Further development of testing computerization should focus on the creation of new types of computer-aided technical facilities that will reduce the range of testing hardware required and extend the range of testable measuring instruments. Given the wide range of measuring instruments in use, we should stress the significance of computer-aided testing algorithms. The production of such software–hardware instrument testing systems should be organized in order to expedite their distribution to enterprises and metrological services. It is also necessary to review the corresponding standards and technical documentation and to refine the required metrological tolerances allowing for the self-testing capabilities of the work-stations.

Finally, the organization of the national metrological service should be reviewed, allowing for the new possibilities provided by computer-aided instrument testing.

LITERATURE CITED

1. G. D. Kolmogorov, Izmer. Tekh., No. 11, 3 (1987).
2. N. V. Gelashvili et al., Izmer. Tekh., No. 6, 3 (1987).
3. A. I. Mekhannikov, Izmer. Tekh., No. 11, 5 (1987).
4. N. N. Vostroknutov et al., Izmer. Tekh., No. 11, 16 (1987).
5. V. A. Nosenko and O. F. Chuevskii, Tekh. Sredstv Svyazi, Ser. RIT, No. 5, 15 (1988).
6. V. A. Nosenko et al., Sredstva,Svyazi, No. 4, 71 (1983).
7. V. A. Nosenko, Sredstva Svyazi, No. 1, 8 (1984).
8. V. M. Zaprudskii and V. M. Malyshev, Izmer. Tekh., No. 11, 18 (1987).
9. MI-187, MI-188, –É9. Admissible Errors in Testing of Measuring Instruments [in Russian], Izv. Standartov, Moscow (1980).
10. GOST 8.009-84. USSR State Standard. Normalized Metrological Characteristics of Measuring Instruments [in Russian].

THEORETICAL FOUNDATIONS OF GUARANTEED ERROR BOUNDS FOR THE SOLUTION OF METROLOGICAL PROBLEMS BY STATISTICAL METHODS

S. F. Levin and A. P. Blinov

UDC 389.14:519.241.2

A number of metrological functions (primary certification of high-precision measuring instruments, construction and prediction of long-term deviations of metrological characteristics given a small number of reference points, separation between systematic and random components of measurement error) involve certain experiments and tests, the processing of which relies on the solution of the so-called initial problems of mathematical statistics – the problems of structural, parametric, and composite identification. The specific feature of these identification problems is that we have no prior information about the characteristics of the random factors and it is fundamentally necessary to check the validity of the assumptions underlying the statistical methods proper. Correct solution of the initial problems of mathematical statistics is the source of that prior information which plays an exceptional role in metrology.

It is well known that probability theory is limited to the description of statistically stable random phenomena [1]. However, so far we do not know the quantitative boundaries that separate completely reproducible phenomena from statistically unstable phenomena [2, 3].

Analysis [4-7] shows that the prevailing situation in mathematical statistics may be termed paradoxical without exaggeration.

On the one hand, the axiomatic and frequency conceptions of probability provide a detailed probability-theoretical apparatus, for which the assumptions of applicability were formulated fairly definitely and were not questioned for a long time. Moreover, it seemed plausible to identify theoretical studies of various types of convergence with extension, by passage to the limit, of the properties of posterior statics of statistical data to their prior dynamics, and the differences between theory and experiment were regarded as a natural manifestation of randomness.

Translated from Izmeritel'naya Tekhnika, No. 12, pp. 5-8, December, 1988.

On the other hand, large-scale computerization of statistical methods has established, not so much that many problems of mathematical statistics are incorrectly specified [8], but that the principles of their solution are largely heuristic [9]. These "20th century myths" [10] include the preference for quadratic criteria of estimation accuracy [11], the universal belief in the increase of estimation accuracy with the increase of sample size [1], the validity of statistical inference using a single sample [12], the acceptance of the assignment of the significance level in hypothesis testing from practical considerations, etc. A special place on this list is occupied by the confidence probability of interval estimates [7]. Similarly to the phenomenon of actual inhomogeneity of data [13], actual guaranteed error bounds may deviate from the declared (e.g., Clopper–Pearson) confidence probabilities by as much as 25%, with maximum error attained for 0.80-0.95 confidence levels.

The essence of the paradox is that interpolation methods of data analysis have proved particularly efficient in this setting, and yet these methods are without complete theoretical foundations [14, 15].

The possible violation of the assumptions of applicability of statistical methods was recognized only in the early 1980s [16] in connection with the marked increase in measurement accuracy requirements with the development of modern technologies and in basic research. Definite progress has been achieved by this time in the development of the theoretical foundations of the solution of the initial problems of mathematical statistics [7, 12, 17-19].

We should specifically mention the cross-validation scheme, considered in different forms, e.g., in [13, 20-24]. This scheme is a passive variant of the empirical logic of statistical inference and after 1982 [25] it has been regarded as a promising direction for the development of statistical methods and, in most cases, solely as a tool for error bias reduction.

However, the available methods of dealing with statistical imhomogeneity of data do not resolve the more fundamental issues of the axiomatic and frequency conceptions of probability: the physical meaning of probability, a quantitative measure of statistical instability, and, on the applied level, the theoretical foundations of guaranteed error bounds in the solution of applied problems by statistical methods. More consistent and at the same time more radical in this respect is the interpolative conception of probability [26], whose basic assumptions can be summarized as follows.

First, the interpolative conception of probability presumes a strictly deterministic solution of the philosoophical problem of the relationship between dynamic and statistical regularities [3, 27], i.e., randomness is interpreted as unpredictability of deviations of the observed or measured characteristics of some processes from the corresponding models. In other words, there is no physical assumption of the existence of quantitative boundaries between completely reproducible, statistically stable, and statistically unstable phenomena. This and, apparently, the method of semantic differential, have stimulated the development of fuzziness theory [28].

Second, a fairly universal description of randomness [7] was obtained by resolving the sophism of controllability of experimental conditions [1] with the aid of the model of "exponential decay of ranked effects" [29]. The postulated interrelationship between controllability of processes and statistical stability of the corresponding factor models has become known as the controllability (observability) principle [26, 7].

Third, statistical stability is a property of the model, and not of the modeled data. A quantitative measure of statistical stability of a model is provided by the compactness of the distribution of the data around the model on prediction intervals (the compactness principle), as expressed by the compactness function [7]. The sample moments of this distribution have been used since the late 1960s [23], and in the 1980s so-called analysis of residuals has become particularly popular with the development of interactive computer systems [16].

Fourth, probability in the interpolative framework is defined as a calculated characteristic of the compactness function of the model: it is not introduced axiomatically, but rather derived from observations and satisfies the reproducibility condition.

Thus, in the interpolative framework, all mathematical models are regarded as approximate, and their accuracy depends not only on the degree of fit between the structure of the model and the process, but primarily on the attained accuracy level of the measurements during identification. Therefore, relatively large measurement errors encourage simplification of fairly complex models, even though they may be correct. This relationship between simplic-

ity and stability of models, making it possible to describe the joint effect of a multitude of factors by a limited number of parameters, is one of the motivations for the use of probability theory [30].

The interpolative conception of probability can be approximated [31] by Kolmogorov's axiomatic system [32]. To this end, the probability in the second axiom of [32] should be defined not as a nonnegative real number, but as a nonnegative real random variable with a corresponding probability distribution function in the ordinary sense or as a nonnegative real stochastic process. The theory of stochastic compactness developed on this base relies on the two-level representation [31] of the distribution density $\tilde{f}(x) = f(x) + E_f(x)$, where $f(x)$ is the ordinary distribution density of the random variable X and $E_f(x)$ is the realization of some random function such that

$$\int_{-\infty}^{+\infty} x^m E_f(x)dx|_{m=0}=0.$$

Sample analysis of these two relationships suggest a concrete form of allowing for the effect of the statistical stability of the model on the accuracy characteristics of the moment estimates for $m \geqslant 1$. Thus, for the quasiuniform probability distribution

$$f(x)=\begin{cases}1/(x_2-x_1), & x\in[x_1,x_2],\\ 0, & x\notin[x_1,x_2],\end{cases} \quad E_f(x)=\sin\omega x,\ \omega=\mathrm{var},$$

the possible deviations of the estimated mean MX and variance DX may exceed respectively 30% of the halfwidth of the distribution support and 60% of the variance for $Ef(x) = 0$. Such prior information, of course, is unavailable. Therefore, the initial problems of mathematical statistics, in addition to ordinary forms of estimation (point and interval), must employ a specific form of estimation – so-called compact estimation. Compact estimation may be interpreted as a modification of interval estimation, which allows for the effect of the degree of statistical stability of the models on the accuracy of the estimates.

Stochastic compactness theory in the interpolative framework plays the same role as Fisher estimation theory does in the axiomatic framework. There are differences, however.

The first difference is the consistent use of model cross selection, which constitutes the decision making procedure that relies on the compactness function of the competing models.

The second difference – preservation of lienar criteria – is traceable to the convections that persisted for more than half a century regarding the analytical properties of the absolute error [17] and were finally established by a theorem proved in 1983: if the probability distribution function $F_X(x)$ of a random variable X is such that $\lim_{x\to\infty} xF_X(x)=0$, then for an arbitrary finite estimtor θ we have

$$M|X-\theta|\equiv M(X-\theta)+2\int_{-\infty}^{\theta} F_X(x)dx. \tag{1}$$

An important corollary of the identity (1) is a minimal definition of the remarkable properties of the median, which is obtained by differentiation with respect to the parameter θ.

The third difference is based on the refinement of Kolmogorov type criteria in the problem of establishing equivalence of the random variables X_1 and X_2 by examining the extrema of the difference $D(x) = F_1(x) - F_2(x)$ of the corresponding distribution functions, as stated by the following theorem: if the distribution densities of two random variables satisfy the equality $f_1(x) = f_2(x) \neq 0$ on a sequence of points x_i, $i = \overline{1,n}$ then for $\inf X_1 < \inf X_2$ the equivalence measure is

$$\varkappa_{S2}\overset{\Delta}{=}\int_{-\infty}^{+\infty}\min_f\{f_1(x),\ f_2(x)\}dx\equiv 1-\sum_{i=1}^{n}(-1)^{i-1}D(x_i). \tag{2}$$

For $n = 1$, this notion of kappa-equivalence for sample distribution is expressible in terms of the Smirnov distribution parameter [33], $\varkappa_{S2}\equiv 1-D_{u,v}$.

These differences make it possible to increase the noise-tolerance of statistical estimation with inhomogeneous data and to eliminate the need for artificial introduction of the significance level in testing of hypotheses.

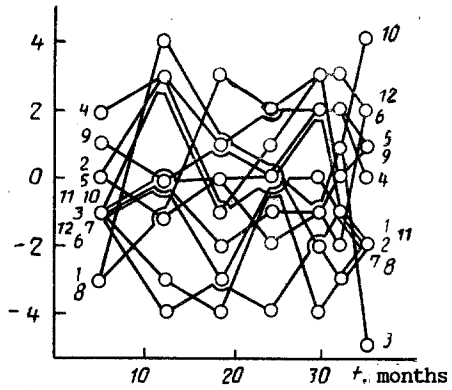

Fig. 1

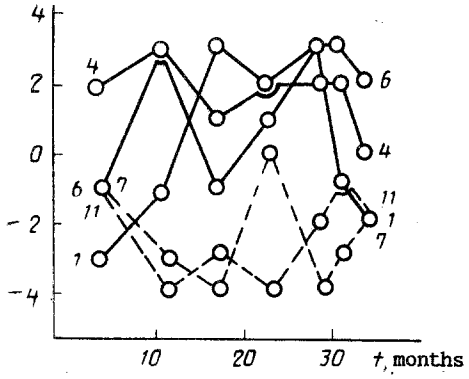

Fig. 2

The approximation by the axiomatic system of [32] was necessary in order to stress the main distinctive feature of stochastic compactness theory in application of metrological problems. The point is that, in general, the measurement result $\hat{\theta}_0$ is an unknown function $\Phi(X, \theta)$ of the set X of factors describing measurement conditions and the true value of the directly measured variable θ. In fact, instead of $\Phi(X, \theta)$ we only know an approximate model $\varphi(X_m, \theta)$, which allows only for part of the factors X_m, and instead of the true values of X_m, θ we only have the estimators $\hat{X}_m$, $\hat{\theta}$. Therefore the measurement error

$$\theta_0 - \hat{\theta}_0 = \Phi(X, \theta) - \varphi(X_m + \chi_m, \theta + \Xi_d)$$

contains three components: Ξ_d (the direct measurement error), χ_m (the error in the control of measurement conditions), and Ξ_M (the error of the model φ for $\Xi_d = \chi_m = 0$). The nature of the errors Ξ_M essentially depends on the structure of the model φ and the properties of the estimators of X_m, i.e., in the interpolative framework under conditions of prior uncertainty the above relationship suggests the following identification problem: it is required to find an analytical structure and parameter values for the model φ, which maximize the reproducibility of the distribution $f_{\Xi M}(\xi)$ by the empirical kappa-test (2).

In stochastic compactness theory, this is the standard formulation of the problem in the maximum compactness (MC) method [7], which in this case assumes consistent application of the stuctural and parametric identification procedures in one of the classes of interpotation polynomials defined by the maximum complexity model. The properties of the resulting solution depend on the properties of the MC-estimators used [7].

It is easy to note that the above problem aims at experimental determination of both the instrumental and the method components of the measurement error, as well as the random and systematic (including "ignored") components. Thus, for instance, the calibration characteristics of quartz temperature sensors were constructed [34] in the class of power series using median algorithms of MC identification. The average absolute error of the resulting MCMED-estimators by (1) was $1.18 \cdot 10^{-3}$°C compared with $0.68 \cdot 10^{-3}$°C for the ordinary least squares (OLS) estimators, i.e., roughly twice as high. However, control measurements show that the reproduction errors were $1.2 \cdot 10^{-3}$°C for MCMED-estimators against $31 \cdot 10^{-3}$ °C for OLS-estimators.

The efficiency of the algorithms of composite MC-identification was demonstrated for the case of analysis of statistically inhomogeneous data (see Fig. 1) represented by the deviation of a metrological characteristic (in relative units) of high-precision measuring instruments obtained, as observed in the context of determining the next certification time. The problem is characterized by a conflict between metrological guarantees and the high cost of individual assignment of the certification times in the framework of "operation by state." No compromise solution of this problem can be obtained with small samples using ordinary homogeneity testing procedures.

The algorithm PUMA using MCMED-models to fit a trend to the metrological characteristics was applied to classify 12 measuring instruments into five groups minimizing the overall reproduction error by the absolute compactness criterion [7]. The first group includes Nos. 1, 4, 6, the second No. 8, the third Nos. 7, 11, the fourth No. 3, and the fifth Nos. 2, 5, 9. 10. Figure 2 shows a realization of the first and the third (broken lines) groups.

High reproducibility or statistical stability of the solutions of standard problems of stochastic compactness theory is thus achieved at the cost of additional processing of data describing the real properties of partially controlled factors. This is a fundamental

point, because "probability theory is often viewed as a kind of a magic wand that produces information from total ignorance" [35]. At the same time, the frequency approach, based on the notion of limiting frequency as the number of trials goes to infinity, cannot be used to justify the application of probability theory in real life, where we invariably deal with a finite number of trials [6].

On the other hand, the interrelationship between dynamic and statistical regularities remains one of the most topical problems in the philosophy of natural sciences, focusing in the most acute form the fundamental conflicts between determinism and indeterminism on the gnoseological base. Nevertheless, past experience shows that real progress in measurement accuracy over the last few centuries has been achieved entirely through concrete consideration of progresively less significant factors.

LITERATURE CITED

1. V. N. Tutubalin, Probability Theory [in Russian], Mosk. Gos. Univ., Moscow (1972).
2. A. Kh. Shen', in: Semiotics and Informatics [in Russian], No. 18, VINITI Akad. Nauk SSSR, Moscow (1982), p. 14.
3. G. Ya. Myakishev, From Dynamics to Statistics [in Russian], Znanie, Moscow (1983).
4. A. N. Kolmogorov, in: Proc. 2nd All-Union Conf. on Math. Stat. [in Russian], Tashkent (1949), p. 240.
5. B. V. Gnedenko, A Course of Probability Theory [in Russian], Nauka, Moscow (1969).
6. A. N. Kolmogorov, in: Semiotics and Informatics [in Russian], No. 18, VINITI Akad. Nauk SSSR, Moscow (1982), p. 3.
7. Issues in Cybernetics [in Russian], No. 94, Akad. Nauk SSSR, Moscow (1982).
8. Methods of Mathematical Modeling, Computer-Aided Processing of Observations, and Their Applications [in Russian], Mosk. Gos. Univ., Moscow (1986).
9. N. Draper and H. Smith, Applied Regression Analysis, Wiley, New York (1966).
10. P. E. El'yasberg, Measurement Data: How Much is Needed? How to Process It? [in Russian], Nauka, Moscow (1983).
11. A. A. Ershov, Avtomat. Telemekh., No. 8, 66 (1978).
12. Yu. I. Alimov, An Alternative to Mathematical Statistics [in Russian], Znanie, Moscow (1980).
13. P. J. Huber, Robustness in Statistics [Russian translation], Mir, Moscow (1984).
14. V. N. Vinogradov, E. V. Gai, and N. S. Rabotnov, Analytical Approximation of Data in Nuclear and Neutron Physics [in Russian], Énergoatomizdat, Moscow (1987).
15. S. F. Levin, A Combination Method of Statistical Modeling [in Russian], Akad. Nauk SSSR, Moscow (1978).
16. I. Vuchkov, L. Boyadzhieva, and E. Solakov, Applied Linear Regression Analysis [Russian translation], Finan. State., Moscow (1987).
17. R. Runion, A Handbook of Nonparametric Statistics [Russian translation], Finan. Stat., Moscow (1982).
18. D. J. Hand, "Recent advances in error rate estimation," Pattern Recogn. Lett., 4, No. 5, 335 (1986).
19. F. Mosteller and J. Tukey, Data Analysis and Regression, Addison-Wesley, Reading, MA (1977).
20. M. Hollander and D. Wolfe, Nonparametric Statistical Methods, Wiley, New York (1973).
21. J. Van Rijsen, ed., Classification and Clustering [Russian translation], Mir, Moscow (1980).
22. V. Ya. Katkovnik, Nonparametric Identification and Smoothing of Data: Local Approximation Method [in Russian], Nauka, Moscow (1985).
23. A. G. Ivakhnenko, Avtomatika, No. 3, 58 (1983).
24. B. Efron, Nontraditional Methods of Multivariate Statistical Analysis [Russian translation], Finan. Stat., Moscow (1988).
25. J. C. Marron, "Will the art of smoothing ever become a science?" Contemp. Math., No. 59, 169 (1986).
26. A. Z. Ivanov and E. P. Nikitina, Inf. Mater. NSK Akad. Nauk SSSR, Kibernetika, No. 6 (121), 11-14 (1982).
27. Philosophical Dictionary [in Russian], Politizdat, Moscow (1987).
28. L. Zadeh, The Concept of Linguistic Variable and Its Application to Approximate Decision Making [Russian translation], Nauka, Moscow (1971).
29. V. V. Nalimov, Theory of Experiment [in Russian], Nauka, Moscow (1971).
30. S. F. Levin, Optimal Interpolation Filtering of Statistical Charcteristics of Random Functions of Statistical Characteristics of Random Functions in the Deterministic Ver-

sion of the Monte-Carlo Method and the Red Bias Law [in Russian], Akad. Nauk SSSR, Moscow (1980).
31. S. F. Levin, in: Statistical Theory of Signal Transmission and Reception [in Russian], Vol. 1, Akad. Nauk SSSR, Moscow (1981), p. 90.
32. A. N. Kolmogorov, Fundamental Concepts of Probability Theory [in Russian], Nauka, Moscow (1974).
33. D. Dugue, Theoretical and Applied Statistics [Russian translation], Nauka, Moscow (1972).
34. A. P. Blinov, Izmer. Tekh., No. 7, 15 (1987).
35. E. S. Ventsel', in: Mathematicians on Mathematics [in Russian], Znanie, Moscow (1982), p. 37.

METHOD OF CORRECTING ERRORS IN MEANS OF MEASUREMENT

V. S. Popov and E. V. Shumarov

UDC 681.2.088.6

There are various ways of correcting the systematic errors in means of measurement. If the instrument characteristics are nonlinear, one needs a multivalued standard-signal measure to implement any of them, for which purpose one usually employs a standard digital–analog converter (DAC). We consider correcting systematic errors by means of DAC in the feedback [1] and a new algorithm for calculating the corrected result, which provides elevated noise immunity and thus resolution. The highest noise suppression is obtained in an instrument employing digital processing with the instantaneous signal values.

Figure 1 shows the structural diagram for an instrument with automatic systematic-error correction. It contains the switch S, scale amplifier A, analog–digital converter ADC, computing unit CU, DAC, and voltage divider D. The working principle is that there is an additional correction cycle for each signal reading.

In the first cycle, the switch is in position 1 and the signal passes through the amplifier to the ADC. The control device (not shown in Fig. 1) causes the ADC to read the instantaneous value for the input signal u_i. The ADC output code N_i goes to the computer. In the correction cycle, the switch is in position 2, and N_i passes from the computer to the DAC. The transfer coefficient C_2 in the feedback circuit is taken as the reciprocal to the nominal value for C_1 in the forward circuit, i.e., $C_2 = 1/C_1$, so the forward circuit receives the feedback signal u_{fbi} close to u_i in the correction cycle. The difference $\Delta u_i = u_i - u_{fbi}$ defines the error in the forward circuit (it is assumed that the error in the feedback circuit is negligible). In the correction cycle, the ADC measures the amplified voltage u_{fbi}, and the corresponding code N_{fbi} also passes to the computer, where the corrected code N_{ic} is derived from

$$N_{ic}=2N_i-N_{fbi}. \tag{1}$$

The performance is dependent on the nonlinearity in the forward circuit and has been examined in detail in [1]. A disadvantage is the relatively low suppression factor for noise referred to the amplifier input (large additive error), which restricts the resolving power.

We estimate the noise immunity provided by (1), where we express N_i and N_{fbi} in terms of their components:

$$\begin{cases} N_i=C_1(u_i+\Delta_i+u_{ni}) \\ N_{fbi}=C_1(N_iC_2+\Delta_i'+u_{ni}'), \end{cases} \tag{2}$$

in which Δ_i and Δ_i' are the absolute errors referred to the amplifier input in the forward circuit due to the multiplicative error in the transfer coefficient correspondingly in the readings on the input signal and the feedback one, while u_{ni} and u_{ni}' are the noise values (additive errors) referred to the amplifier input and corresponding to the times when the readings are taken on the input signal and the feedback.

Noise induced in the external signals leads (communication lines) is here not eliminated, but this is not a source of additional error in measurement in alternating-current circuits because such instruments should measure the entire input, i.e., the sum of the

Translated from Izmeritel'naya Tekhnika, No. 12, pp. 8-10, December, 1988.

 0543-1972/88/3112-1150$12.50