partial x_j) u(x_i x_j). \tag{2}$$

also constitutes an expansion of the sample variance (1) of the output variable y into a series in the *input variables* x_i. Here $y = f(x_i)$ is a transfer function (indirect measurement "equation").

It is difficult to say why the authors of [1] found it necessary to expand the output variance into such a series. Of course, the influencing factors could have been taken into account by augmenting the variance of the output variable itself, and not through the use of the variances $u^2(x_i)$ and the covariance $u(x_i, x_j)$ of a series of correlated input variables. Moreover, here the focus is on unknown influence factors and on factors that cannot be measured and introduced into the function $y = f(x_i)$.

The principal reason for the fact that the input variables x_i and x_j are correlated derives from the dependence on one and the same effect (manifestation of a "slow" influencing factor") in practically simultaneous measurement of the variables [4]. This mechanism of interconnectedness as a consequence of the influence of a factor on both quantities was not taken into account in [1], where the input random variables are generally considered to be independent. This means that when applying a law of propagation of uncertainties, it is necessary to know not only the variances $u^2(x_i)$ but also the covariances $u(x_i, x_j)$. However, we cannot depend on published studies or specifications for an answer to this question, and not only because such data are absent. Even when these data are available, they would not have been suitable due to the lack of the correspondence between measurement conditions and conditions in the laboratory that uses and extends these data in estimating uncertainty. At the same time, the experimental determination of correlation properties of a set of input variables is a problem that is *far more complex than the direct estimation of the uncertainty of a measurement of a single output variable*.

The total influence of the variability of measurement conditions characteristic of testing and measurement laboratories on the statistical characteristics of an output variable may be established, but only empirically. And, in addition, it is entirely essential that an estimate of uncertainty consist in a determination of the influencing factors and subsequently in the summation of these factors with the use of the law of propagation. Of course, it is possible to at once establish an *accumulating* factor that would combine in a natural way all the influencing factors, including the unknown factors. The time needed to perform repeated measurements on the preparatory stage that was referred to earlier is such a factor for internal laboratory measurements and calibrations. Thus, let us now consider whether it is possible to use the concept of total noise when attempting to find the uncertainty and necessity of an experimental estimation of the properties of noise on a preparatory stage. As a result of numerous measurements performed on a consistent specimen in the course of a lengthy interval of time that reproduces the physical conditions under which measurements are performed, values of the uncertainty for different segments of the range of measurement will be obtained and interpolated. A corresponding curve for the *entire* range of measurements may be used subsequently in assignment of uncertainty to the "working" measured values. Thus, the procedure of measuring uncertainty does not involve performing any working measurements. The estimation of uncertainty is transformed

from a complex problem which, according to the Guide [1], requires the intellectual efforts of highly skilled metrological specialists, basically into a routine procedure involving control of the consistency of measurements, i.e., into a procedure which, one way or the other, must be carried out independently of the estimation of uncertainty. Difference between the procedures relate to the processing of the measured values.

As for the "monotone" factors, primarily drift, these are not sources of random error and *do not have any effect on the uncertainty.* However, in measurements performed on the preparatory stage, they will, unfortunately, induce a progressive *bias* in the measured values which must be taken into account when estimating the overall mean and the uncertainty itself. For this purpose, it would be useful, for example, throughout the preparatory stage, to approximate the measured values by a straight line (line of sample means) in a graph that depicts the dependence of the measured values on time. Next, in the process of estimating uncertainty, this (in the general case, inclined) line may be used as the axis of the overall mean from which the random errors of both individual measurements that have already been realized on the preparatory stage as well as future "working" measurements must then be computed.

The problem which is the focus of the present article does not include the description of the details of an experiment that relate, in particular, to estimation of an initial pivotal value at the start of the preparatory stage and to adaptive estimation of the mean value and standard deviation. Neither are we considering the elimination of outliers or the need to take into account the influence of a decrease in the frequency of isolated measurements (or of several series of repeated observations) throughout the preparatory stage or to the need to take into account on this stage the number of measurements of the object, performed in a "working" regime. These details will be different for different measurements.

Recall that in the case of measurements of several series, the sample mean is the mean value of n observations in a series. If a judgement as to correspondence in the course of testing or calibration is adopted as the basis of a single measurement, then this sample mean in particular will be the measured value. On a set of several series (of measured values) the sample mean is a random variable the mean value of which is called the overall mean. The sample variance, defined over n observations in a series (cf. formula (1)) is also a random variable. The variance of the measured values would be equal to $(1/n)$th part of the sample variance of the observations if the observations may be considered to be independent.

We would like to add that in statistical models it is not necessary to distinguish between the concept of observation and that of measurement. The former cannot be applied in general, since the question whether y_z is the result of averaging secondary observations in a series or is the value of a single measurement becomes inessential. And in fact, all questions related to estimation of type A uncertainty become superfluous [1].

Pivotal Values and the Invariance of Scatter. Formalized proofs demonstrating that the approach to estimation of uncertainty described above, which relies on measurements in a preparatory stage, are closely related to the relationship between the concept of uncertainty and that of precision. By [2], precision is a properry of multiple measurements that characterizes the proximity of the results of these measurements. The degree of proximity, moreover, depends on the conditions under which the measurements are performed that are in effect in the particular laboratory or in a set of laboratories, while the influencing factors are basically determined by a sample from the above "limiting" set of sources of uncertainty presented in [1].

Let us now turn to the question touched on at the start of the article, that is, why is it that in estimating uncertainty it is necessary to use mean values as "values that may be assigned to a measurable quantity? In principle, there is a simple answer. It is a consequence of Bayes' theorem. From this theorem it also follows that under identical measurement conditions, the differential distributions of the random variables $\Delta y_g = y_{gz} - y_0$ (y_{gz} is a random value representing values that may be assigned to a measured quantity y_0) and $\Delta y = y_z - \bar{y}$ (a quantity that describes a random correction) are antisymmetric relative to the midpoint of the interval $(y_0, \bar{y})$ [5]. Moreover, the two random values Δy_g and Δy, which serve as the basis for uncertainty and precision, are equal in terms of modulus and opposite in terms of sign. Hence, it follows that all the measures of the scatter, including the limiting and the sample measures for the uncertainty and the precision, as well as the fact that the variance of the uncertainty (precision) and other characteristics of the scatter (standard deviation and confidence interval) may, in fact, be determined from values that are measured on the preliminary stage (for variance, by means of (1)) if it is not the *mean values* but instead the *overall means* which are used as the pivotal values.

That the variance of uncertainty and the variance of precision are identical may also be proved by relying not on (1), but instead on the derivation of the law of propagation of uncertainty. As a result, one and the same expression on the right-

hand side of (2) is obtained in both cases. At the same time, if it is assumed that in (2), the pivotal values are not equal to the mean and contain nonzero additional terms, additional terms will appear on the right-hand side of (2). This means that the word, "scatter" in the statement of the concept of uncertainty [1] does not have any relationship to the true, objective, or any other values of the measurable quantity, except for the mean value.

Questions related to the fact that the characteristics of uncertainty and precision coincide are considered in greater detail in [5, 6].

Let us now turn our attention to the fact that the characteristics of uncertainty (like the characteristics of precision) reflect not so much the properties of the measurement object, for example, the dimension of a measurable quantity, as imperfections in the measurement procedure. This circumstance forces us to compare the scales of typical (for example, standard) deviations with the scales of bias that induce significant variations in the scatter, and recall that a typical overall error in a measured value is 10^{-2}–10^{-3} part of a measurable quantity and, consequently, the random error is even less. There question arises, what bias is needed to achieve a significant variation in the random error?

In the techniques used to estimate precision described in [2], most of the focus is on experimental ascertainment of the dependence of the precision indices on the values of the measurable quantity in the measurement range. Significant variations in the indices are observed only in the transition from one of four or five levels to another in the measurement range, i.e., in a bias that amounts to roughly 20% of the measurement range. This means that the *values of the uncertainty*, like the value of any ordinary measure of the scatter (variance, standard deviation, confidence interval) that relies on the mean value, *do not greatly vary with bias in the measured value within the measurement level, i.e., on intervals that significantly exceed the maximal scatter* (principle of invariance of scatter of measured values relative to bias). It is not hard to overstate the importance of this circumstance, since the true, objective, and mean values of a measurable quantity and all the initial and working pivotal values are found within these intervals. Thus, scatter, which is estimated in multiple measurements of any consistent specimen, preserves its value for all specimens of a given level of measurements, in particular, for all references and for all *standard specimens.* And *from the standpoint of uncertainty estimation*, a reference of any grade is no better than any consistent substitute (equivalent) of the same measurement level. Moreover, the value of a consistent specimen (reference) cannot be used to estimate uncertainty; only the mean value of the results of measurements of this specimen can, and must, be used.

In principle, the interval of values of a measurable quantity on which there occurs a significant variation in the standard deviation of the uncertainty (precision) may serve as an estimate of the width of the measurement level.

Besides the fact that it is possible to construct indices of the scatter of a measured quantity on sets of measured values that accumulate influencing factors and the recognition of the dependence of the variance of the added uncertainty (intermediate precision) on the conditions under which the measurements are performed, the principle of invariance of the scatter also enables us to create a consistent interpretation, a system of indices, and a general technique of a priori *measurement* of uncertainty for any mode of measurement, testing, and calibration.

Only through the use of sample means as possible values of a measurable quantity is it possible to distinguish measurements of scatter from measurements of bias, since, in particular, it assures the mutual independence of the procedures and results of any estimate of precision (uncertainty), on the one hand, and the calibration of references at different levels, on the other hand. The mean values, and only the mean values, may become pivotal values in the measurement of the scatter [2] and are not related to measurements of the bias such that references may not be replaced by any kind of mean value. Measurements of scatter do not infringe in any way on the principles underlying the organization of measurements associated with estimation of the "true" value of a measurable quantity, compensation of drift, or traceability. At the same time, the principles of definition, use, and registration of *statistical* indices related to measurements (such as uncertainty, precision, and reproducibility) and the methods and procedures used for comparison and estimation of correspondence form a unified domain of metrology related to the calculation of scatter. This domain has just as much a right to be referred to as statistical metrology as do the analogous appied statistical disciplines that we discussed earlier.

Stages in the Measurement Process and Added Uncertainty. Formula (2) was obtained long before the publication of the Guide [1] for use in the expression of the variance of the random error of an output quantity in indirect measurements in terms of the characteristics of the random errors of the input quantities (cf. [7]). Random errors, moreover, are understood

in the traditional sense, as distortions within a series of observations. Of course, this formula is of a more general nature, however, since it may relate to any variance of an output quantity. In particular, it is valid for different combinations of influencing factors and for different sets, for example, in a transition from time scales of a series of observations to time scales of a cycle of measurements on a preliminary stage. Formula (1) is no less general in nature. However, the values of the variance of uncertainty (precision) obtained with the use of this formula will naturally differ for different measurement conditions (combinations of influencing factors and their manifestation over time), for example, in a transition from one laboratory to another.

Thus, it has been found that there has to be very many uncertainty indices, in fact, as many as there are precision indices, and that they must differ in the way described in [2, 8]. Recall that the indices correspond to different quantities and combinations of "included" (i.e., in effect in the course of measurements) influencing factors that determine the measurement conditions. For testing and calibration laboratories, such factors are the "operator," "instrument," "calibration regime," and "external measurement conditions." For example, if a single model of an instrument is used under the ordinary working conditions of a laboratory and if the measurements are performed by a single operator and in the single calibration regime, the corresponding factors "are not included" and there will remain on the preparatory regime a single complex factor – the external measurement conditions.

Earlier, it was remarked that an equivalence between the conditions under which measurements are performed is required in order to assure equality between the variance of precision and the variance of uncertainty. However, this condition itself cannot be satisfied without some modification in the concept of precision and in the concept of uncertainty (since the measures of the scatter are formed on different sets). But this means that it is not possible to estimate the variance of the uncertainty empirically. On the one hand, the concept of the variance of uncertainty that depends on external factors is not used in [1] and has a single value corresponding to a given concrete chain of dissemination of the unit of measurement and the set of influencing factors that is in effect in all the links of this chain. On the other hand, the variance, say of the reproducibility (or limiting precision) characterizes a stage of the measurement process and is related to a set of links (calibration or testing laboratories) of a single level for different measurement links.

The requirement to encompass all the stages of a measurement process and the requirement that it be possible to use the formula

$$s^2(y) = \frac{1}{n-1}\sum_z (y_z - \bar{\bar{y}})^2 \tag{3}$$

instead of (2), where $\bar{\bar{y}}$ is the overall mean in a measurement of a given specimen, may be combined only with the introduction of the concept of the *added* uncertainty that arises in each individual link (on each individual stage) of the chain of dissemination of a unit of measurement. In this case, the added uncertainty will be equal to the variance of the precision for a given link, while the overall variance of the uncertainty becomes the sum of the added uncertainties. This means that each link (each laboratory) in this chain may estimate its proper added variance of the uncertainty as the variance of the precision on the basis of repeated measurements of a consistent specimen (in different realizations of the measurement conditions) and process the results according to (1). It is, however, not necessary to resort to "extraneous" sets, for example, derived models of the testing, calibration, and measurement instruments. To obtain the overall variance of the uncertainty, it is necessary to add the added variance to the overall variance of all the preceding stages.

In fact, in the present article the process of constructing a description of the experimental estimation of uncertainty on the preparatory stage has been related precisely to the added uncertainty. It is important to bear in mind that in accordance with the principle of uncertainty (precision) of the scatter, the variance that is added (augmented) on each stage of the measurement process may be related both to an *unknown* value of the measureable quantity, as well as to the measured value (in particular, to the result of a single measurement), to which the value of the index of uncertainty must be assigned.

Note that our suggested replacement of multiple pivotal values corresponding to several individual series of observations by a single pivotal value, i.e., an overall mean, in the measurement of a given specimen on the entire preparatory stage, places the pivotal value in the domain of "true" values of and suggests estimating uncertainty as the standard deviation from the effective value. However, a proof the possibility of making such an estimate based on the degree of proximity of the overall mean and the effective value for each concrete case still entails the need to find the overall mean. In such types of statis-

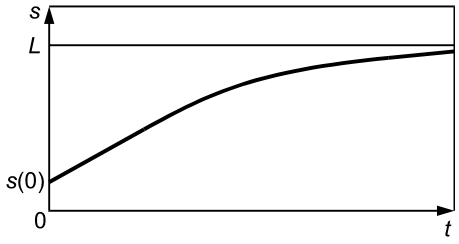

Fig. 1. Graph of standard deviation.

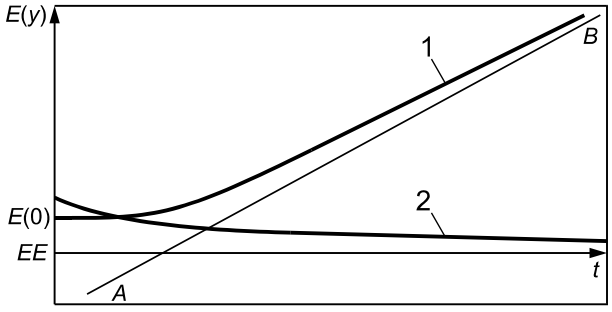

Fig. 2. Graphs of pivotal values in the absence of drift (curve *1*) and where drift is present (curve *2*).

tical problems involving the determination of the scatter, it is important to bear in mind that the overall mean is always more preferable, since for high n even a slight "systematic variation" that is the same for all n terms on the right-hand side of (3) may substantially increase the variance.

If in the course of testing or calibration, the procedure used to estimate the measureable quantity does not reduce to a single measurement, and instead requires several (a series of) *measurements* with subsequent averaging of the results, estimation of the uncertainty on the preliminary stage must be performed on the basis of such a procedure. Thus, it is not the result of a single measurement that should be considered the measurable quantity, but instead the mean value of such a series. In this case, for both single as well as multiple measurements the general nature of the dependence of the estimated uncertainty on the length of the preparatory stage is described by the curves presented in Figs. 1 and 2. Here may be seen graphs that describe the variation of the standard deviation and mean of a random variable that has been measured on intervals $(0, t)$ with increasing length t of the preparatory stage. In Fig. 2, curve *1* corresponds to the extreme case of absence of drift; in addition, the asymptote *EE* (line of overall mean) is selected as the horizontal axis, while curve *2* corresponds to the general case of the presence of drift and inclined asymptotes. Moreover, curves *1* and *2* and the asymptotes may be situated entirely or in part in the upper as well as the lower half-plane. The two figures should be considered only as illustrations of trends, and the physical curves will not be smooth. The objective of a construction of the graphs is to estimate the asymptotic value L and the position of the asymptote *AB*, the axis of the overall mean. The standard deviation $s(t)$ and the sample mean $E(y) = \bar{y}$ estimated on the interval $(0, t)$ tend to these asymptotes.

In conclusion, let us demonstrate the importance of the principle of invariance using as an example the process of estimating uncertainty of analytic measurements of inhomogeneous media. According to current views, for such media it is not possible to propose or produce a representative specimen measurement of which would make it possible to estimate uncertainty [9]. However, according to the principle of invariance for an estimate of uncertainty, *arbitrary* fragments (probes) may be used as the measurable specimen; it is only necessary that the value of the measurable parameter of the medium not exceed the common value for all fragments of the measurement level. Measurements of other fragments must reduce approximately to the same value of the uncertainty. The degree of approximation, moreover, may be insignificant, indeed, in every

case not comparable with the precision of the measurement itself. As a result, once the uncertainty of one fragment has been estimated, this uncertainty may be assigned to the other fragments. This also corresponds to the assumption that the concept of uncertainty characterizes above all a procedure and conditions of a measurement, but not to the measurement object.

Thus, it is the present writer's opinion that a reliable estimate of the uncertainty of a measurement cannot be achieved without repeated measurements that encompass different measurement conditions. And the procedure used to estimate uncertainty that is recommended by the Guide [1] and is based on the use of (in general case, correlated) input quantities is more complex and less accurate than the procedure proposed in the present article. At the same time, with the proposed approach it becomes possible to take everything into account, including unknown influencing factors, and to separate the process of estimating uncertainty from that of estimating the measurable quantity and avoid the use of *references and standard specimens.* As a result, it becomes possible to assign in simple fashion different values of the uncertainty obtained from the data of the preparatory stage to the measured values and to arrive at practical decisions, in particular, a correspondence in the course of subsequent (repeated) performance of "working" measurements, tests, and calibrations.

REFERENCES

1. *Guide to the Expression of Uncertainty in Measurement*, ISO, 1993 [Russian translation] (1999).
2. ISO/IEK Guide 99, *International Vocabulary of Metrology. Basic and General Concepts and Associated Terms (VIM)* (first edition).
3. ISO 5725: 1994-1998 (GOST R ISO 5725-2002), *Accuracy (trueness and precision) of Measurement Methods and Results. Parts 1–6*.
4. E. A. Golubev, "A comparison of different approaches to estimating uncertainty," *Izmer. Tekhn.*, No. 3, 6 (2008); *Measur. Techn.*, **51**, No. 3, 232 (2008).
5. E. A. Golubev, "Uncertainty in measurements and GOST R ISO 5725," *Zav. Lab. Diagnos. Mater.*, No. 6, 63 (2007).
6. E. A. Golubev, "A comparison of uncertainty and precision," *Metrologiya*, No. 8, 3 (2007).
7. S. G. Rabinovich, *Error in Measurements* [in Russian], Energiya, Leningrad (1978).
8. E. A. Golubev and L. K. Isaev, *GOST R ISO 5725. Basic Assumption. Questions of Adoption and Introduction. Measurements, Control, Quality* [in Russian], FGUP Standartinform, Moscow (2005).
9. R. L. Kadis, "The metrological and statistical meanings of the concept of "precision" in chemical analysis," *Zav. Lab. Diagnos. Mater.*, No. 12, 53 (2005); No. 2, 53 (2006).